Object-based saccadic selection during scene perception

Citation for published version:
Pajak, M & Nuthmann, A 2013, 'Object-based saccadic selection during scene perception: Evidence from viewing position effects' Journal of Vision, vol 13, no. 5, 2, pp. 1-21. DOI: 10.1167/13.5.2

Digital Object Identifier (DOI):
10.1167/13.5.2

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Publisher's PDF, also known as Version of record

Published In:
Journal of Vision

Object-based saccadic selection during scene perception: Evidence from viewing position effects

Maciej Pajak
Neuroinformatics Doctoral Training Centre, Institute for Adaptive and Neural Computation, School of Informatics, University of Edinburgh, UK

Antje Nuthmann
Psychology Department, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, UK

The goal of the present study was to further test the hypothesis that objects are important units of saccade targeting and, by inference, attentional selection in real-world scene perception. To this end, we investigated where people fixate within objects embedded in natural scenes. Previously, we reported a preferred viewing location (PVL) close to the center of objects (Nuthmann & Henderson, 2010). Here, we qualify this basic finding by showing that the PVL is affected by object size and the distance between the object and the previous fixation (i.e., launch site distance). Moreover, we examined how within-object fixation position affected subsequent eye-movement behavior on the object. Unexpectedly, there was no refixation optimal viewing position (OVP) effect for objects in scenes. Where viewers initially placed their eyes on an object did not affect the likelihood of refixating that object, suggesting that some refixations on objects in scenes are made for reasons other than insufficient visual information. A fixation-duration inverted-optimal viewing (IOVP) effect was found for large objects: Fixations located at object center were longer than those falling near the edges of an object. Collectively, these findings lend further support to the notion of object-based saccade targeting in scenes.

Introduction

During natural scene perception, we move our eyes about three times each second via rapid eye movements (saccades) so that the object of interest is centered on the high-resolution fovea. Accordingly, a large body of research on real-world scene perception and search has been addressed to the question of what determines where we attend and where we look in scenes (for reviews, see Henderson, 2003, 2011; Rayner, 2009). In particular, there is an ongoing debate about the relative influence of bottom-up, stimulus-driven factors and top-down, conceptually driven factors in determining fixation locations in scenes (e.g., Foulsham & Underwood, 2007; Henderson, Malcolm, & Schandl, 2009; see Schütz, Braun, & Gegenfurtner, 2011, for review).

Units of attentional selection in scene perception

A key question for models of eye guidance in complex scenes concerns the units of saccadic selection. As the visual world around us is full of objects, it can be construed that saccade target selection is driven by objects. In the scene perception literature, however, the dominant theoretical and computational framework to emerge has been image saliency, in which low-level properties of the stimulus play a crucial role in guiding attention and the eyes (see Tatler, Hayhoe, Land, & Ballard, 2011, for review). In those models, a saliency map is computed from basic image features such as luminance, color, and orientation (Itti & Koch, 2000, 2001; Itti, Koch, & Niebur, 1998; see also Torralba, Oliva, Castelhano, & Henderson, 2006). For each visual feature, a separate conspicuity map is computed. These individual maps are then combined in some manner to create a unified saliency map. The saliency map makes explicit the locations of the most visually distinct regions in the image. Attention is then allocated to those locations (in order of decreasing saliency), using a winner-takes-all principle.

Importantly, saliency works on low-level image features and has no knowledge about objects. Yet there is the possibility that saliency does not drive attention

Citation: Pajak, M., & Nuthmann, A. (2013). Object-based saccadic selection during scene perception: Evidence from viewing position effects. *Journal of Vision*, *13*(5):2, 1–21, http://www.journalofvision.org/content/13/5/2, doi:10.1167/13.5.2.

doi: 10.1167/13.5.2 Received October 31, 2012; published April 1, 2013 ISSN 1534-7362 © 2013 ARVO

and gaze directly but through its correlation with object properties (e.g., Henderson, Brockmole, Castelhano, & Mack, 2007). Surprisingly, very few studies have directly investigated whether objects can predict gaze better than low-level image features. Einhäuser, Spain, and Perona (2008) found that fixated locations were better predicted by object-level information than by image saliency. Using a receiver-operating characteristic analysis, objects predicted fixations with an accuracy of about 65%, whereas the predictive level of saliency was below 60% yet better than chance (50%). As a novel approach, Nuthmann and Henderson (2010) investigated landing positions[1] within fixated objects (and saliency proto-objects) to make inferences about attentional selection in natural scenes. The article reported data from a study in which participants viewed semantically rich color photographs of real-world scenes under three instruction sets: memorization, preference judgment, and visual search. The data were suggestive of a preferred viewing location (PVL) for real objects in scenes: Viewers preferred to send their eyes to the center of objects within naturalistic scenes, supporting the conclusion that the eye-movement control system directs the eyes in terms of object units during scene viewing. There were several other findings. First, the PVL was also present when restricting analyses to first fixations or single fixations in first-pass viewing, demonstrating that the PVL was not driven by refixations. Second, when landing position distributions were decomposed by saccade direction, we demonstrated that, in fact, saccades tended to undershoot the center of the object, consistent with behavior in reading (McConkie, Kerr, Reddix, & Zola, 1988). An additional set of analyses lent support to the notion that bottom-up saliency acts through its correlation with objects (cf. Einhäuser et al., 2008). Compared with the PVL for real objects, there was less evidence for a PVL for human fixations within saliency proto-objects—identified by an extension to the saliency map model (Itti et al., 1998) proposed by Walther and Koch (2006). There was no evidence for a PVL when only saliency proto-objects that did not spatially overlap with annotated real objects were analyzed (Nuthmann & Henderson, 2010).

The present article is an extension of our prior work reported in Nuthmann and Henderson (2010). Specifically, we aimed at collecting further evidence for our working hypothesis that objects are important units of saccade targeting and, by inference, attentional selection in real-world scene perception and search. We examined the effect of two further factors that Nuthmann and Henderson (2010) reasoned may modulate the PVL for objects in scenes: object size and launch site distance. Moreover, we investigated whether other landing position-related phenomena observed in reading generalize from words in sentences to objects in scenes, notably the refixation optimal viewing position (OVP) effect (e.g., McConkie, Kerr, Reddix, Zola, & Jacobs, 1989; Nuthmann, Engbert, & Kliegl, 2005) and the fixation-duration inverted-optimal viewing position (IOVP) effect (e.g., Nuthmann et al., 2005; Vitu, McConkie, Kerr, & O'Regan, 2001). The present work draws on a rich literature on the selection of words in reading, which is reviewed below.

Viewing-position effects in reading

In reading research, there are several well-documented findings related to landing positions in words. First, there are systematic tendencies with respect to where the eyes typically land within a word (PVL). In addition, landing position influences refixation probability (OVP effect) and fixation duration (IOVP effect).

Preferred viewing location

Distributions of landing positions within words resemble Gaussian distributions that are truncated at word boundaries. The mean of the distribution is the PVL, and it is typically found to be somewhat left of word center (Rayner, 1979; see also Rayner, Liversedge, Nuthmann, Kliegl, & Underwood, 2009). Word length has a moderate effect on the PVL: For longer words, the PVL moves somewhat to the left, toward a location between the beginning and the middle of a word (Rayner, 1979). Where readers fixate in a word can be viewed not only as a landing site for that word but also as the launch site for the next saccade. Launch site distance exhibits a strong influence on the mean (and standard deviation) of the Gaussian PVL (McConkie et al., 1988; Nuthmann et al., 2005; Rayner, Sereno, & Raney, 1996). If the launch site for a saccade landing on a target word is far from that word (e.g., seven letter spaces), the landing position will be shifted to the left. Likewise, if the distance is small (e.g., one letter space), the landing position is shifted to the right. Moreover, the standard deviation of the landing site distribution increases with the distance of the launch site from the target word.

It is generally argued that readers attempt to target the center of words as an optimal viewing location but that the eyes tend to deviate from that location because of two sources of error (McConkie et al., 1988). First, a systematic saccadic range error is responsible for the launch site contingent shifts in mean landing site. Saccadic eye movements are thought to be biased toward mean saccade amplitude, which causes systematic undershoot of far target locations and systematic overshoot of near target locations. Second, a random error component produces the spread in Gaussian landing position distributions. More recently, Engbert and Krügel (2010) suggested that the launch site effect is based on Bayesian estimation of saccade target positions.

Refixation probability OVP effect

The refixation probability OVP effect refers to the observation that the likelihood of making more than one fixation on a word before moving to another word (i.e., refixation probability) is lowest when the eyes initially fixate on the middle of the word (e.g., McConkie et al., 1989; Nuthmann et al., 2005; Vitu et al., 2001). As landing position moves further toward word boundaries, refixation probability increases. This rise is not linear but shows positive acceleration with distance from word center. The observation that refixation probability is minimal around word center is a piece of evidence for considering word center as the optimal viewing position in continuous reading. Although the functional dissociation of different subpopulations of refixation saccades is an ongoing research issue (cf. McDonald & Shillcock, 2004; Nuthmann & Engbert, 2009), it is undisputed that the occurrence of refixations is related to both visual acuity constraints and ongoing processing demands. First, some proportion of refixations results from failure to identify the word because of visual acuity limitations. Visual predictors involved in refixation planning include suboptimal initial landing positions within word, word length, and launch site distance (McConkie et al., 1989; Nuthmann, 2006). Initial landing position gives rise to the parabolic shape of the refixation probability curve, whereas word length and launch site distance modulate parameters of the refixation probability function. Second, there is a substantial independent influence of word frequency on the occurrence of refixations, suggesting that some proportion of refixations are due to word identification problems at higher processing levels (McConkie et al., 1989; Nuthmann, 2006).

Fixation-duration IOVP effect

In view of the refixation OVP effect, one could imagine that both refixation probabilities as well as fixation durations would be lowest when the eyes are located near the center of a word. Empirical reading data are, however, suggestive of an Inverted-OVP effect such that when the eyes are optimally placed at word center, fixation durations are longest rather than shortest (Hyönä & Bertram, 2011; McDonald, Carpenter, & Shillcock, 2005; Nuthmann et al., 2005; Vitu et al., 2001; White & Liversedge, 2006). The IOVP effect was first documented by Vitu and colleagues (2001), who subsequently examined various possible reasons for this seemingly counterintuitive effect (Vitu, Lancelin, & d'Unienville, 2007). They settled on the central argument that, for perceptual economy reasons, the oculomotor system plans on keeping the eyes longer at locations in words (e.g., the center of words) where a greater amount of information is anticipated. The authors reasoned that perceptual-economy principles may be universal and could be at work in any task that involves visual information intake (Vitu et al., 2007). Furthermore, they proposed a relationship between refixation OVP and fixation-duration IOVP effects. The *full perceptual-economy hypothesis* assumes that both OVP and IOVP effects result from perceptual-economy processes. The alternative *temporal perceptual-economy hypothesis* states that the OVP effect derives from the IOVP effect, which itself results from perceptual-economy processes. It assumes a functional relationship between the duration of the current fixation and the amplitude of the subsequent saccade (Yang & Vitu, 2007).

In contrast, the *mislocated fixation explanation* suggests that mislocated fixations are the primary source of the fixation-duration IOVP effect (Nuthmann et al., 2005; Nuthmann, Engbert, & Kliegl, 2007). Using modeling techniques, it was demonstrated that many fixations are mislocated such that they do not fall on the intended target word because of overshoots (for fixations falling on the beginning of words) and undershoots (for fixations falling at the end of a word) of the oculomotor system. It was proposed that the eyes respond to mislocated fixations with the start of a new, potentially error-correcting saccade program. Because mislocated fixations occur most frequently at word boundaries, this mechanism generated the typical inverted *u*-shaped pattern for fixation durations as a function of landing position. The IOVP model relies on low-level perceptual-oculomotor mechanisms unrelated to word recognition and was implemented and validated with the SWIFT model of saccade generation during reading (Engbert, Nuthmann, & Kliegl, 2007; Engbert, Nuthmann, Richter, & Kliegl, 2005). IOVP effects were also observed in *z*-string scanning, conceptualized as an oculomotor control condition to normal reading (Nuthmann & Engbert, 2009; Nuthmann et al., 2007). The *mislocated fixation* IOVP model qualitatively reproduced the strong IOVP effect in *z*-string scanning (Nuthmann et al., 2007), and so did the SWIFT model (Nuthmann & Engbert, 2009).

Generalizability of viewing-position effects

PVL, OVP, and IOVP effects were also observed for symbols in a reading-like sequential search task (Trukenbrod & Engbert, 2007). Notably, the viewing-position effects were found to generalize to simple object-viewing (Henderson, 1993) and object-search

(Foulsham & Underwood, 2009) tasks. For example, in the Henderson (1993) study, participants viewed rectangular arrays of four line drawings of objects in a prescribed sequence. Participants preferred to send their eyes to the center of objects. Moreover, landing positions were associated with both a refixation OVP effect as well as a fixation-duration IOVP effect. As objects rarely occur in isolation in the real world, it was important to show that PVL effects are also found for objects embedded in naturalistic scenes (Nuthmann & Henderson, 2010).

Present study

The goal of the present work was to further investigate the nature of object-based saccade targeting in naturalistic scene viewing. The reported analyses extend our prior work (Nuthmann & Henderson, 2010) in two important ways. First, we investigated the effect of two further factors that may modulate the PVL for objects in scenes: object size and launch site distance. It is possible that very small objects are not fixated optimally such that a clear PVL at object center may emerge only for larger objects. For objects that show a PVL, the effect of object size may parallel the effect of word length in reading: For larger objects, the incoming saccades may land closer to the edge (the edge that is closest to the launch site) of the object than for smaller objects. Moreover, if the selection of objects in scenes shares common mechanisms with the selection of words in reading, the distributions of within-object fixation positions should be modulated by launch site distance. We tested two launch site categories (near and far), contrasting objects targeted in central as opposed to peripheral vision. Near launch sites ($<5°$) represented instances in which the center of the object was situated in high-acuity central vision when a saccade was first directed to it. Far launch sites ($>5°$) corresponded to cases in which the object was located in low-acuity peripheral vision. If the launch site for a saccade landing on a target object is far from that object, the landing position should be shifted closer to the edge of the object. Alternatively, the identification of object boundaries may require foveal ($\sim1°$ eccentricity) and parafoveal ($\sim5°$ eccentricity) vision, at least for smaller objects. In that case, we should not find a PVL for far launch sites as preprocessing of the object in parafoveal vision is possible only for near launch sites. In the present study, the joint consideration of launch site and object size effects focused on horizontal saccades.

Second, we examined how within-object fixation position affected subsequent eye-movement behavior on the object, namely, whether landing position was associated with a refixation OVP effect and a fixation-duration IOVP effect. If objects in scenes are selected in a similar manner as words in reading, then the probability of a refixation on the same object will be the lowest in the center of the object, that is, in the middle of the horizontal axis and in the middle of the vertical axis. Moreover, we may find a systematic relationship between fixation location and fixation duration such that mean fixation durations on objects are longer when the eyes are near the center of the object (around the intersection of both main axes) than when the eyes are placed toward object boundaries (two-dimensional IOVP effect).

A large data set is needed to allow simultaneous data splitting by object size and launch site distance. Therefore, the sample size of the eye-tracking corpus was increased from 36 (Nuthmann & Henderson, 2010) to 72.

Methods

Participants, apparatus, and materials

The data from the 36 participants (mean age = 22.2 years, 12 men) reported in Nuthmann and Henderson (2010) were complemented by data from another 36 participants (mean age = 23.1 years, 22 men). Each participant viewed 135 unique full-color photographs of real-world scenes from a variety of scene categories. Scenes were presented for 8 s each on a 21-inch CRT monitor with a screen resolution of 800 × 600 pixels. Participants were seated 90 cm from the monitor, and a chin rest with a head support was used to minimize head movement. During scene presentation, participants' eye movements were recorded using an SR Research EyeLink 1000/2K system with high spatial resolution. The first batch of participants was tested with the monocular Tower Mount system, tracking the right eye at a monocular sampling rate of 1000 Hz. The second batch of participants was tested with the binocular Desktop Mount system, tracking both eyes at a binocular sampling rate of 1000 Hz. There is no difference in accuracy (0.25° to 0.5°) or spatial resolution (0.01°) between the Tower Mount and Desktop Mount systems. The experiment was implemented in SR Research Experiment Builder.

Gaze data analysis

Data from the right eye were analyzed. Raw data were converted into a fixation sequence matrix using SR Research Data Viewer. Fixations were excluded from analysis if they were preceded by or co-occurred with blinks, were the first or last fixation in a trial, or had durations less than 50 ms or longer than 1200 ms.

Procedure

The 135 scenes were divided into three blocks of 45 scenes. In each block, participants performed a different viewing task. Scenes were rotated through task and order across groups of participants. For the purpose of this article, only two of the three tasks were analyzed: Memorization and Preference Judgment.[2] In the Memorization task, the participants had to encode the scene to prepare for an old/new recognition test administered at the end of the experiment. In the Preference Judgment task, participants were asked to evaluate how much they liked each scene. For a full account of the procedure, see Nuthmann and Henderson (2010). To maximize data power, the pooled data from the memorization and preference judgment tasks were analyzed. Eye movements from the full 8-s presentation period were analyzed.

Object annotation

To analyze viewing-position effects for objects in scenes, an independent annotator who was naïve to the purpose of the research labeled objects in the corpus scenes. She did not tag the scenes exhaustively. Object selection criteria included choosing objects that were fully visible and not occluded by other scene elements and disregarding objects that were located at the center of the image (see Nuthmann & Henderson, 2010, for details). The position and spatial extent for each object were defined via a rectangular box (for an example scene and tagged objects, see figure 2a in Nuthmann & Henderson, 2010). Altogether, 730 objects were annotated ($M = 5.4$ objects per scene, $SD = 1.9$). Reflecting the real world, tagged objects varied in size. The average object width was 2.31° of visual angle ($SD =$ 1.24°, range: 0.61°–8.86°), and the average object height was 2.64° ($SD = 1.35°$, range: 0.64°–10.22°). The shape of annotated objects is described by their aspect ratio. A width:height aspect ratio of 1 is indicative of squared objects, whereas a smaller (larger) ratio describes objects that are vertically (horizontally) elongated. For the aspect ratios of corpus objects, the first, second, and third quartiles were 0.63, 0.85, and 1.20, respectively (range of aspect ratios: 0.25–3.77).

Normalized landing positions within objects

A given fixation location FL has a horizontal (x) and a vertical (y) component. The eye tracker provided fixation locations in screen coordinates. For the purpose of the present analyses, the original fixation sequence matrix was reduced to valid fixations that were placed on objects. The locations of these fixations were recalculated as within-object landing positions: The horizontal component was calculated relative to the left border of the object box and the vertical component relative to the upper border of the object (for visualization, see figure 1d in Nuthmann & Henderson, 2010). To allow data averaging across objects of different sizes, the two-dimensional landing positions were normalized to fit into a unit square (1 × 1; Nuthmann & Henderson, 2010). Specifically, the horizontal landing position axis was standardized by dividing the horizontal landing position component by the width of the object box, leading to landing positions ranging between 0 and 1. Likewise, vertical landing positions were normalized by dividing them by the height of the object, again leading to landing positions ranging between 0 and 1. Accordingly, a two-dimensional within-object fixation location $FL_{xy} = [0.5, 0.5]$ refers to the center of the object.

Results and discussion

The results and their discussion are presented in four main sections. The first two sections report effects of object size and launch site distance on the PVL for objects in scenes. The third section investigates the relationship between the initial landing position on an object and the probability of refixating the object. The final section explores the relationship between landing position and fixation duration.

Influence of object size on the PVL

To explore the effect of object size, three object-size categories were created based on the empirical distribution of object sizes in the corpus data. For tagged objects, there was a positive correlation between object width and object height ($r = 0.53$, $p < 0.001$). Object width and height followed similar distributions. The two distributions were therefore merged. We determined the two points that divided the distribution into three parts, each containing a third of the objects. However, large objects received considerably more fixations than did medium-sized or small ones (Figure 1). Therefore, the terciles were slightly adjusted to ensure comparable data power across object-size categories. The resulting two points that created the three object size categories were 1.85° and 3.15°. For the reported analyses, objects were selected such that both their width and their height fell into the same size category. This does not imply that

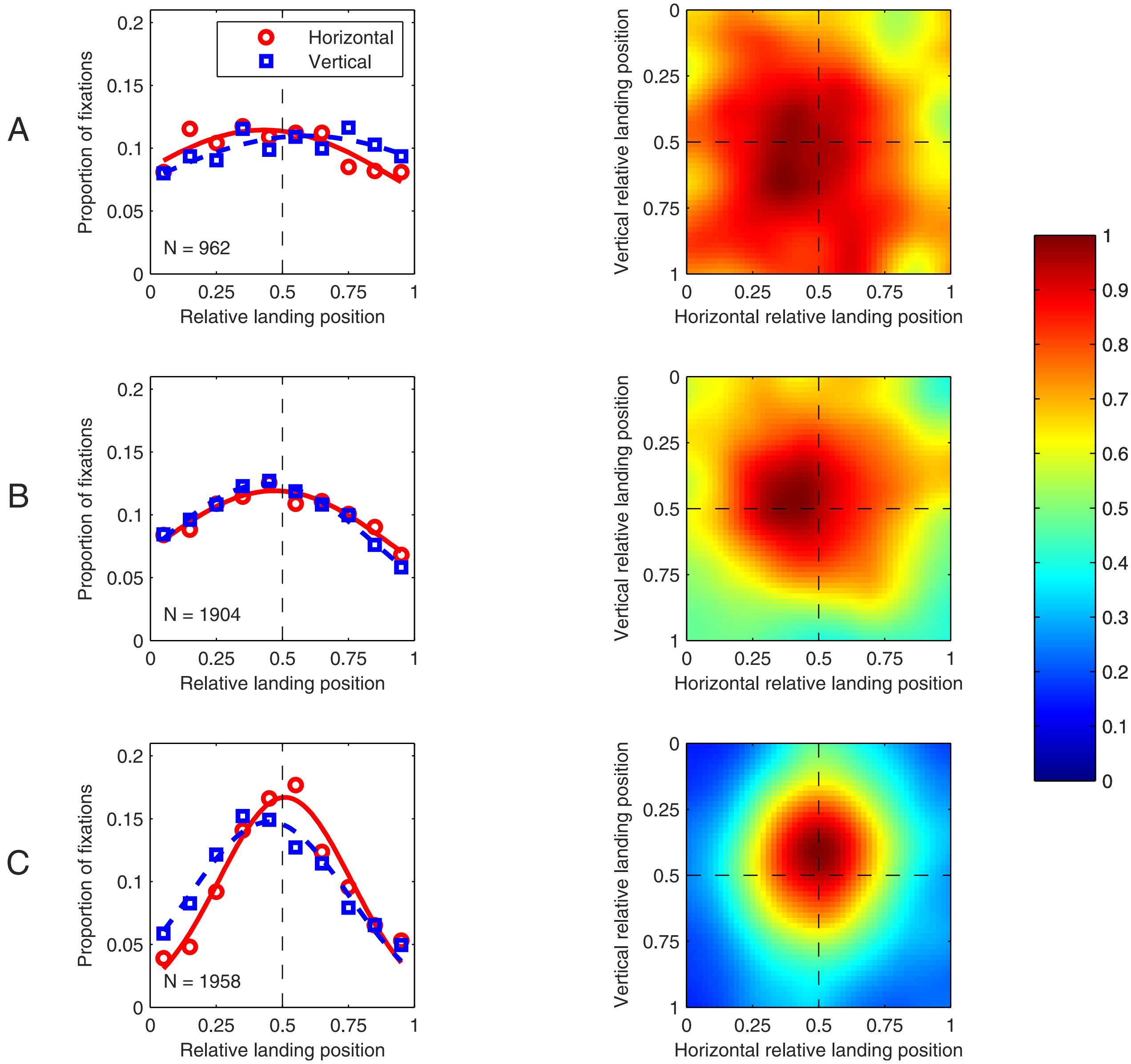

Figure 1. Effect of object size on the preferred viewing location for objects in scenes. Distributions of fixation locations within objects are displayed for small (A), medium-sized (B), and large (C) objects. (Left) Distributions for the horizontal (red circles) and vertical (blue squares) components of landing positions within objects. Experimentally observed distributions were fitted using truncated Gaussians. The vertical broken line marks a landing position at the center of the object. (Right) Corresponding smoothed two-dimensional viewing location histograms. The legend represents a relative scale with one corresponding to the maximum fixation probability observed for a given object size. The intersection of the two broken lines marks the center of the object.

objects within a given category were necessarily square shaped (see Table 1 and Supplementary Figure S1). Rather, this procedure eliminated objects with more extreme aspect ratios, where object width and height fell into different categories (but see Figure 9). Furthermore, objects with a width or height smaller than 1° were disregarded ($n = 67$) to accommodate the inaccuracy of the eye tracker as well as the inaccuracy of the visual system. The size of large objects was capped at 6° to exclude objects that covered a large proportion of the scene ($n = 29$). As a result of this selection procedure, the large object category comprised relatively fewer objects ($n = 64$) than the small ($n = 119$) or medium category ($n = 117$). To summarize, for small objects, both object width and height were smaller than 1.85°. The width and height of medium-sized objects ranged between 1.85° and 3.15°. Large objects had an object width and height greater than 3.15°. The Euclidean distance between object center and image center was similar across object size categories (Table 1). Therefore, observed effects cannot be accounted for by the placement of objects relative to the center of the scene.

Object size	Small (<1.85°)	Medium (1.85°–3.15°)	Large (>3.15°)
n	119	117	64
Width	$M = 1.38°$ ($SD = 0.24°$)	$M = 2.39°$ ($SD = 0.36°$)	$M = 4.17°$ ($SD = 0.74°$)
Height	$M = 1.47°$ ($SD = 0.23°$)	$M = 2.45°$ ($SD = 0.36°$)	$M = 4.29°$ ($SD = 0.76°$)
Ratio of object width to object height	$Q_1 = 0.78$, $Q_2 = 0.91$, $Q_3 = 1.12$	$Q_1 = 0.84$, $Q_2 = 0.98$, $Q_3 = 1.14$	$Q_1 = 0.78$, $Q_2 = 1.00$, $Q_3 = 1.15$
Euclidean distance between object center and image center	$Q_1 = 5.48°$, $Q_2 = 8.36°$, $Q_3 = 10.98°$	$Q_1 = 5.77°$, $Q_2 = 7.29°$, $Q_3 = 9.79°$	$Q_1 = 5.53°$, $Q_2 = 8.09°$, $Q_3 = 10.20°$

Table 1. Description of the three object-size categories used in the study. *Note*: M denotes mean; SD, standard deviation; Q_1, first quartile; Q_2, second quartile (median); and Q_3, third quartile.

In reading, landing-position distributions are typically based on initial fixations, defined as the first fixation on a word regardless of whether it is the only fixation on a word or the first of multiple fixations on a word (e.g., McConkie et al., 1988; Nuthmann et al., 2005). Accordingly, the present analyses considered first fixations in first-pass viewing, that is, saccades that were launched from outside a given object and led to a within-object fixation, irrespective of whether it was followed by an immediate refixation or not.

Figure 1 displays the distributions of fixation locations within objects for small (Figure 1A), medium sized (Figure 1B), and large (Figure 1C) objects. The horizontal and vertical components of within-object fixation locations were examined separately (left panels) and jointly (right panels). In a first step, for objects of a given size, horizontal and vertical components of landing positions were analyzed separately. Distributions were calculated based on 10 bins of equal size. The results demonstrate that the distribution of fixation locations within objects is modulated by object size. For small objects, both horizontal and vertical landing positions more or less followed a uniform distribution, suggesting that there was no evidence for a PVL (Figure 1A, left panel). For medium-sized objects, within-object landing position distributions for both horizontal and vertical components were Gaussian in shape and peaked around the center of the object (Figure 1B, left panel). Thus, the data are suggestive of a PVL such that viewers preferred to send their eyes to the center of objects. This preference for object centers over object ends was even more pronounced for large objects (Figure 1C, left panel).

In a second step, distributional analyses were performed in two-dimensional space. For the data that went into the analyses presented in the left panels of Figure 1, the right panels show the corresponding two-dimensional fixation location histograms. Those histograms were created using narrow bins and additional smoothing. In either direction (horizontal and vertical), 50 histogram bins were used. The smoothing of the two-dimensional histograms was based on an algorithm by Eilers and Goeman (2004), with the smoothing parameter λ set to 10 (values close to zero lead to a plot that is essentially just the raw data, and higher values lead to more smoothing; see Eilers & Goeman, 2004).

The frequency information is displayed as variations in color (Figure 1, right panels), with colors ranging from blue (few fixations) to red (many fixations) and passing through cyan, yellow, and orange. For all three object-size categories, more fixations were located around object center ($FL_{xy} = [0.5\ 0.5]$) than at object boundaries. For small objects, there was only a slight tendency for such central clustering. However, a central "hot spot" emerged for medium-sized objects and was clearly defined for large objects. The results thus confirm the separate analyses of horizontal and vertical landing position components.

For the data from each object-size category, repeated-measures analyses of variance (ANOVAs) were conducted with the specific aim of testing for any systematic effect of horizontal and/or vertical relative landing zone on fixation probability. Polynomial trends of fixation probabilities over fixation position inform us about the shape of the landing position curve. Negative quadratic trends (i.e., convex curvature) indicate a preference for the center of objects. Linear trends indicate a significant preference for targeting either end of the object. The results of the ANOVAs are presented in Table 2 and can be summarized as follows: For small objects, there was no systematic effect of horizontal or vertical relative landing zone on fixation probability. For medium-sized and large objects, the probability of fixations differed between landing zones; this was true for both horizontal and vertical landing positions. In addition, there were significant quadratic trends, indicating a PVL at the center of objects (Table 2). A comparison of effect sizes revealed that the quadratic trend was stronger for large as compared with medium-sized objects, supporting that the preference for fixating object centers over object ends was stronger for large objects.

To summarize, fixation location distributions for medium-sized and large objects showed evidence for a PVL. To better describe the PVL, the experimentally

	Horizontal		Vertical	
	F	η^2	F	η^2
Small objects				
Overall effect	1.6	0.022	0.7	0.010
Quadratic trend	6.93*	0.089	0.59	0.008
Linear trend	1.99	0.27	2.83	0.038
Medium-sized objects				
Overall effect	4.78***	0.063	6.83***	0.088
Quadratic trend	29.2***	0.291	68.36***	0.491
Linear trend	1.37	0.019	9.61**	0.119
Large objects				
Overall effect	35.61***	0.334	18.74***	0.209
Quadratic trend	201.34***	0.739	113.76***	0.616
Linear trend	0.755	0.011	11.19**	0.136

Table 2. Effect of horizontal and vertical landing position on fixation probability within objects of different sizes. *Note*: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

observed distributions were fitted using truncated Gaussians. To obtain estimates for the mean (μ) and standard deviation (σ) of the horizontal and vertical landing position distributions, a grid search method (in steps of 0.01) with a minimum χ^2 criterion was used (cf. Nuthmann et al., 2005; Nuthmann & Henderson, 2010). Best-fitting lines are shown in Figure 1. Mean and standard deviation of the fitted normal curve characterize the PVL curve. The actual PVL is indexed by the mean of the fitted normal curve. How much the eyes deviate from the PVL is indexed by the standard deviation of the Gaussian distribution. Two important points are clear from the fit parameters (medium-sized objects: horizontal $PVL_{M,SD} = [0.47, 0.47]$, vertical $PVL_{M,SD} = [0.44, 0.42]$, $n = 1904$; large objects: horizontal $PVL_{M,SD} = [0.51, 0.25]$, vertical $PVL_{M,SD} = [0.45, 0.30]$, $n = 1958$). First, they confirm that the PVL is close to object center for both object-size categories. However, this finding cannot be taken to suggest that object size has no effect on where the PVL is situated. Previous analyses, collapsing data across object-size categories, demonstrated that the global PVL distribution with a peak at the center of the object resulted from the summation of distributions dependent on the orientation of the incoming saccade (Nuthmann & Henderson, 2010). Specifically, it was shown that saccades coming from the left, right, top, or bottom tended to undershoot the center of the object on the axis of the movement. Analyses in the next section will take the direction of the incoming saccade into account, thereby allowing one to assess whether object size affects where the PVL is situated. Second, the fit parameters explicate that the distributions for the horizontal and vertical components of landing positions showed a smaller spread for large objects than for medium-sized objects.

Influence of launch site distance and object width on the PVL

Center-based launch site distance is the distance between the launch site of the last saccade and the center of the target word (in reading) or object (in scene viewing). In reading, the eyes tend to move from left to right through a line of text. In comparison, the direction of eye movements people make when viewing or searching a natural scene is less predictable. For analyses, incoming saccades were categorized according to the angle between the horizontal plane and the vector connecting the center of the object and the launch site (Figure A1). Nuthmann and Henderson (2010) showed that viewers' eyes preferred to enter objects along the horizontal axis, that is, from the left or from the right. Therefore, the present analyses focused on horizontal landing positions, considering instances in which the eyes entered the object from the left or from the right. Specifically, object boxes were divided into two directional segments, each subtending 45° (toward either side) around the two horizontal axes connecting to the center of the object (left, right). For example, saccades launched from outside a given object and leading to a within-object fixation were classified as entering the object from the right if the angle between the launch site and the center of the object was between –45° and +45° (Figure A1). For the purpose of the present analyses, the data were collapsed across the two horizontal categories. For instances in which the eyes entered the object from the right, the relative horizontal fixation location was reverse coded. Specifically, the fixation location F_x was recalculated as $F'_x := 1 - F_x$. For objects of a given size, a paired two-sample t test confirmed that there was no significant difference in relative horizontal fixation location for cases in which the eyes entered the object from the left or right, respectively (all p's > 0.05).

Analyses are reported for the object-size categories that showed a PVL, that is, medium-sized and large objects (Figure 1). For medium-sized objects, data were divided into two groups, depending on the launch site of the saccade leading into the object. Near ($<5°$) and far ($>5°$) launch sites contrasted objects targeted in central as opposed to peripheral vision. The results revealed an interesting dissociation. Although there was a horizontal PVL for near launch sites (Figure 2a), there was no PVL for far launch sites (Figure 2b), suggesting that medium-sized objects were too small to perceive their boundaries without the high acuity present in central vision. For near launch sites, the probability of fixations differed between landing zones, $F(7.10, 500.72) = 4.36$, $p < 0.001$. The significant quadratic trend, $F(1, 71) = 30.46$, $p < 0.001$, captured the convex shape indicative of a preference for targeting the object center. Although the overall F test

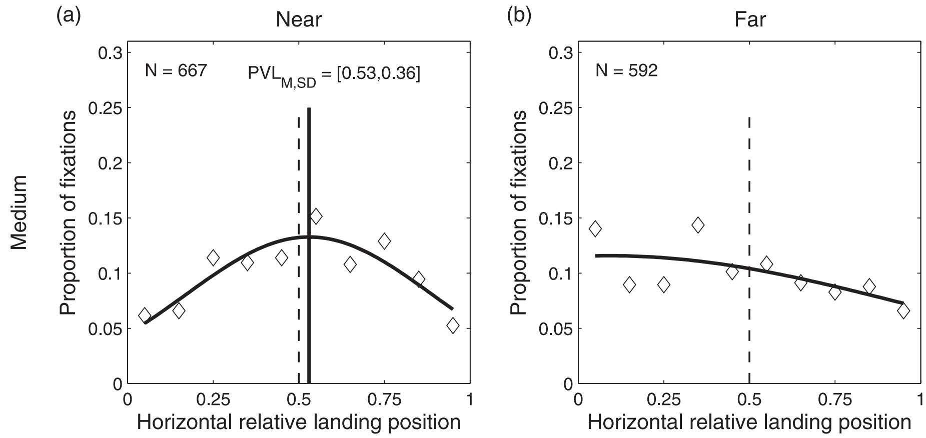

Figure 2. Effect of launch site distance on the preferred viewing location for medium-sized objects in scenes. For near (a) and far (b) launch sites, distributions of horizontal landing positions are shown, based on instances in which the eyes entered the object along the horizontal axis, that is, from the left or from the right (the data were collapsed across the two groups). In the left panel, the best-fitting normal curve is additionally presented. The mean of the fitted curve, depicted by the vertical solid line, denotes where the PVL is situated. In both panels, vertical broken lines mark a landing position at the center of the object.

was also significant for far launch sites, $F(7.31, 518.81) = 3.37$, $p = 0.001$, trend analyses identified a significant linear trend, $F(1, 71) = 17.36$, $p < 0.001$, but, importantly, no quadratic trend, $F(1, 71) = 0.89$, $p = 0.35$.

Such dissociation was not present for large objects as a horizontal PVL was observed for both near and far launch sites. To assess the effects of launch site distance and object width on the PVL for these objects simultaneously, the data were split according to a 2 × 2 analysis scheme. Based on the median launch site distance observed in the data, two launch site categories were defined: For near/far launch sites, the Euclidean distance between the location of the previous fixation (outside of the object) and the center of the object was shorter/longer than the median launch site distance of 5.7° (which is close to the 5° central-vision criterion). To study the effect of object width, large objects (3.15°–6°, see above) were further divided into two subcategories of "smaller large" objects with a width smaller than 4.5° ($n = 44$) and "larger large" objects with a width exceeding 4.5° ($n = 20$). The resulting decomposed distributions of horizontal landing positions are depicted in Figure 3. First, for far launch sites, the distribution of landing positions was shifted to the left, toward the edge closest to the launch site (compare panels column-wise). Second, a similar leftward shift was observed for larger objects as compared with smaller objects (compare panels row-wise). A two-way repeated-measures ANOVA was conducted to test for any systematic effects of launch site distance and object width on mean relative horizontal landing position. There were significant main effects of launch site distance, $F(1,62) = 16.63$, $p < 0.001$, and object width, $F(1, 62) = 8.37$, $p = 0.005$. The interaction between the two variables was not significant, $F(1, 62) = 2.63$, $p = 0.110$.

Influence of initial fixation location on refixation probability

The next set of analyses examined the relationship between the position of the initial fixation on an object and the likelihood of making another immediate fixation within the same object. If objects in scenes are refixated in a similar manner as words in texts, refixations should be least likely if the first fixation is optimally placed close to object center. As first-fixation position moves further toward object boundaries, refixation probability should increase. Refixations initiated from suboptimal locations near object boundaries should bring the eyes closer to object center (see Figure 5 for visualization of this hypothesis).

Refixation probability was found to depend on object size such that refixations were more likely for large objects (23%) compared with medium-sized objects (15%) and small objects (15%). Of main interest was whether this probability was modulated by initial landing position. The focus was on large objects, but qualitatively similar results were obtained for small and medium-sized objects. There were 580 refixation cases on large tagged objects. In 79% of these instances, the eyes made exactly two successive fixations on the object ($n = 457$). Analyses considered those two-fixation cases. The effect of initial fixation location on refixation probability was examined separately for horizontal and vertical initial landing positions (Figure 4a), based on seven landing-position bins of equal size. Refixation probabilities considered

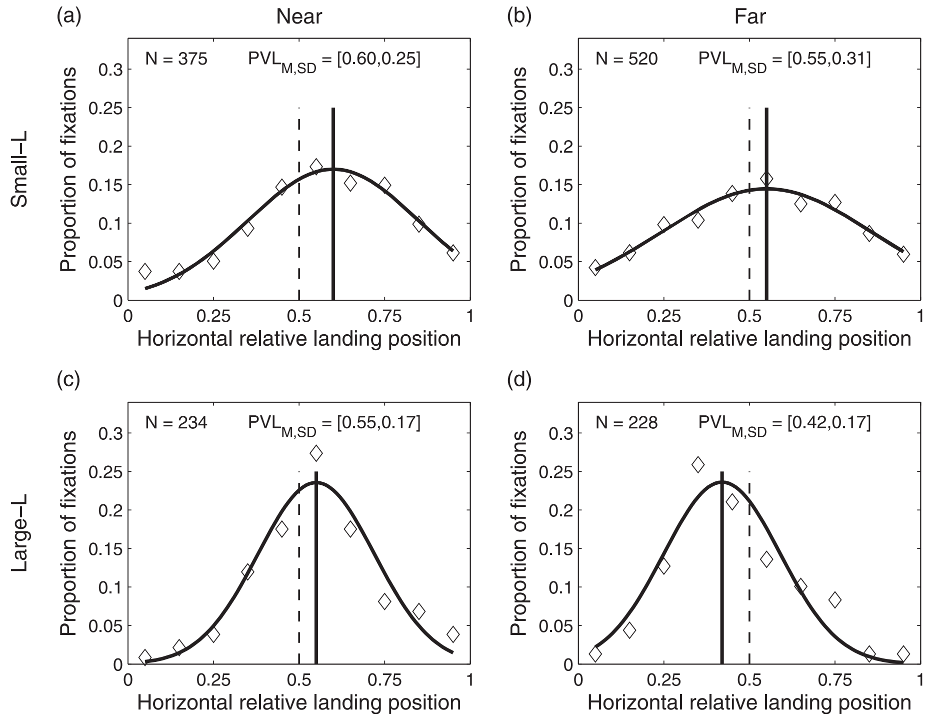

Figure 3. Distributions of horizontal landing positions for two different object sizes (rows) and near and far launch site distances (columns). Analyses considered large objects, which for the purpose of this analysis were split into two subcategories of "smaller large" and "larger large" objects. Also presented is the best-fitting normal curve for each distribution. Mean and standard deviation of the fitted normal curve characterize the PVL curve. The mean of the fitted Gaussian, highlighted by the vertical solid line, denotes where the PVL is situated. For comparison, vertical broken lines mark a landing position at the center of the object.

the proportion of initial fixations at different landing positions that were immediately followed by a refixation on the object. Regardless of initial horizontal or vertical landing position, the probability of making an immediate refixation on the object was about 0.25 (Figure 4a). Repeated-measures ANOVAs with relative landing zone as a within-subject factor confirmed that there was no systematic effect of horizontal, $F(5.20, 368.89) = 0.29$, $p = 0.92$, or vertical, $F(4.62, 327.99) = 0.55$, $p = 0.73$, initial landing position on refixation probability. In addition, refixation probability was examined in two-dimensional space (Figure 4b). The normalized object was divided into a 4×4 grid, resulting in 16 same-sized initial landing position cells. For each cell, the average refixation probability was calculated. For the purpose of visualization, the data

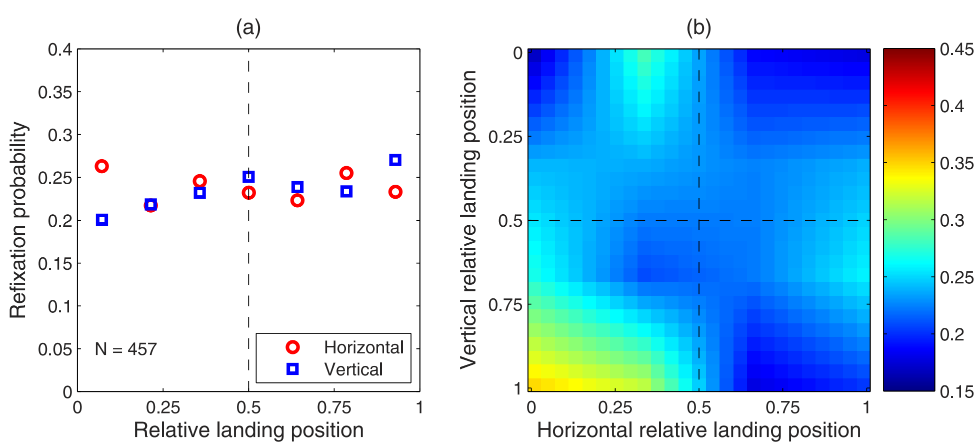

Figure 4. The effect of initial fixation location on refixation probability for large objects. (a) Mean refixation probability as a function of horizontal (red circles) and vertical (blue squares) initial landing position. (b) Corresponding two-dimensional distribution of refixation probability.

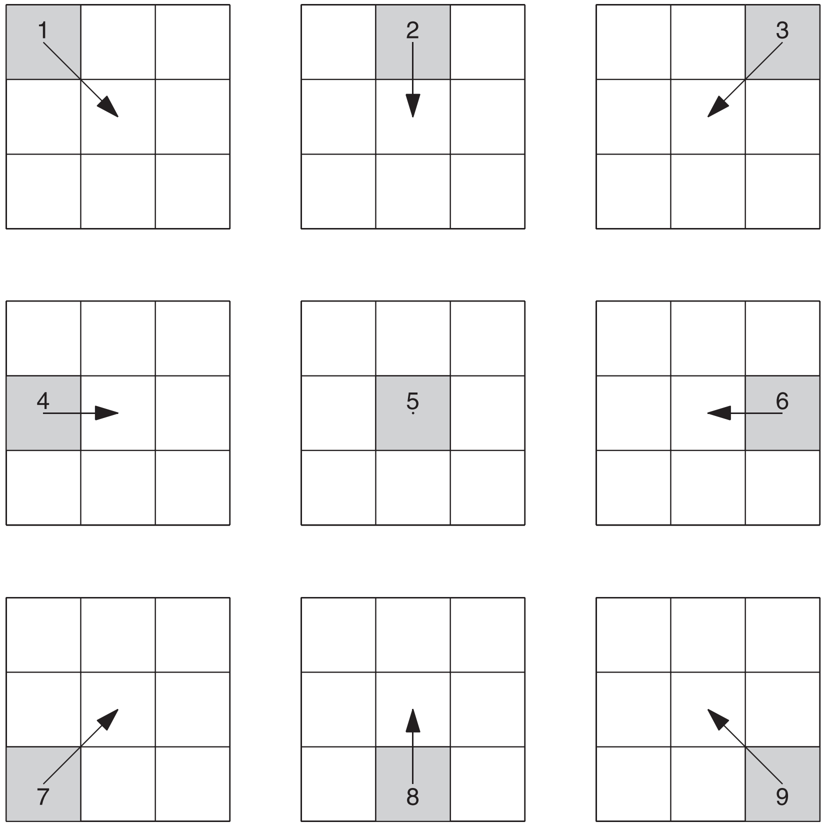

Figure 5. To analyze spatial properties of refixation saccades, the normalized object was divided into a 3 × 3 grid, resulting in nine same-sized launch cells. Each of the nine figure panels displays the grid, with the current cell numbered and colored in light gray. The arrow lines depict hypotheses regarding the direction of refixation saccades launched from a given grid cell. See text for details.

were extrapolated and smoothed with a simple bilinear filter. Again, initial fixation position within the object does not appear to influence refixation probability in a systematic manner (Figure 4b).

Drawing on the analogy between objects in scenes and words in texts, word shapes can be described as horizontally elongated rectangles. We therefore tested whether a refixation probability OVP effect would emerge for horizontally elongated objects. These additional analyses considered objects with a width larger than 3.15° and a width:height aspect ratio greater than 1.5 ($n = 41$). For those objects, refixation probability remained unrelated to the position of the initial fixation along the horizontal axis, $F(3.17, 44.38) = 1.74$, $p = 0.17$.

As this was an unexpected finding, further analyses were conducted to explore the origin of the flat refixation probability functions. First fixation (FF) position corresponds to the launch site of the refixation saccade, whereas second fixation (SF) position describes the landing site of the refixation saccade. Landing positions for both first fixations (as part of Figure 1c) as well as second fixations (Figure 7, top panel) cluster around object center. To understand the eyes' refixation behavior on objects in scenes, we must illuminate the relationship between PVL_{FF} and PVL_{SF} on the one hand and the flat refixation probability functions on the other hand. To analyze spatial and directional properties of refixation saccades, the normalized object was divided into a 3 × 3 grid, resulting in nine same-sized launch cells (Figure 5). Each first fixation was assigned to one of the nine launch cells, according to the fixation's within-object position. Then, for a given launch cell, the directional distribution of refixation saccades was computed and visualized with a polar plot (Figure 6). Probability densities were computed from 12 bins. For the eight noncentral launch cells, the results adhere to the prediction that refixation saccades were directed away from object boundaries toward a more central viewing position (compare distributions in Figure 6 with arrow lines in Figure 5). In contrast, refixations launched from a central position went in all directions, with a slight directional bias toward upwards-directed refixations. Although those data are suggestive, they only speak to the launch site contingent direction of refixation saccades and provide no information about their landing sites.

Do refixation saccades bring the eyes closer to object center, or to a location beyond object center? To

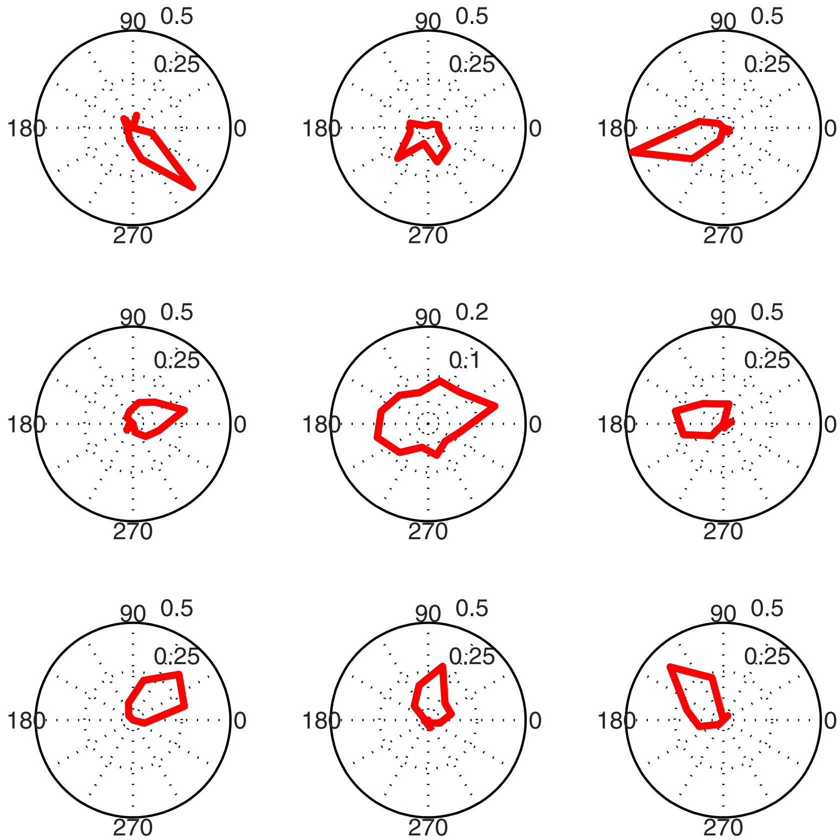

Figure 6. Directional distribution of refixation saccades. The angular orientation of refixation saccades was calculated contingent on the launch site of the refixation saccades within the 3 × 3 grid (Figure 5). For example, the bottom-left panel displays the results for refixation saccades that were launched from the bottom-left corner of objects. Analyses considered large objects.

answer this question, a complementary analysis examined partial landing position distributions for second fixations, again contingent on the refixation saccade's launch cell in the grid. Accordingly, for a given launch cell, Figure 7 displays the corresponding two-dimensional second fixation (refixation) landing position distribution. For example, the bottom-left panel (panel 7 in Figure 5) displays the results for refixation saccades that were launched from the bottom-left corner of objects. For most, but not all, of the eight noncentral launch cells, the data show a hot spot within or close to the central cell representing the area around object center. This confirms that refixations launched close to object boundaries have a tendency to correct the suboptimal initial viewing location. Interestingly, for refixations that were launched from a central position, the results also reveal a central hot spot, suggesting that a lot of those refixations kept the eyes relatively close to object center. Pooling of the nine launch cell contingent partial landing position distributions leads to the composite distribution displayed in the top panel of Figure 7.

Taken together, in keeping with our predictions we observed corrective refixations elicited from initial viewing positions close to object boundaries. However, in one out of four cases, viewers also refixated the object when the initial fixation was placed close to object center, which is thought to be optimal for object processing. The majority of those refixations kept the eyes in proximity to object center. Those two types of refixations contributed to the observed flat refixation probability function.

Influence of within-object fixation location on fixation duration: IOVP effect

A final set of analyses investigated the effect of within-object fixation location on fixation duration. In agreement with the PVL analyses, first fixations were analyzed. For consistency with the reading literature (Nuthmann et al., 2005; Vitu et al., 2001), additional analyses considered single fixations. Single fixations are fixations on objects that were fixated exactly once during first-pass viewing. The results for small and medium-sized objects can be summarized such that fixation duration was not modulated by within-object fixation location. However, IOVP effects were observed for large objects, and those are described below. First, effects of horizontal and vertical landing position were

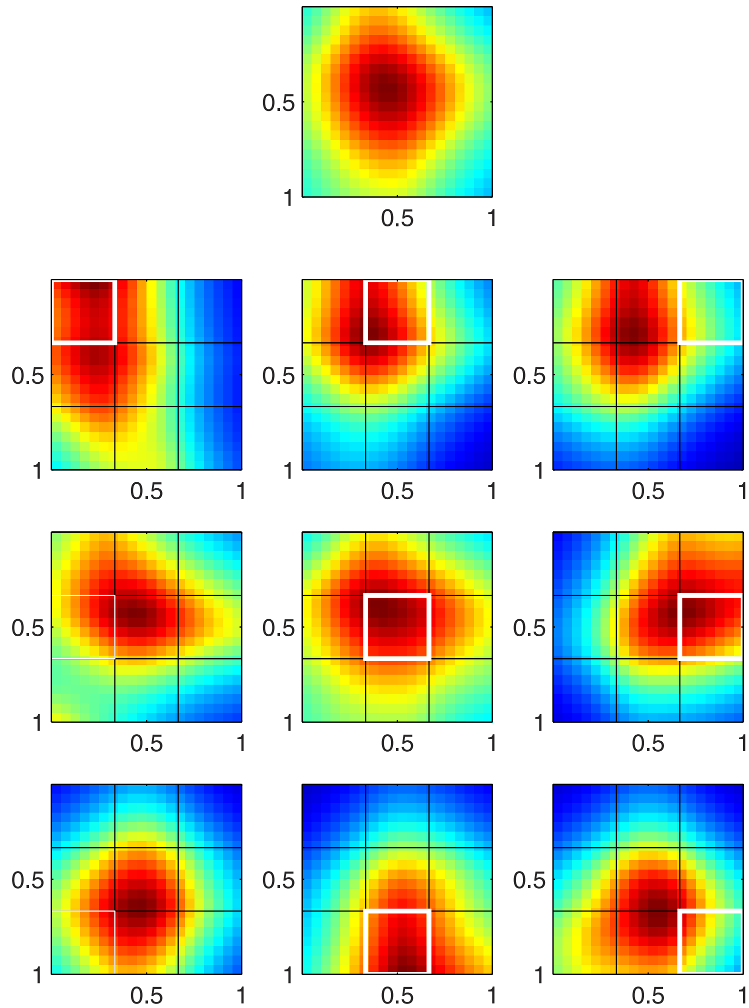

Figure 7. Partial landing position distributions for second fixations in refixation cases. The nine panels forming a square depict smoothed two-dimensional landing position distributions for second fixations in refixation cases, each contingent on the refixation saccade's launch cell in the 3 × 3 grid (Figure 5). The respective launch cell is highlighted in white. The top panel displays the composite two-dimensional landing position distribution for second fixations across launch sites. Analyses considered large objects.

analyzed separately, based on seven landing-position bins of equal size. For both first and single fixations, fixation durations were longer when the eyes landed in the center of the object than when they landed near object boundaries (left panels in Figure 8). In addition, the effect of landing position on fixation duration was examined in two-dimensional space (right panels in Figure 8). The normalized object was divided into a 4 × 4 grid, and the corresponding mean fixation durations were extrapolated and smoothed with a simple bilinear filter. The duration of fixations is displayed as variations in color, with colors ranging from blue (short fixation durations) to red (long fixation durations) and passing through cyan, yellow, and orange. For both first and single fixations, fixation locations at object center or in proximity to object center were associated with the longest fixation durations. The closer fixations were placed toward object boundaries, the shorter was the fixation duration. The results thus confirm the analyses in which effects of horizontal and vertical landing position were analyzed separately.

For a given fixation type, a repeated-measures ANOVA tested for any systematic effect of relative landing zone on fixation duration. The results are presented in Table 3. For first fixations, for both horizontal and vertical landing positions the average fixation duration differed between landing zones. Only the quadratic (but not the linear) trend was reliable, indicating a fixation-duration IOVP effect, with longest fixation durations at object center. For single fixations, only the effect of vertical within-object fixation location was significant, including a significant quadratic trend.

To estimate the IOVP effect quantitatively, we approximated the effect with the vertex form of the quadratic polynomial, that is,

$$y = A - B(x - C)^2 \qquad (1)$$

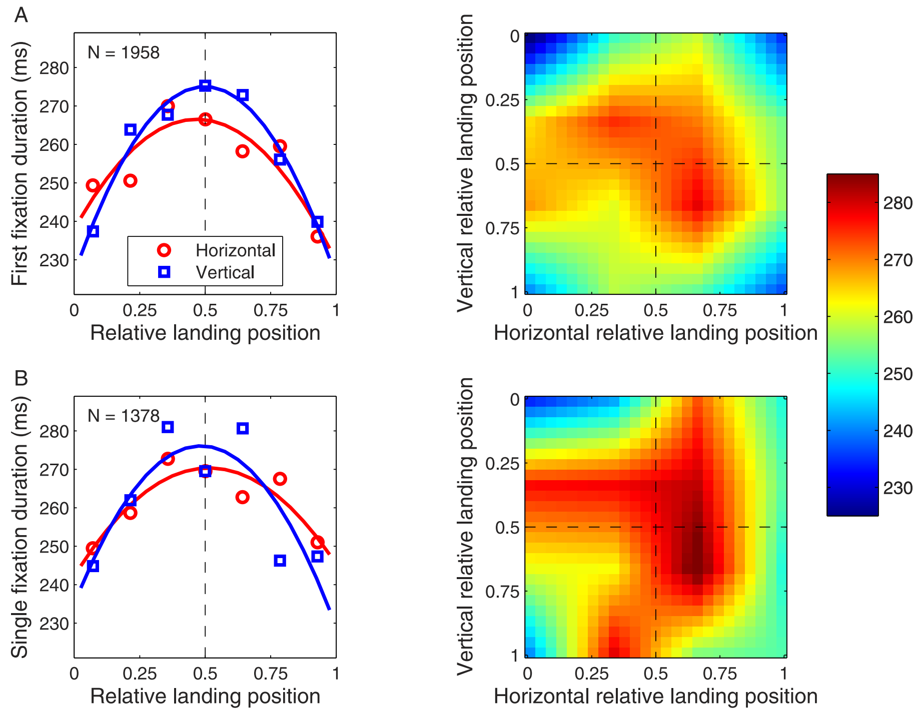

Figure 8. The effect of landing position on subsequent fixation duration for first fixations (A) and single fixations (B). Analyses considered large objects. (Left) Average fixation duration as a function of horizontal (red circles) and vertical (blue squares) landing position. (Right) Corresponding two-dimensional visualization.

where x denotes landing position (relative landing zone) and y is fixation duration (Nuthmann et al., 2005). In the vertex form, the coefficients C and A may be interpreted as the Cartesian coordinates of the vertex of the parabola. That is, C is the x-coordinate of the axis of symmetry, and A is the maximum value of the quadratic function. In the present context, A indicates maximum fixation duration, and C denotes the relative landing position at which the maximum fixation duration is observed. B is the slope of the parabolic curve, describing how fixation duration decreases with more eccentric landing positions. The fitted curves are depicted in Figure 8, and estimates of A, B, and C are presented in Table 4. Parameter B indicates that the IOVP effect was stronger for first fixations than for single fixations, which accords with findings from reading studies (Nuthmann et al., 2005; Vitu et al., 2001). In addition, for both first and single fixations, maximum fixation duration was estimated to be very close to the center of the object (0.5, 0.5), as indicated by parameter C.

	Horizontal		Vertical	
	F	η^2	*F*	η^2
First fixations				
Overall effect	2.50*	.034	4.08**	.054
Quadratic trend	10.98**	.134	21.87***	.235
Linear Trend	.82	.011	.006	.000
Single fixations				
Overall effect	1.28	.018	3.63**	.049
Quadratic trend	5.59*	.073	14.59***	.170
Linear Trend	.08	.001	.32	.004

Table 3. Effect of horizontal and vertical landing positions on first and single fixation durations for objects in scenes. *Note*: *$p < 0.05$, **$p < 0.01$, and ***$p < 0.001$.

General discussion

The aim of the present study was to further test the hypothesis that objects are important units of saccade targeting and, by inference, attentional selection in real-world scene perception and search. To investigate the object-based nature of eye guidance, we examined various phenomena related to where viewers place their eyes when fixating objects in naturalistic scenes. The present work built on our previous demonstration that the finding of a preferred viewing location generalizes from words in sentences (McConkie et al., 1988; Rayner, 1979) to objects embedded in naturalistic scenes (Nuthmann & Henderson, 2010). Here, this basic finding was further qualified by novel results on how the PVL is modulated by object size and the distance between the object and the previous fixation (i.e., launch site distance). Moreover, we examined how

	First fixation durations				Single fixation durations			
	A	*B*	*C*	χ^2	*A*	*B*	*C*	χ^2
Horizontal	267	−131	0.47	157	270	−106	0.51	89.9
Vertical	275	−196	0.50	49.2	276	−176	0.48	413.7

Table 4. Quadratic fit of IOVP curves for first and single fixation durations on objects in scenes: estimates of parameters *A* (ms), *B* (ms/U^2), and *C* (U). *Note*: χ^2 denotes sum of squared residuals.

within-object fixation position affected subsequent eye-movement behavior on the object. Unexpectedly, there was no refixation OVP effect for objects in scenes. A fixation-duration IOVP effect was observed for large objects.

Influence of object size and launch site distance on the PVL

Analyses of within-object fixation position distributions considered three object size categories (small, medium, large) and two horizontal launch site categories (near vs. far). An initial set of analyses demonstrated that the PVL varied with object size (Figure 1). For both medium-sized and large objects, the data showed a PVL such that viewers preferred to send their eyes to the center of objects. This preference for object centers over object ends was stronger for large objects than for medium-sized objects. For small objects, however, there was no evidence for a PVL, suggesting that they were not targeted or fixated optimally. A second set of analyses demonstrated additional effects of launch site distance, which interacted with object size effects in a very specific and interesting manner. Analyses considered saccades that entered objects along the horizontal axis, that is, from the left or from the right (the data were collapsed across the two horizontal categories). For far launch sites ($>5°$), there was no PVL for medium-sized objects (Figure 2b), suggesting that those objects were too small to perceive their boundaries without the high acuity present in central vision. For near launch sites ($<5°$), however, there was a horizontal PVL for medium-sized objects (Figure 2a), suggesting that object boundaries could be identified in parafoveal vision. A different pattern of results was observed for large objects, which were further divided into two subcategories ("smaller large" vs. "larger large" objects). First, a PVL was also observed for far launch sites ($>5.7°$), suggesting that those objects were sufficiently large to target them optimally in peripheral vision. Moreover, launch site distance and object width affected the mean of landing position distributions in an additive manner (Figure 3). For far launch sites, the distribution of landing positions was shifted to the left, toward the edge closest to the launch site. A similar leftward shift was observed for "larger large" as compared with "smaller large" objects. Numerically, the launch site effect was greater for the largest objects. When the eyes were relatively close to the target object (near launch sites), we observed an overshoot such that the PVL was situated somewhat beyond object center (Figure 3c). In contrast, when the eyes were farther away (far launch sites), we observed a tendency to undershoot object center (Figure 3d). In addition, the data tentatively suggest that hypometric saccades, in which the eyes undershoot the target, are less common in complex-scene viewing than in reading (McConkie et al., 1988; Rayner, 1979) or simple oculomotor aiming tasks (e.g., Henson, 1979).

In reading, launch site distance is the strongest predictor of within-word landing position, and it dominates the effect of word length (McConkie et al., 1988). In comparison, the launch site effect observed for large objects was modest in size (Figure 3). For medium-sized objects, it showed as a breakdown of the PVL when the center of the object was situated in low-acuity peripheral vision (Figure 2b). In sum, launch site distance and object size affected the PVL for objects in scenes, albeit in a manner specific to the processing of objects in scenes. Future empirical and theoretical work is required to further investigate the interplay of object size and launch site distance and to test whether oculomotor mechanisms such as the saccadic range error (McConkie et al., 1988, for reading) or Bayesian estimation of target positions (Engbert & Krügel, 2010, for reading) play a role in the process of target selection in natural scenes.

Influence of initial fixation location on refixation probability

Whether an initial fixation on an object was followed by an immediate refixation on the same object was found to depend on object size such that refixations were more likely to occur on large objects compared with small and medium-sized objects. However, refixation probability was not modulated by where viewers initially placed their eyes on an object. This flat refixation probability function contrasts with a *u*-shaped refixation probability function for words in texts (e.g., McConkie et al., 1989; Nuthmann et al.,

2005) and isolated objects (Foulsham & Underwood, 2009; Henderson, 1993). In those tasks and settings, refixations were least likely at word or objects center and increased as the initial landing position moved further out.

The way the uniform refixation probabilities for large objects in scenes (Figure 4) differ from results for long words in reading (McConkie et al., 1989; Nuthmann et al., 2005) can be described as follows: Fewer refixations were launched from object boundaries and more from object centers. Let us first consider refixations launched from object or word ends. The slope of the quadratic refixation probability function in reading is thought to indicate the rate of drop off of visual information necessary for reading as a function of retinal eccentricity (McConkie et al., 1989). As the initial fixation location moves away from the center of the word, the amount of clear visual information provided by the word is reduced, which is increasingly compensated for by a (corrective) refixation. In keeping with the reading results, refixations launched close to object boundaries showed a tendency to correct suboptimal initial viewing locations (Figures 6 and 7). Compared with words in reading, those refixations appear to occur less frequently on objects in scenes. The visual span in real-world scene perception and search corresponds to 8° in each direction from fixation (Nuthmann, 2012), which is larger than the perceptual span in reading (Rayner, 2009). We speculate that the larger span in scene perception may reduce the need for corrective refixations.

Contrary to predictions, refixation probability was not reduced for initial fixations that were located optimally at the center of objects. Given a central PVL for first fixations in two-fixation cases, the process of transforming refixation probabilities back into absolute numbers inevitably means that most refixations were launched from central positions. Object center is thought to be optimal for visual object processing. Therefore, refixation probability at object center is indicative of refixations that were made for reasons other than insufficient visual information. According to our analyses, the majority of those refixations kept the eyes in proximity to object center (Figure 7). As a matter of fact, across launch cells, refixation saccades were preferentially directed toward locations close to object center (Figure 7, top panel). At present, we can only speculate about the origin of centrally launched refixations. There is the possibility that those refixations readjust the eyes to fixate on parts of the object that are particularly informative. At least some proportion of those refixations may be caused by difficulties at higher processing levels, thus keeping the eyes from advancing to the next scene element until the observer is ready for new visual input. Our own CRISP model of fixation durations in scene viewing (Nuthmann, Smith, Engbert, & Henderson, 2010) offers an additional explanation. An autonomous timer, conceptualized as a random walk process, generates commands to initiate saccade programs. Once the random walk timing signal reaches threshold, a new saccade program is initiated. Processing difficulties can inhibit and thus modulate saccade timing. In such a framework, some of the centrally launched refixation saccades would simply be a consequence of autonomous saccade timing. In sum, further research is needed to investigate the basis for such refixations.

Influence of within-object fixation location on fixation duration

Only for large objects did we find a systematic relationship between fixation location and fixation duration that paralleled the fixation-duration IOVP effect found for words in sentences (e.g., Nuthmann et al., 2005; Vitu et al., 2001) and isolated objects (Henderson, 1993). Mean fixation durations on those objects were longest when the eyes were placed near the center of the object. As landing position moved further toward object boundaries, fixation duration decreased. This drop showed negative acceleration with distance from object center.

In the reading literature, the IOVP effect has been attributed to perceptual economy reasons (Vitu et al., 2007; Vitu et al., 2001) or mislocated fixations (Nuthmann et al., 2005, 2007). According to the perceptual-economy account, the oculomotor system plans on keeping the eyes longer at optimal locations where greater amounts of information are anticipated. This principle is assumed to be universal, and its descriptive nature may indeed account for any IOVP effect in any visual task, including the IOVP effect for objects in scenes observed here. However, the account has difficulty explaining why the effect was found for large objects only (as object center should be favored regardless of object size). In an interesting extension of the account, a functional relationship between refixation OVP and fixation-duration IOVP effects was proposed (Vitu et al., 2007). Two alternative relations were put forward: (a) both OVP and IOVP effects result from perceptual-economy processes (full perceptual-economy hypothesis), and (b) the OVP effect derives from the IOVP effect, which itself results from perceptual-economy processes (temporal perceptual-economy hypothesis). Thus, both hypotheses rely on the finding that the word refixation curve has a local minimum at word center (e.g., McConkie et al., 1989; Nuthmann et al., 2005). For objects embedded in natural scenes, however, the refixation function has no local minimum at object center (Figure 4). Rather, the findings for objects in scenes lead to the following paradox: Object center is optimal in the sense that saccades to medium-sized and large objects prioritize this location (Figure 1). At the same time, however, those instances of optimal saccade targeting are frequently accompanied by immediate refixations.

Nuthmann et al. (2005, 2007) linked the fixation-duration IOVP effect to mislocated fixations. Distributions of within-word landing positions are relatively broad and show overlapping tails, which suggests that a significant proportion of fixations is mislocated and falls on words to the left or right of the selected target word (McConkie et al., 1988). Methods for the quantitative estimation of the likelihood of mislocated fixations in reading analyze the overlapping tails of the landing position distributions for adjacent words (Engbert & Nuthmann, 2008; Engbert et al., 2007; Nuthmann et al., 2005). For objects in scenes, Gaussian landing position distributions are truncated at object boundaries (see also Nuthmann & Henderson, 2010), suggesting that some fixations are mislocated such that they fall short of the intended target object. However, estimating the prevalence of mislocated fixations for objects in scenes is complicated by a number of problems. Compared to scene viewing, reading takes place in a highly structured visual environment. In alphabetic scripts, words are arranged along horizontal lines, and the spaces between words provide cues for word boundaries. Unlike words in texts, objects in real-world scenes are not organized in a tight and regular grid. Instead, natural scenes are often cluttered and contain objects that partially occlude each other. Moreover, what constitutes an object may depend on the task and mind-set of the observer (see Nuthmann & Henderson, 2010, for discussion). In sum, although distributions of within-object fixation locations suggest that mislocated fixations exist in scene viewing, it is presently unclear as to whether they contribute to the fixation-duration IOVP effect observed for large objects.

Outlook

Although the present data confirm that attentional selection in scenes has a strong object-based component, it is clear that not all fixations in scenes fall on objects (see Nuthmann & Henderson, 2010, for discussion). For instance, there is the possibility that object-based selection is, at least partially, combined with center-of-gravity saccade averaging (e.g., Vitu, 1991, 2008, for reading; Zelinsky, Rao, Hayhoe, & Ballard, 1997, for search). On first inspection of a given scene region, the eyes may not aim for specific objects but may be drawn toward the center of gravity of the visual configuration formed by a smaller group of objects. Those off-object fixations (see Zelinsky, 2012, for search) may be followed by accurate fixations to individual objects. Testing such hypotheses most likely requires experimental manipulation of object configurations in highly controlled 3D-rendered images of real-world scenes.

Similarly, controlled experiments may follow up some of the corpus-analytical results reported here. For example, when assessing effects of object size, we excluded objects with extreme aspect ratios. However, it is possible that the width and height of objects modulate the PVL independently in that very tall objects (e.g., a tower) may be targeted in a different manner than wide objects (e.g., a bench). To facilitate such research, we conducted an exploratory analysis based on the present corpus data. The previously introduced three object-size categories (small, medium, large) were used to create a full design in which three object-width categories (small, medium, large) were crossed with three object-height categories (small, medium, large). In Figure 9, each row plots data for a fixed object height category, with height increasing from small (bottom row) to large (top row). In contrast, each column plots data for a fixed object width category, with width increasing from small (left column) to large (right column). Consequently, each panel depicts one combination of object width and object height. The panels in the diagonal (Small–Small, Medium–Medium, Large–Large) replicate the data displayed in Figure 1. With regard to the differential effects of object width and height, the data can be summarized as follows: The more the width of the object increased, the more evidence there was for a horizontal PVL emerging at a central location (Figure 9, for a given row from left to right). Likewise, the more the height of the object increased, the more evidence there was for a vertical PVL (Figure 9, for a given column from bottom to top). The separate analyses of horizontal and vertical landing position components (Figure 9) were confirmed by corresponding analyses in two-dimensional space (Supplementary Figure S2). Collectively, the data seem to suggest that object size along a given axis affects the PVL in the same axis. However, the post hoc design of the analysis poses limitations on the interpretation of results. Objects within a given category were not fixed in size and aspect ratio (see Supplementary Figure S1). Object width and height were (still) correlated (see above). Some categories comprised few objects (see Figure 9). Therefore, the present results derived from corpus analyses should be followed up by an experiment in which object width and object height are independently varied.

Finally, the reported PVL analyses considered low-level visuomotor variables such as object size and launch site distance. Overall, the data support the assumption that the geometric center of an object affords optimal visual processing. This is not to

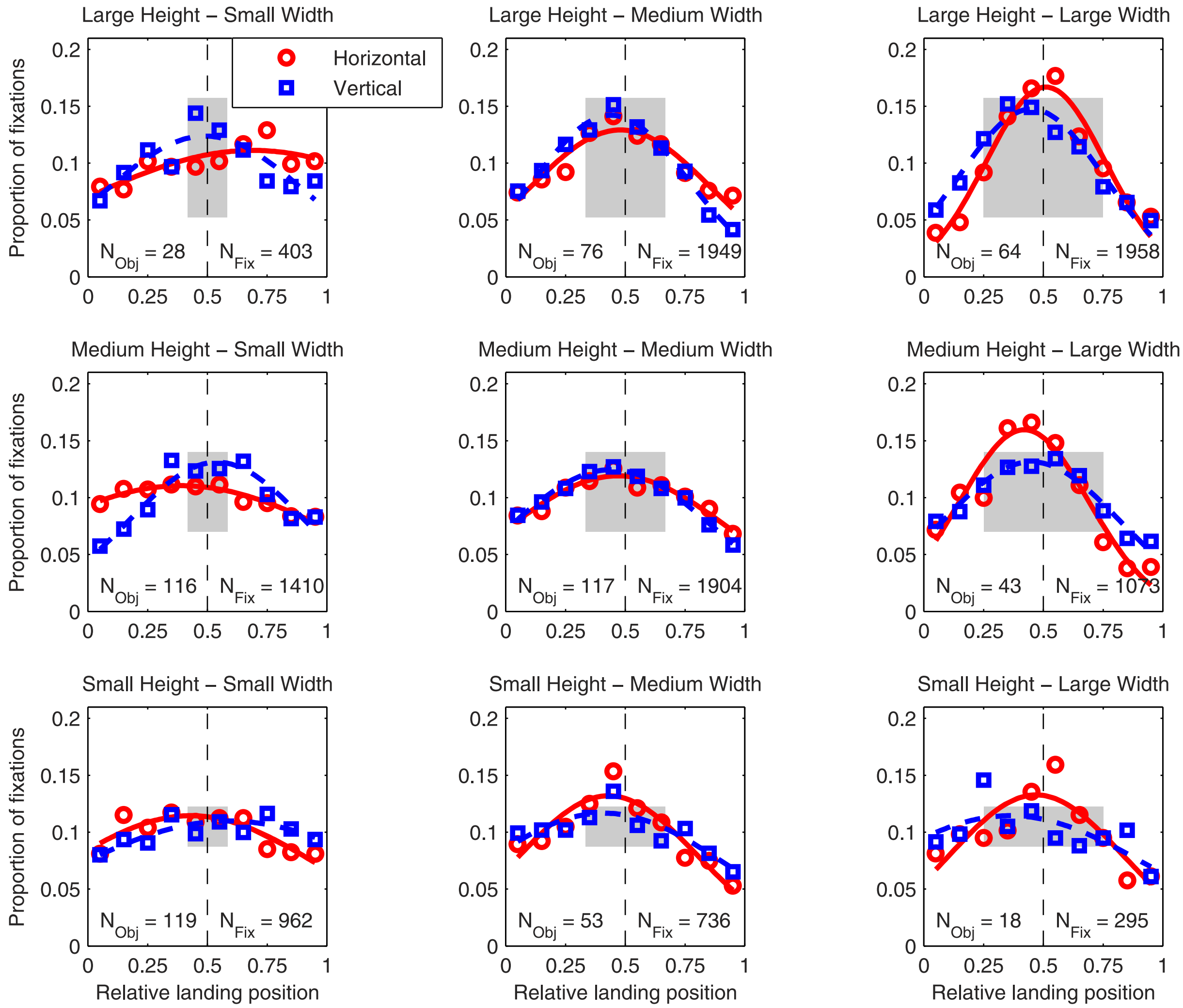

Figure 9. Effects of object width and object height on the preferred viewing location for objects in scenes. Three object width categories were crossed with three object height categories. For the resulting nine combinations of object width and height, the figure shows the distributions of fixation locations within objects. Horizontal (red circles) and vertical (blue squares) components of within-object landing positions were analyzed separately. Experimentally observed distributions were fitted using truncated Gaussians. The vertical broken line marks a landing position at the center of the object. In the figure, each row plots data for a fixed object height category, with height increasing from small (bottom row) to large (top row). Each column plots data for a fixed object width category, with width increasing from small (left column) to large (right column). The panels in the diagonal replicate the data displayed in Figure 1. To aid comprehension, the respective cell of the design is described by the panel title and schematically visualized by the light gray rectangle in the panel background.

discount the possibility that within-object landing positions are affected by higher-level variables such as the informativeness of regions within objects or object affordances. For words in reading, for example, subtle linguistic influences on the PVL have been observed (e.g., White & Liversedge, 2004).

Conclusion

In summary, the present work further qualified the nature of object-based saccade targeting in real-world scene perception. Informed by a rich literature on the selection of words in reading, we examined whether various landing position-related phenomena (PVL, OVP, IOVP) generalized from words in sentences to objects embedded in naturalistic scenes. Compared to reading, the object-in-scene data revealed qualitative similarities and differences. As predicted, object size and launch site distance affected the PVL for objects in scenes, albeit in a manner specific to the processing of objects in scenes. The word refixation OVP effect was found not to generalize to objects embedded in naturalistic scenes, suggesting that some refixations on objects in scenes are made for reasons other than

insufficient visual information. A fixation-duration IOVP effect was found for large objects only. The results support the conclusion that objects play an important role in eye-movement control during scene viewing. To a certain extent, the selection of objects in scenes parallels the selection of words in reading.

Keywords: naturalistic scenes, objects, fixation location, refixation, fixation duration

Acknowledgments

Portions of this research were conducted by M.P. as part of an undergraduate honors project at the University of Edinburgh and were presented at the BPS Scottish Branch Undergraduate Student Conference (Glasgow, Scotland, March 2012). M.P. was supported in part by grants EP/F500385/1 and BB/F529254/1 to the University of Edinburgh Doctoral Training Centre in Neuroinformatics and Computational Neuroscience (www.anc.ac.uk/dtc) from the UK EPSRC, BBSRC, and MRC research councils. We are grateful to John M. Henderson for making the eye-tracking corpus data available and thank Robin Hill for assistance in data collection. Data collection was supported by a grant from the Economic and Social Research Council (ESRC) of the UK to JMH (RES-062-23-1092).

Commercial relationships: none.
Corresponding author: Antje Nuthmann.
Email: antje.nuthmann@ed.ac.uk.
Address: University of Edinburgh, Psychology Department, Edinburgh, UK.

Footnotes

[1] The terms landing position and viewing location are synonyms for fixation location.

[2] In the third task, participants searched for an object in the scene. Participants were presented with a target word prior to scene presentation and pressed a button as soon as the object was located in the scene.

References

Eilers, P. H. C., & Goeman, J. J. (2004). Enhancing scatterplots with smoothed densities. *Bioinformatics, 20*, 623–628.

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14):18, 1–26, http://www.journalofvision.org/content/8/14/18/, doi:10.1167/8.14.18. [PubMed] [Article]

Engbert, R., & Krügel, A. (2010). Readers use Bayesian estimation for eye movement control. *Psychological Science, 21*, 366–371.

Engbert, R., & Nuthmann, A. (2008). Self-consistent estimation of mislocated fixations during reading. *PLoS ONE, 3*, e1534.

Engbert, R., Nuthmann, A., & Kliegl, R. (2007). An iterative algorithm for the estimation of the distribution of mislocated fixations during reading. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain*. (pp. 319–337). Oxford, UK: Elsevier.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review, 112*, 777–813.

Foulsham, T., & Underwood, G. (2009). Does conspicuity enhance distraction? Saliency and eye landing position when searching for objects. *Quarterly Journal of Experimental Psychology, 62*, 1088–1098.

Foulsham, T., & Underwood, G. (2007). How does the purpose of inspection influence the potency of visual salience in scene perception? *Perception, 36*, 1123–1138.

Henderson, J. M. (1993). Eye movement control during visual object processing: Effects of initial fixation position and semantic constraint. *Canadian Journal of Experimental Psychology, 47*, 79–98.

Henderson, J. M. (2011). Eye movements and scene perception. In I. D. Gilchrist, S. P. Liversedge, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 593–606). Oxford, UK: Oxford University Press.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*, 498–504.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford, UK: Elsevier.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review, 16*, 850–856.

Henson, D. B. (1979). Investigation into corrective saccadic eye movements for refixation amplitudes of 10 degrees and below. *Vision Research, 19*, 57–61.

Hyönä, J., & Bertram, R. (2011). Optimal viewing position effects in reading Finnish. *Vision Research, 51*, 1279–1287.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*, 194–203.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*, 1489–1506.

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.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28*, 245–253.

McConkie, G. W., Kerr, P. W., Reddix, M. D., Zola, D., & Jacobs, A. M. (1989). Eye movement control during reading: II. Frequency of refixating a word. *Perception & Psychophysics, 46*, 245–253.

McDonald, S. A., Carpenter, R. H. S., & Shillcock, R. C. (2005). An anatomically-constrained, stochastic model of eye movement control in reading. *Psychological Review, 112*, 814–840.

McDonald, S. A., & Shillcock, R. C. (2004). The potential contribution of preplanned refixations to the preferred viewing location. *Perception & Psychophysics, 66*, 1033–1044.

Nuthmann, A. (2012). *On the visual span during object search in real-world scenes*. Manuscript submitted for publication.

Nuthmann, A. (2006). *The "where" and "when" of eye fixations in reading* (Doctoral dissertation, University of Potsdam). Retrieved from http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-7931

Nuthmann, A., & Engbert, R. (2009). Mindless reading revisited: An analysis based on the SWIFT model of eye-movement control. *Vision Research, 49*, 322–336.

Nuthmann, A., Engbert, R., & Kliegl, R. (2007). The IOVP effect in mindless reading: Experiment and modeling. *Vision Research, 47*, 990–1002.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45*, 2201–2217.

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision, 10*(8):20, 1–19, http://www.journalofvision.org/content/10/8/20, doi:10.1167/10.8.20. [PubMed] [Article]

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*, 382–405.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8*, 21–30.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62*, 1457–1506.

Rayner, K., Liversedge, S. P., Nuthmann, A., Kliegl, R., & Underwood, G. (2009). Rayner's 1979 paper. *Perception, 38*, 895–906.

Rayner, K., Sereno, S. C., & Raney, G. E. (1996). Eye movement control in reading: A comparison of two types of models. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1188–1200.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5):9, 1–23, http://www.journalofvision.org/content/11/5/9, doi:10.1167/11.5.9. [PubMed] [Article]

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5):5, 1–23, http://www.journalofvision.org/content/11/5/5, doi:10.1167/11.5.5. [PubMed] [Article]

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 113*, 766–786.

Trukenbrod, H. A., & Engbert, R. (2007). Oculomotor control in a sequential search task. *Vision Research, 47*, 2426–2443.

Vitu, F. (2008). About the global effect and the critical role of retinal eccentricity: Implications for eye movements in reading. *Journal of Eye Movement Research, 2*, 1–18.

Vitu, F. (1991). The existence of a center of gravity effect during reading. *Vision Research, 31*, 1289–1313.

Vitu, F., Lancelin, D., & d'Unienville, V. M. (2007). A perceptual-economy account for the inverted-optimal viewing position effect. *Journal of Experimental Psychology: Human Perception and Performance, 33*, 1220–1249.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: an inverted optimal viewing position effect. *Vision Research, 41*, 3513–3533.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19*, 1395–1407.

White, S. J., & Liversedge, S. P. (2006). Linguistic and non-linguistic influences on the eyes' landing positions during reading. *Quarterly Journal of Experimental Psychology A, 59,* 760–782.

White, S. J., & Liversedge, S. P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16*, 52–78.

Yang, S.-N., & Vitu, F. (2007). Dynamic coding of saccade length in reading. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 293–317). Oxford, UK: Elsevier.

Zelinsky, G. J. (2012). TAM: Explaining off-object fixations and central fixation tendencies as effects of population averaging during search. *Visual Cognition, 20*, 515–545.

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8*, 448–453.

Appendix

Launch Site Distance for Horizontal Saccades Directed Toward Objects in Scenes

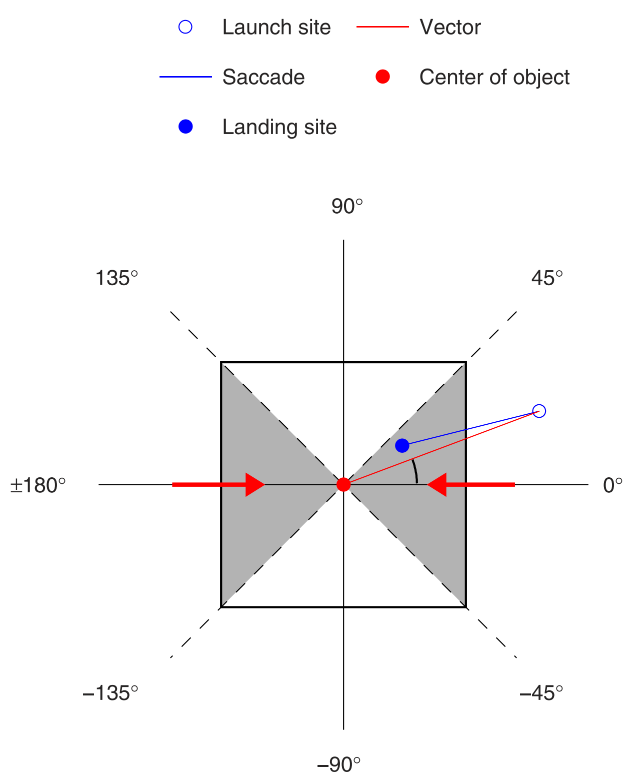

Figure A1. The effect of launch site distance on within-object fixation positions was evaluated for instances in which the eyes entered the object along the horizontal axis, that is, from the left or from the right. The figure shows the normalized object and an example saccade that was launched from outside the object and led to a fixation on the object. The saccade entered the object from the right. Launch site distance was defined as the vector connecting the launch site of the saccade and the center of the targeted object. This is different from saccade amplitude, denoting the distance the eyes traveled from the launch site to the landing site. The direction of initial saccades toward objects was examined as the angle between the horizontal plane and the vector denoting center-based launch site distance. For analyses, object boxes were divided into two directional segments, each subtending 45° (toward either side) around the two horizontal axes connecting to the center of the object (left, right). For example, saccades were classified as entering the object from the right if the angle between the launch site and the center of the object was between −45° and +45°.