On the Relative (Un)importance of Foveal Vision during Letter Search in Naturalistic Scenes

Adam C. Clayden[1, 2], Robert B. Fisher[3] and Antje Nuthmann[1, 4]

[1] Psychology Department, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, UK

[2] School of Engineering, Arts, Science and Technology, University of Suffolk, UK

[3] School of Informatics, University of Edinburgh, UK

[4] Institute of Psychology, University of Kiel, Germany

Author Note

Antje Nuthmann http://orcid.org/0000-0003-3338-3434

The T.E.A is available at https://github.com/AdamClayden93/tea.

The project was supported by the Engineering and Physical Sciences Research Council (UK). Portions of this research were presented at the 18th European Conference on Eye Movements (Vienna, Austria, 2015).

Correspondence concerning this article should be addressed to Antje Nuthmann, University of Kiel, Institute of Psychology, Olshausenstr. 62, 24118 Kiel, Germany. Email: nuthmann@psychologie.uni-kiel.de

27

Abstract

The importance of high-acuity foveal vision to visual search can be assessed by denying foveal vision using the gaze-contingent Moving Mask technique. Foveal vision was necessary to attain normal performance when searching for a target letter in alphanumeric displays, Perception & Psychophysics, 62 (2000) 576-585. In contrast, foveal vision was not necessary to correctly locate and identify medium-sized target objects in natural scenes, Journal of Experimental Psychology: Human Perception and Performance, 40 (2014) 342-360. To explore these task differences, we used grayscale pictures of real-world scenes which included a target letter (Experiment 1: T, Experiment 2: T or L). To reduce between-scene variability with regard to target salience, we developed the Target Embedding Algorithm (T.E.A.) to place the letter in a location for which there was a median change in local contrast when inserting the letter into the scene. The presence or absence of foveal vision was crossed with four target sizes. In both experiments, search performance decreased for smaller targets, and was impaired when searching the scene without foveal vision. For correct trials, the process of target localization remained completely unimpaired by the foveal scotoma, but it took longer to accept the target. We reasoned that the size of the target may affect the importance of foveal vision to the task, but the present data remain ambiguous. In summary, the data highlight the importance of extrafoveal vision for target localization, and the importance of foveal vision for target verification during letter-in-scene search.

243 words

Keywords: naturalistic scenes; visual search; target size; eye movements; foveal vision

1. Introduction

How important is the availability of high-acuity foveal vision to visual search? This question has been investigated with different search tasks, ranging from letter search in alphanumeric displays (Bertera & Rayner, 2000) to object-in-scene search (Nuthmann, 2014), either highlighting the relative importance (letter search) or unimportance (scene search) of foveal vision. The aim of the present work was to combine design features from both search paradigms to better understand these task differences. In Experiment 1, observers searched for the letter "T" embedded in grayscale pictures of real-world scenes, with or without foveal vision. In Experiment 2, we added a letter recognition component to the search task ("Is it a T or an L?"). In both experiments, we also varied the size of the letter target to investigate the degree to which the importance of foveal vision depends on the size of the search target.

Visual acuity is highest at the fovea before declining rapidly as it approaches the periphery (Strasburger, Rentschler, & Jüttner, 2011, for review). Whereas the foveal region is typically defined as the central 2° of vision, the parafoveal region extends from the foveal region out to about 5° from fixation; the fovea and parafovea together are commonly referred to as central vision. The peripheral region is everything beyond the parafoveal region. During each eye fixation, information may be extracted from foveal, parafoveal, and peripheral regions of the visual field.

The importance of foveal vision was first studied in sentence reading by means of the gaze-contingent *Moving Mask* technique. To this end, Rayner and Bertera (1979) aligned a visual mask with the reader's gaze to wipe out the text in view. The size of the mask ranged between 1 and 17 characters (1° = three characters). Simulating reading without a fovea in that manner reduced the reader's reading speed by increasing the number of fixations, fixation duration, and reducing saccade length. Moreover, reading comprehension suffered. The same authors also investigated the importance of foveal vision in visual search (Bertera

& Rayner, 2000). In this study, participants searched for the target letter "y" within a randomly arranged array of alphanumeric characters, with or without a simulated scotoma. Five different scotoma sizes, ranging from 0.3º to 3º, were tested. As the mask size increased, the lower the search accuracy, the longer the search time, and the more fixations were made. Geringswald, Baumgartner, and Pollmann (2012) investigated the impact of a large simulated central scotoma (diameter: 9º) on contextual cueing in visual search. Participants searched for a T-shaped target among L-shaped distractors. Blocking out central vision eliminated the search facilitation which is oftentimes observed for targets appearing in repeated configurations (see also Geringswald & Pollmann, 2015).

Interestingly, visual search studies involving naturalistic scenes have found rather different results (McIlreavy, Fiser, & Bex, 2012; Nuthmann, 2014). In the study by Nuthmann (2014), participants searched for a specific object in a colored image of a real-world scene (e.g., a blender in a kitchen scene). Search was cued with a word label and search objects had an average size of 2.5º × 2.5º (medium size). When searching the scene with artificially impaired foveal or central vision[1], search performance was surprisingly unimpaired. *Foveal* vision was not necessary to attain normal search performance. When searching without *central* vision, participants' gaze data revealed that they were not impaired in locating the search object in the scene, but in verifying that the target was in fact the target. In the study by Nuthmann (2014), the scene image contained contextually relevant search targets (cf. Torralba, Oliva, Castelhano, & Henderson, 2006). McIlreavy et al. (2012) excluded such contextual guidance towards the target by asking observers to look for spatial distortions (Bex, 2010), which were embedded at random places in grayscale images of natural scenes. The results for search times were similar to the ones by Nuthmann (2014).

---

[1] The size (i.e., radius) of the scotoma was manipulated as the standard deviation of the two-dimensional Gaussian distribution that was used to mix the high-resolution foreground with a low-resolution background image; foveal scotoma: $\sigma_{x,y}$ = 1.6º, central scotoma: $\sigma_{x,y}$ = 4.1º.

Searching with a foveal scotoma ($\sigma_{x,y}$ = 1º) had no detrimental effect on performance. Only the largest central scotoma condition ($\sigma_{x,y}$ = 4º) led to a significant increase in mean search time.

During overt search of any kind, the information extracted during eye fixations subserves both a peripheral selection task as well as a central discrimination task (Hooge & Erkelens, 1999; Shen, Reingold, Pomplun, & Williams, 2003). The peripheral selection task determines the target location for the next saccade, whereas the central discrimination task involves an accept/reject decision about whether the fixated object is the target. Since foveal analysis allows for encoding fine perceptual detail, making foveal vision unavailable should be disruptive to the central discrimination task. However, such reasoning ignores the fact that the processing of the fixated object or region can begin prior to the start of fixation via extrafoveal processing (Reichle & Reingold, 2013; Reingold & Glaholt, 2014). Thus, simple search and scene search may differ in the way extrafoveal processing enables the extraction of information that is required to reject distractors and to accept the target. Moreover, the relation between foveal analysis and peripheral selection may be task dependent (cf. Shen et al., 2003).

Target size is a feature that may be relevant in this regard. Both McIlreavy et al. (2012) and Nuthmann (2014) discuss that target size could be an important mediating factor for their findings on the (un)importance of foveal vision. Before elaborating on this argument, we briefly review research on size and eccentricity effects in (normal) visual search. A common paradigm is to use fairly small simple displays which observers search covertly in the absence of eye movements. Using this approach, Duncan and Humphreys (1989) investigated the effect of stimulus size and showed that search is more difficult for small letters than for large letters. A related finding is the eccentricity effect: search performance deteriorates as the target is presented at farther peripheral locations (Carrasco, Evert, Chang, & Katz, 1995; Geisler & Chou, 1995). This reduction in search efficiency may be due to the poorer spatial resolution in the periphery. Consistent with this view, enlarging the stimuli according to the cortical magnification factor (Rovamo & Virsu, 1979) eliminated the eccentricity effect (Carrasco & Frieder, 1997; Carrasco, McLean, Katz, & Frieder, 1998; but see Wolfe, O'Neill, & Bennett, 1998, Experiment 4a). The eccentricity effect is also observed in the presence of eye movements (Scialfa & Joffe, 1998; Zelinsky, 2008).

In the context of visual search in real-world scenes, the effect of target size has received little systematic investigation. Wolfe, Alvarez, Rosenholtz, Kuzmova, and Sherman (2011, Experiment 1) had observers search for annotated objects in photographs of real-world scenes. The objects showed a natural variability in size and eccentricity and search times were found to increase for both smaller as well as more eccentric targets. Miellet, Zhou, He, Rodger, and Caldara (2010) asked both Eastern and Western observers to search for animals in zoo photographs. In the experiment, target size and the size of a gaze-contingent moving mask were parametrically manipulated (size/ diameter: 2º, 5º, or 8º). Search performance was better for larger targets. As the simulated scotoma got larger, performance increasingly suffered (cf. McIlreavy et al., 2012; Nuthmann, 2014). Importantly, there was an interaction between mask size and target size such that the deleterious effect of mask size was more pronounced for smaller targets. In the 2º-*Blindspot* condition, making foveal vision unavailable, search performance was reduced for 2º targets but not for 8º targets. Although suggestive, any findings involving target size in this study need to be treated cautiously because target salience (Itti & Koch, 2000) was not controlled for. Other potential confounds are target eccentricity (i.e., distance from scene center) and contextual guidance.

The goal of the present research was to further investigate the importance of foveal vision to visual search. Stimuli were grayscale pictures of real-world scenes in which a target letter was inserted (Experiment 1: T, Experiment 2: T or L). Four letter sizes, ranging from

0.25º to 1.5º in width, were crossed with the presence vs. absence of foveal vision. To control for visual salience, the letter was algorithmically placed for each scene in a location for which there was a medium change in local contrast when inserting the letter. Letter targets were used for a number of reasons. The small to large animal targets in Miellet et al. (2010) were all part of different scenes. Our approach allowed us to place letter targets of variable size at the same location within a given scene. In addition, using context-free letter targets rather than contextually relevant search targets prevents observers from using their knowledge about the likely positions of targets to guide their eye movements (cf. McIlreavy et al., 2012). Our task still approximates natural behavior because there are real-world searches for which there is minimal guidance by scene context (e.g., search for a fly). Perhaps more importantly, scene processing and object identification are not totally suppressed when searching for a "T" overlaid onto the scene (T. H. W. Cornelissen & Võ, 2017). In Experiment 1, on each trial participants were asked to look for the letter "T". In Experiment 2, we added a recognition component to the task. The target was either a "T" or an "L", and—once they found the letter—participants had to indicate which one it was. We chose these two letters because they share exactly the same features (strokes) and differ only in their spatial arrangement (Duncan & Humphreys, 1989). Because we used participants' eye-movement data to verify that targets had indeed been found, there were no target-absent trials (Nuthmann, 2013, 2014; Nuthmann & Malcolm, 2016).[2]

If foveal vision is necessary to achieve normal search performance during letter-in-scene search, then we should observe a reduction in performance—lower search accuracy and longer search time—when searching the scene with a simulated foveal scotoma, compared

---

[2] In a typical laboratory search experiment, the observer's task is to establish whether the target is present or absent amongst other distractor items (Wolfe, 2014). In the present experiments, observers are asked to acquire the target with their eyes so the task has been referred to as target acquisition rather than search (Zelinsky, 2008).

with a normal-vision control condition. Moreover, we expected to find effects of target size, with better performance for larger targets. Critically, the experimental design allowed us to investigate whether the importance of foveal vision depended on the size of the search target (cf. Miellet et al., 2010). Why would size matter? Here, our hypotheses concern two separate sub-processes of search: scanning for the target and accepting the target. The scanning process involves the localization of the target in space, the duration of which (scanning time) is indexed as the time between the first saccade and the first fixation on the target (Malcolm & Henderson, 2009). Similarly, verification time is the elapsed time between the beginning of the first fixation on the target and search termination.

The possibility exists that the actual search process, indexed by the scanning time, is slowed down when foveal vision is absent. The reject decision during scanning epoch fixations may be impaired if the extraction of information in extrafoveal vision (on the previous, but also on the current fixation) cannot compensate for the lack of foveal analysis. Moreover, the difficulty of central discrimination may affect the efficiency of peripheral selection, if the two tasks share resources (see Shen et al., 2003, for discussion). Alternatively, blocking out foveal vision may only affect the verification process, as explained next.

Upon fixation with a foveal scotoma, all of the target—or some part of it—will be covered by the scotoma. The extent of this masking depends on both the size of the target and the initial fixation position on the search target (Nuthmann, 2014). If the available information is not sufficient to make the accept decision, the eyes may move off the target to unmask the letter and to process it in parafoveal or peripheral vision (cf. Nuthmann, 2014). Such behavior would increase verification times. We hypothesized that any detrimental effect of the foveal scotoma may only occur for smaller targets, or may be more pronounced for those. Moreover, in Experiment 2 we changed the task to involve not only target detection but also target identification. At least for small letters, letter identification may require the extraction of fine detail via foveal analysis. Therefore, we reasoned that any adverse effect of the foveal scotoma, and its interaction with target size, may be stronger in Experiment 2 than in Experiment 1.

2. Methods

2.1. Participants

Thirty-two participants (12 males) between the ages of 18 and 27 (mean age 20 years) participated in Experiment 1. Thirty-two different participants (8 males) between the ages of 18 and 27 (mean age 22 years) participated in Experiment 2. All participants had normal or corrected-to-normal vision by self-report. They gave their written consent prior to the experiment and either received study credit or were paid at a rate of £7 per hour for their participation. Ethics approval was obtained from the Psychology Research Ethics Committee of the University of Edinburgh.

2.2. Apparatus

Working with gaze-contingent displays requires minimizing the latency of the system. This was achieved by using (a) an eye tracker with high temporal resolution, (b) modern graphics hardware, and (c) a monitor with a high refresh rate. Stimuli were presented on a 21-inch CRT monitor with a refresh rate of 140 Hz at a viewing distance of 90 cm, taking up a 24.8º × 18.6º (width × height) field of view. A chin and forehead rest was used to keep the participants' head position stable. During stimulus presentation, the eye movements of the participants were recorded binocularly with an SR Research EyeLink 1000 Desktop mount system with high accuracy (0.15º best, 0.25-0.5º typical) and high precision (0.01º RMS). The Eyelink 1000 was equipped with the 2000 Hz camera upgrade, allowing for binocular recordings at a sampling rate of 1000 Hz per eye. The experiments were programmed in MATLAB 2013a (The MathWorks, Natick, MA) using the OpenGL-based Psychophysics

Toolbox 3 (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007) which incorporates the EyeLink Toolbox extensions (F. W. Cornelissen, Peters, & Palmer, 2002). A game controller was used to record participants' behavioral responses.

2.3. Stimuli

In Experiment 1, stimuli consisted of 120 grayscale images of naturalistic scenes (800 × 600 pixels), which came from a variety of categories; 104 of these photographs were previously used as colored images in Nuthmann (2014). Example scenes are shown in Figures 1 and 3. Eight additional images were used as practice scenes. Image processing techniques (See Section Target Embedding Algorithm below) were used to insert the letter T in four sizes at the same location within a given scene, such that the chosen location was of median salience, as explained below. Note that in the experiment, each participant viewed a given scene only once, in one of the four target size conditions (and either with or without foveal vision).

In Experiment 2, 128 (+ 8 practice) grayscale images of real-world scenes were used, 120 of which were from experiment 1 with 8 new images. The new images were chosen because the experimental design required an equal number of T- and L-scenes in each target-size condition. The search target was either a letter T or L that was again algorithmically placed into the scene at a median salience location.

2.4. Design

Both experiments used a 2 × 4 within-subjects design with 2-level factor foveal vision (present vs. absent) and 4-level factor target size. The factor foveal vision refers to the implementation of a foveal scotoma. In the scotoma condition, foveal vision was blocked by a gaze-contingent moving mask (foveal vision absent, or scotoma on). This was contrasted with a normal-vision control condition (foveal vision present, or scotoma off).

In both experiments, the presence or absence of foveal vision was crossed with four target sizes. In Experiment 1, they were equally spaced as follows: S - Small (letter width 0.25º), M - Medium (0.66º), L - Large (1.08º), and XL - Extra Large (1.5º). The XL target size was chosen such that the foveal scotoma, which had a radius of 1º, completely obscured the target when observers fixated the center of the letter. In Experiment 2, we removed the XL targets; instead, we added targets of intermediate size (0.41º) halfway between the small and medium targets. These adjustments were informed by the results obtained in Experiment 1: search efficiency was much worse for small targets compared with medium-sized targets, while performance differences between large and extra-large targets were much less pronounced.

In Experiment 1, the 120 T-scenes were assigned to eight lists of 15 scenes each. The scene lists were rotated over participants, such that a given participant was exposed to a list for only one of the eight experimental conditions created by the 2 × 4 design. There were eight groups of four participants, and each group of participants was exposed to unique combinations of list and experimental condition. To summarize, participants viewed each of the 120 scene items once, with 15 scenes in each of the eight experimental conditions. Across the 32 participants, each scene item appeared in each condition four times.

For Experiment 2, each of the 128 original scene images was submitted to the Target Embedding Algorithm to produce four T-scenes and four L-scenes, one for each target size. In the experiment, half of the original scenes were used as T-scenes, the other half as L-scenes. Since the algorithm placed the Ts and Ls of four different sizes in the same location, there were a few cases where the horizontal bar of the T or the vertical bar of the L blended with a dark scene background. Therefore, the decision about which scenes to use in either category was guided by visual inspection. We then created eight scene lists, each comprising eight T-scenes and eight L-scenes. Apart from that, the same counterbalancing procedure as in Experiment 1 was used to control for item effects.

The foveal vision manipulation was blocked so that participants completed two blocks of trials in the experiment: in one block observers' foveal vision was available, in the other block it was obstructed by a gaze-contingent scotoma. Each block started with four practice trials, one for each target size condition. The order of blocks was counterbalanced across subjects. Within a block, scenes with targets of different sizes and types (Experiment 2 only) were presented randomly.

2.4.1. Target Embedding Algorithm – T.E.A.

It is important to manipulate target size within scenes rather than between scenes. Regarding target placement, different degrees of randomness are conceivable. If the target was placed randomly on a given trial, targets of different sizes would be located at varying eccentricities in a given scene. Moreover, the degree to which the target stands out from its neighboring regions (i.e., its visual salience) would differ widely between scenes and between target sizes per scene. Therefore, it is important to place targets of different sizes at the same location within a given scene. In principle, this common location can be picked randomly (McIlreavy et al., 2012). When using letter targets, random placement would inevitably lead to considerable differences in target salience between scenes. To reduce this variability, we developed a target embedding algorithm (T.E.A.) that took target salience into account.

While there are many methods of constructing salience maps for images of real-world scenes (Borji, Sihite, & Itti, 2013), it is widely held that simple stimulus features such as color, orientation and intensity (luminance contrast) contribute to the computation of visual salience (Itti & Koch, 2000). Using the output of a computational model of visual salience as input for our algorithm would be prohibitively computationally expensive. As a practical alternative, we used a version of root-mean-square (RMS) contrast: when stepping through the scene, the standard deviation of luminance values of all pixels in the evaluated region was divided by the mean luminance of the image. Calculating luminance contrast this way is consistent with measures of detectability in natural scenes (Bex & Makous, 2002), and with filter properties of early vision (Moulden, Kingdom, & Gatley, 1990). Moreover, it has been used in experimental studies on fixation selection in scenes (e.g., Nuthmann & Einhäuser, 2015; Reinagel & Zador, 1999).

The target was placed at an image position that caused a median RMS contrast change. To compute this, a rectangular region that was slightly larger than the target moved pixel-by-pixel through the image. The RMS contrast M_o was calculated at each position. Afterwards, the target was inserted and the RMS contrast M_w was computed at each position. By computing $\Delta C = M_w - M_o$ at each pixel, we obtained an image map comprising the contrast difference values within the image. After calculating the contrast difference map for each target size, the four resultant maps were summed together to obtain a final summed difference map. This summing acted as a way for the algorithm to compute a single location for all target sizes, as the values of each individual difference map varied slightly. The distribution of values from the summed map was computed. From the distribution different contrast levels could be selected to control the desired amount of contrast change arising from placing the letter in the scene. We used the median contrast difference as a compromise between harder (smaller contrast difference) and easier (larger contrast difference) target positions.

This final map was then probed by our algorithm to locate all pixel (i.e., potential target) positions with the median change in contrast. Some positions were eliminated by the following two criteria. First, locations within 3º of visual angle from the center were excluded from evaluation due to the central region being the initial location of both the participant's gaze and the gaze-contingent scotoma. Participants were not aware of this constraint. Second, locations at the boundaries of the image were also excluded to avoid truncation of the letter. From all remaining possible median contrast target positions, one was selected at random as the location of the target for that stimulus.

For Experiment 2, the algorithm was extended to handle multiple target letters. In this case, a new ‘TL’ contrast difference map was generated by computing:

$$\Delta C(r,c) = \sum_{\mathcal{L}_s} \left| \Delta C_{\mathcal{L}}^{[s]}(r,c) - t_{\mathcal{L}_s} \right|$$

where $\Delta C_{\mathcal{L}}^{[s]}$ is the difference map for a given font size [s] and letter $\mathcal{L} \in$ {T,L}, with [r,c] denoting the map’s rows and columns. Each of its values were then subtracted by the median contrast of a given map, denoted by $t_{\mathcal{L}_s}$. This process was repeated for both letters and all four scales before adding the resultant image maps together. By subtracting $t_{\mathcal{L}_s}$, the lowest value in this new map (with a minimum of zero) is the pixel closest to the target value $t_{\mathcal{L}_s}$, and the coordinates of this pixel defined the target position for that image. As before, central and boundary pixel positions were eliminated from consideration. Figure 1 provides an illustration by depicting the contrast difference map and the algorithmic probing.

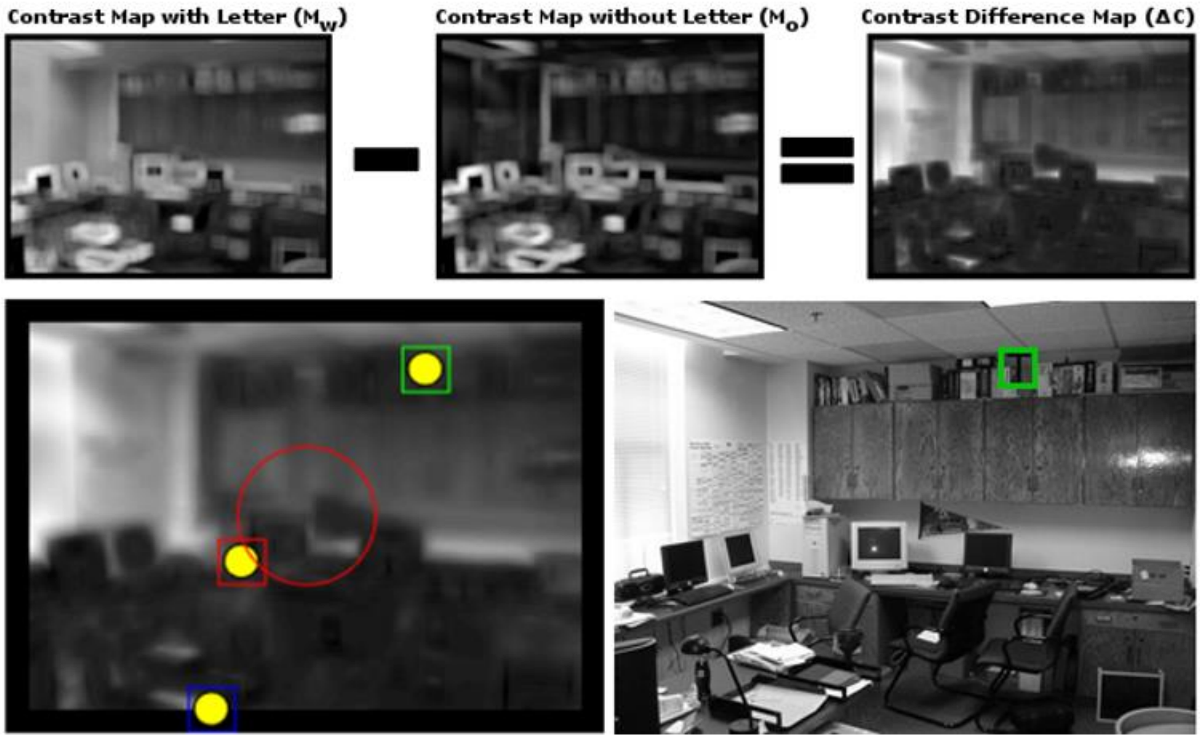

Figure 1. Illustration of the T.E.A. depicting the initial creation of the contrast difference map. The T.E.A. creates 3 contrast maps (from left to right): contrast with the letter placed at each pixel position, contrast without the letter, and the difference between them. Using the contrast difference map, the algorithm then probes the scene, excluding locations near the boundary of the screen (example: blue boxed dot) and inside the central circle (example: red boxed dot). If multiple positions are found (example: yellow dots), one satisfying the above two constraints is chosen at random (example: green boxed dot) for target insertion.

2.4.2. Creation of gaze-contingent scotoma

The foveal scotoma was created using texture-mapping and OpenGL (Open Graphics Library). This technique provides various blending operations that enable simple image combinations to take place via an image’s alpha channel (see Duchowski & Çöltekin, 2007, for details on the general technique). The scotoma was a symmetric circular mask with a radius of 1°. The scotoma size was chosen to completely obscure foveal vision. The foveal mask moved concomitantly with the participant’s gaze. To this end, the average horizontal and vertical position of the two eyes (Nuthmann, 2013, for discussion) was continuously evaluated online. Updating the display contingent on the viewer's gaze required 1 ms to receive a sample from the eye tracker, less than 1 ms to draw the image textures and up to 7 ms to refresh the screen. Thus, the display was updated depending on observers' gaze position in close to real time. A detailed account of the gaze-contingent implementation is provided in Nuthmann (2013, 2014).

There are some subtle differences between the implementation of the foveal scotoma in a previous study from our lab (Nuthmann, 2014) and here. Nuthmann (2014) used full-color images, and foveal vision was degraded by applying a very strong low-pass filter to the currently fixated scene region (the foveal scotoma was only one of six conditions with degraded vision). Moreover, a Gaussian mask was used, and the size of the scotoma was defined as the standard deviation of the two-dimensional Gaussian distribution (1.6° for the foveal scotoma, or small *Blindspot*). In the present experiments, using grayscale images, we used a circular mask drawn in gray. To avoid a sharp-boundary mask and to reduce perceptibility of slight mask position jitter, the perimeter of the circular mask was slightly faded through low-pass filtering, while the interior remained untouched. When investigating the importance of foveal vision (i.e., a relatively small region of the visual field), it seems more appropriate to define the size of the moving mask as the radius of a circle rather than the standard deviation of a Gaussian.

2.5. Procedure

At the beginning of the experiment, a 9-point calibration procedure was performed, followed by a 9-point calibration accuracy test (validation). At the beginning of each trial a fixation cross was presented at the center of the screen for 600 ms, and acted as a fixation check. The fixation check was deemed successful if gaze position, averaged across both eyes, continuously stayed within an area of 40 × 40 pixels (1.24° × 1.24°) for 200 ms. If this condition was not met, the fixation check timed out after 500 ms. In this case, the fixation check procedure was either repeated or replaced by another calibration procedure. If the fixation check was successful, the scene image appeared on the screen. Once subjects had found the target letter, they were instructed to fixate their gaze on it and press a button on the controller to end the trial (cf. Glaholt, Rayner, & Reingold, 2012; Nuthmann, 2014). In experiment 1, participants could press any button to indicate that they had found the T. Upon identifying the target in Experiment 2, observers pressed one of two triggers on the controller corresponding to either “T” or “L”. Trials timed-out 15 s after stimulus presentation if no response was made. There was an inter-trial interval of 1 s before the next fixation cross was presented.

2.6. Data analysis

The SR Research Data Viewer software with default settings was used to convert the raw data obtained by the eye tracker into a fixation sequence matrix. The behavioral and eye-movement data were further processed and analyzed using the R system for statistical computing (R Development Core Team). Figures were created using MATLAB (Figures 1 and 3) or the *ggplot2* package (Wickham, 2016) supplied in R (remaining figures).

The T.E.A. was programmed in MATLAB. When using the T.E.A. to prepare the stimulus material for Experiment 1, due to an input error the target was not inserted into an adequate scene location for eight of the scenes. Moreover, the algorithm did not catch that one scene had a different aspect ratio. As a result, nine scenes were excluded when analyzing the data from Experiment 1.

Analyses of fixation durations and saccade lengths excluded fixations that were interrupted with blinks. Analysis of fixation durations disregarded fixations that were the first or last fixation in a trial. Fixation durations that are very short or very long are typically discarded, based on the assumption that they are not determined by on-line cognitive processes (Inhoff & Radach, 1998). This precaution was not followed in the present study because the presence of a foveal scotoma may affect eye movements (e.g., fixations were predicted to be longer than normal).

Distributions of continuous response variables were positively skewed. In this case, variables are oftentimes transformed to produce model residuals that are more normally distributed. To find a suitable transformation, we estimated the optimal λ-coefficient for the Box-Cox power transformation (Box & Cox, 1964) using the *boxcox* function of the R package *MASS* (Venables & Ripley, 2002) with $y(\lambda) = (y^{\lambda} - 1)/\lambda$ if $\lambda \neq 0$ and $log(y)$ if $\lambda = 0$. For all continuous dependent variables, the optimal λ was different from 1, making transformations appropriate. Whenever λ was close to 0, a log transformation was chosen.

Non-linear transformations distort the ratio scale properties of the measured variables (Stevens, 1946). As a result, the significance of main effects can change, although this rarely happens (Kliegl, Masson, & Richter, 2010). Perhaps more importantly, some interactions can be transformed away, making them non-interpretable (Loftus, 1978; Wagenmakers, Krypotos, Criss, & Iverson, 2012). Here, we analyzed both untransformed and transformed data. As a default, we report the results for the raw untransformed data and additionally supply the results for the transformed data when they differ from the analysis of the untransformed data.

2.7. Statistical analysis using mixed models

Continuous response variables were analyzed using linear mixed-effects models (LMM), and search accuracy was analyzed using binomial generalized linear mixed-effects models (GLMM) with a logit link function. The analyses were conducted with the *lme4* package (version 1.1.-23; Bates, Maechler, Bolker, & Walker, 2015) supplied in R, using the bobyqa optimizer for LMMs, and a combination of Nelder-Mead and bobyqa for GLMMs. Separate (G)LMMs were estimated for each dependent variable.

A mixed-effects model contains both fixed-effects and random-effects terms. Fixed-effects parameters were estimated via contrast coding for which we used the nomenclature and example code provided by the UCLA Statistical Consulting Group (2011). For the factor scotoma, simple coding was used (-0.5/ +0.5, reference: no scotoma). To test effects of target size, Helmert coding was used to compare each level of the factor target size to the mean of the subsequent levels. The first contrast compared the mean of a given DV for S-targets with the mean for all larger targets (Experiment 1: M-, L-, and XL-targets). For Experiment 1, the second target-size contrast compared the mean for M-targets with the mean across L- and XL-targets, and the third contrast compared the mean for L-targets with the mean for XL-targets. Three additional interaction terms allowed for testing whether the scotoma effect was significantly different for different target-size contrasts. Given that the fixed effects were centered around zero, the intercept of the models reflected the grand mean of the DV.

The mixed models included subjects and scene items as crossed random factors. In experimental research, it is common to treat subjects as the sole random factor in the analysis (Judd, Westfall, & Kenny, 2012). However, in research on real-world scene perception and search, the variance introduced by stimulus sampling cannot be ignored (e.g., Nuthmann & Einhäuser, 2015; Nuthmann & Malcolm, 2016). We used counterbalancing to assign scene items to experimental conditions and refrained from placing the search target randomly in the scene. While algorithmic target placement reduces between-scene variability, it does not eliminate it completely. Therefore, scene items were included as random factor.

The overall mean for each subject and scene item were estimated as random intercepts. In principle, the variance-covariance matrix of the random effects not only includes random intercepts but also random slopes as well as correlations between intercepts and slopes. Random slopes estimate the degree to which each main effect and/or interaction varies across subjects and/or scene items.

To select an optimal random-effects structure for (G)LMMs, we pursued a data-driven approach using backward model selection. To minimize the risk of Type I error, we started with the maximal random-effects structure justified by the design (Barr, Levy, Scheepers, & Tily, 2013). However, the maximal random-effects structure would require estimating 72 parameters (by subject: random intercept, 7 random slopes, 28 correlation terms; by item: same as by subject). Across experiments, none of these maximal models converged (maximal number of iterations: 10^6). To reduce model complexity without taking the risk of inflating the Type I error, we proceeded to fit zero-correlation parameter (zcp) models in which the random slopes are retained but the correlation parameters are set to zero (Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017; Seedorff, Oleson, & McMurray, 2019). The full random-effects structure of the zcpLMM required 16 variance components to be estimated. This random-effects structure was backwards-reduced using the *step* function of the R package *lmerTest* (version 3.1-2; Kuznetsova, Brockhoff, & Christensen, 2017) to arrive at a model that was justified by the data. For GLMMs we report random intercept models, because random slope models did not converge. Due to the way GLMMs are estimated, model non-convergence tends to be a much larger issue than with LMMs (Seedorff et al., 2019).

LMMs were estimated using the restricted maximum likelihood criterion. GLMMs were fit by Laplace approximation. For the coded contrasts, coefficient estimates (b) and their standard errors (SE) along with the corresponding t-values (LMM: $t = b/SE$) or z-values (GLMM: $z = b/SE$) are reported. For GLMMs, p-values are additionally provided. For LMMs, a two-tailed criterion ($|t| > 1.96$) was used to determine significance at the alpha level of .05 (Baayen, Davidson, & Bates, 2008).

For the (G)LMM, data were not averaged, and modelled at the level of individual observations instead. For the data depicted in Figures 4 and 6, means were calculated for each subject, and these were then averaged across subjects.

3. Results and Discussion

The results of the two letter-in-scene search experiments are presented in three main sections. First, different measures of search accuracy were analyzed as indicators of search efficiency. Second, the time to find the target was analyzed. Behavioral search times were then decomposed based on participants' gaze data to illuminate disruptions in specific sub-processes of search (e.g., Malcolm & Henderson, 2009; Nuthmann, 2014). Third, we examined saccade amplitude and fixation duration across the viewing period as general eye-movement measures.

3.1. Search Accuracy

The first set of analyses examined the likelihood of finding the target letter in the scene. Performance for each experimental condition was divided into probabilities of "hit", "miss", and "timeout" cases (Nuthmann, 2014). Since we used a target acquisition task, a target was present on all trials. A response was scored as a "hit" if the participant indicated to have located the target by button press and his or her gaze was within the rectangular area of interest (AOI) comprising the target. In signal detection experiments, including yes-no search tasks, trials in which a non-target stimulus is identified as a target are labelled as "false alarms" (Palmer, Verghese, & Pavel, 2000; Tanner & Swets, 1954). In our experiments, incorrect responses included true false alarms where participants were fixating a non-target location and their eyes were not in the vicinity of the target when the button-press response was made. Incorrect responses also included cases where participants fixated near the target but their fixation did not fall within the AOI. Given the difficulty in distinguishing between these two cases, all trials with incorrect responses were labelled as "misses." The third category comprised trials in which the participant had not responded within 15 s. Trials with no responses were coded as "timeouts."

The size of an AOI that can be given to target stimuli is limited by (a) the spatial (in)accuracy and (im)precision of the eye tracker, and (b) the inaccuracy of the visuo-oculomotor system when targeting relatively small objects (Pajak & Nuthmann, 2013). For high-end eye-trackers like the EyeLink 1000, the minimum AOI size is about 1 to 1.5º, and the recommendation has been made to add a buffer of that size around any target object (Holmqvist & Andersson, 2017). Here, we chose to use the same AOI for all target sizes; this AOI was somewhat larger than the XL target letter with an additional 0.5º of padding to either side (2.9º × 2.9º in total).

The search accuracy results for both experiments are depicted in Figure 2. The GLMM results are summarized in Tables 1 and 2 for Experiments 1 and 2, respectively.

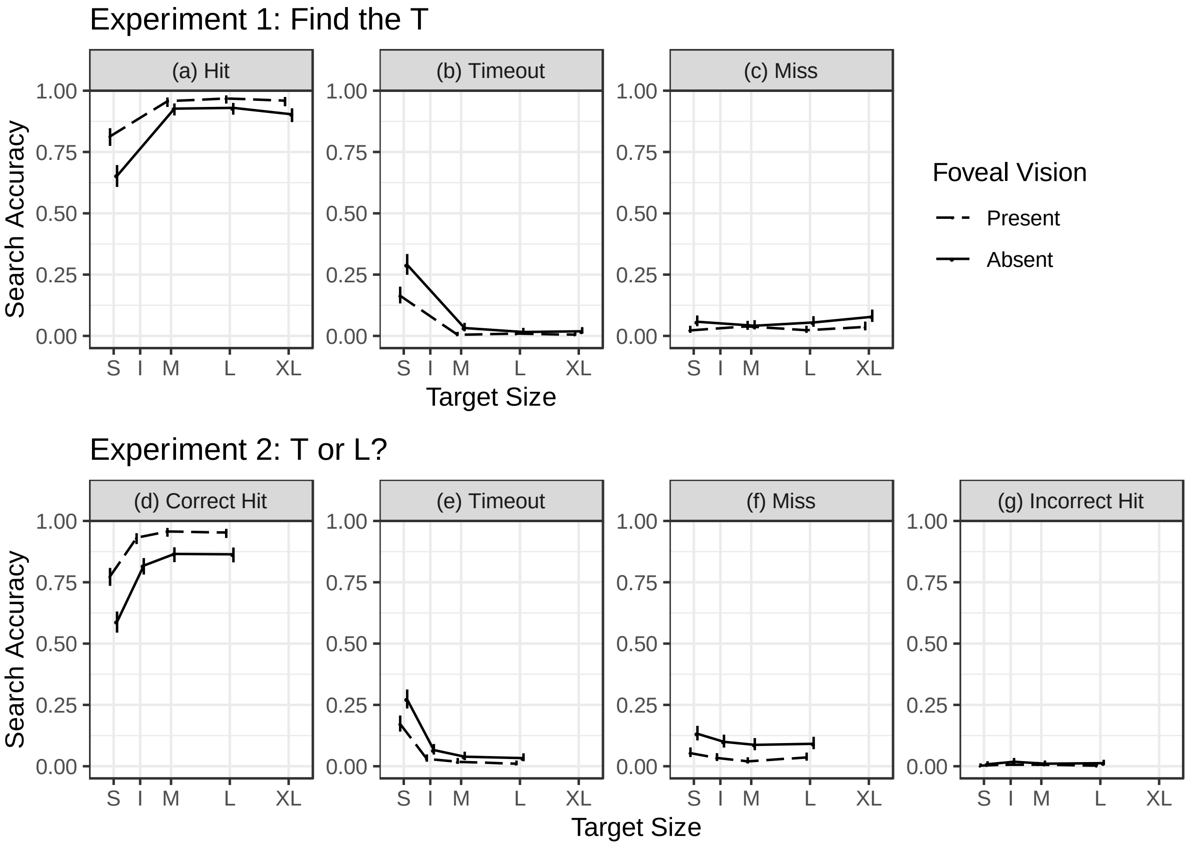

Figure 2. Measures of search accuracy for Experiment 1 (top row) and Experiment 2 (bottom row). Each column presents a designated dependent variable, which is specified in the panel

title (see text for definitions). Target sizes on the *x*-axis are described by letters (S: Small, I: Intermediate - Experiment 2 only, M: Medium, L: Large, XL: Extra Large - Experiment 1 only; see text for actual sizes in degrees of visual angle). The *x*-axis is scaled to show all target sizes across both experiments; the spacing on the *x*-axis preserves the relative distances between target sizes. Data points are binomial proportions, error bars are 95% binomial proportion confidence intervals (Wilson, 1927).

3.1.1. Experiment 1

There was a significant effect of scotoma on the probability of "hitting" the target such that participants were less likely to correctly locate and accept the target when foveal vision was not available, $b = -0.82$, $SE = 0.14$, $z = -5.88$, $p < .05$ (Figure 2a). Moreover, mean search accuracy was significantly lower for S-targets compared to the mean of M- through XL-targets, $b = -1.85$, $SE = 0.12$, $z = -15.26$, $p < .05$; the other target-size contrasts were not significant (Table 1). Scotoma and target size did not interact (Table 1). The drop in performance for small targets was due to an increase in timed out trials (Figure 2b). Timeout probability was low for all other target sizes, with or without a foveal scotoma. The probability of missing the target was low, with and without a scotoma (Figure 2c).

Insert Table 1 about here

3.1.2. Experiment 2

Experiment 2 included an additional letter recognition component (is the target a "T" or an "L"?). Therefore, we distinguished between hit trials with correct and incorrect recognition responses. The probability of incorrect hits was very low in all experimental conditions (Figure 2g). For correct hit trials, there was a significant effect of scotoma such that participants were less likely to locate and correctly identify the target without foveal vision, *b* = -1.15, *SE* = 0.11, *z* = -10.36, *p* < .05 (Figure 2d). Accuracy was lower for smaller targets; specifically, the contrasts testing S-targets and I-targets against respective larger targets were significant (S-targets vs. mean for I-, M-, and L-targets: *b* = -1.62, *SE* = 0.10, *z* = -15.96, *p* < .05; I-targets vs. mean for M- and L-targets: *b* = -0.41, *SE* = 0.14, *z* = -2.98, *p* < .05). Scotoma and target size did not interact (Table 2). The drop in performance for search without foveal vision also shows in increased probabilities of missing the target (Figure 2f) and not responding within 15 s (Figure 2e).

Insert Table 2 about here

3.2. Search time and its subcomponents

Search behavior was analyzed further for correct trials ("hits") only. Search time is the overall time taken from scene onset to a user response terminating the search. We then used participants' gaze data to divide search time into three behaviorally defined epochs: search initiation time, scanning time, and verification time (e.g., Malcolm & Henderson, 2009; Nuthmann, 2014; Nuthmann & Malcolm, 2016; Spotorno, Malcolm, & Tatler, 2015). This was done to test how the availability of foveal vision as well as the size of the target would affect different sub-processes of search. Search initiation time is the interval between scene onset and the initiation of the first saccade (i.e., initial saccade latency, or time to move). This epoch measures the time needed to choose a target location for the first saccade. Scanning time (or time to target) is the time from the first eye movement until the participant's gaze enters the target's area of interest (minus the first saccade). The scanning time measure reflects the process of localizing the target in space (Malcolm & Henderson, 2009), with longer times indicating weaker target guidance. The sum of search initiation time and scanning time represents the latency to first fixate the target (Castelhano, Pollatsek, & Cave, 2008). Our main objective in removing search initiation time from the target latency was to obtain a “clean” measure of scanning time. Finally, the verification process is indexed by the time taken from first entering the target interest area until the participant confirms their decision via button press. This component of search may also include time spent subsequently exploring other scene regions to be sure that they do not contain the target (Castelhano et al., 2008). The segmentation of search time by oculomotor behavior is visualized in Figure 3.

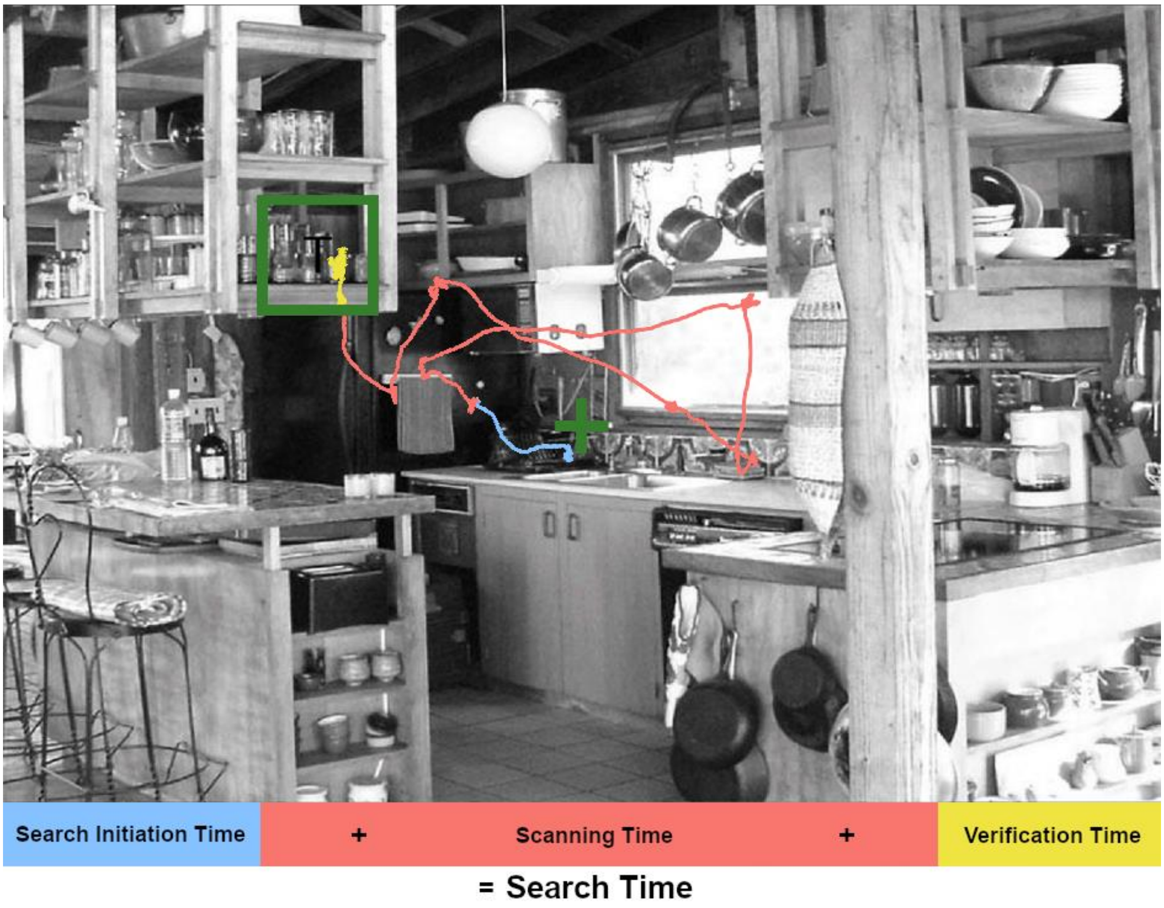

Figure 3. Gaze-based decomposition of search time. For an example search trial, the scene image is presented together with the raw gaze data from one observer (curvy lines are saccades, clustered data points are fixations). Visualizing the division of search time, blue

represents search initiation (i.e., initial saccade latency); red, scanning time; and yellow, verification time. When summed, they yield the total search time. The blue segment includes saccade execution to visualize the change in gaze position during the first eye movement. The green box is the interest area around the target letter “T”.

As outlined in the Introduction, our hypotheses concerned the scanning and verification time epochs, but not search initiation. Our main objective was to explore the degree to which scanning times and/or verification times are lengthened when foveal vision is unavailable. Moreover, we wanted to test whether target size affects the importance of foveal vision to the task; in particular, we hypothesized that verification times may reveal an interaction between target size and scotoma. A final question was whether any effects on sub-processes of search—each operating on a different timescale—were large enough to drive corresponding effects on overall search times. The results are depicted in Figure 4; the LMM results are summarized in Tables 1 and 2.

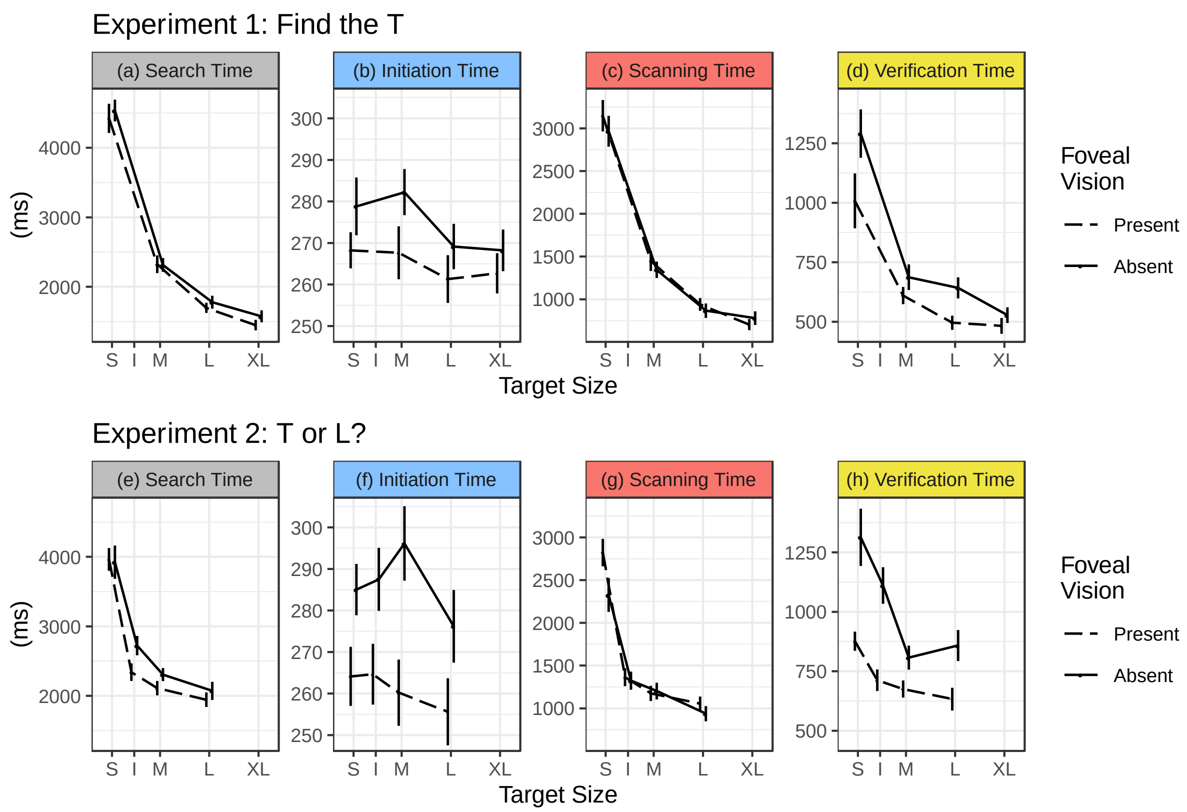

Figure 4. Search time and its three epochs for Experiment 1 (top row) and Experiment 2 (bottom row). Each column presents means obtained for a designated dependent variable (see panel title). For a given dependent variable, the *y*-axis has been normalized across plots for ease of comparison between the two experiments; but note the different *y*-axis scales for the different measures. For the three sub-processes of search (initiation, scanning, verification), subplot titles use the color scheme from Figure 3. Solid bold lines represent the scotoma condition in which foveal vision was absent; dashed lines represent the control condition in which foveal vision was present. Target sizes on the *x*-axis are described by letters (S: Small, I: Intermediate - Experiment 2 only, M: Medium, L: Large, XL: Extra Large - Experiment 1 only). The *x*-axis is scaled to show all target sizes across both experiments; the spacing on the *x*-axis preserves the relative distances between target sizes. Error bars are within-subjects standard errors, using the Cousineau-Morey method (Cousineau, 2005; Morey, 2008).

3.2.1. Experiment 1

The search-time difference between the foveal scotoma and control condition was significant, *b* = 170.38, *SE* = 80.14, *t* = 2.13; for the transformed data, this difference was not significant, *b* = 0.0013, *SE* = 0.0008, *t* = 1.62. Moreover, search times became progressively faster for larger targets, with all three target-size contrasts yielding statistically significant differences (Table 1). Scotoma and target size did not interact (Table 1).

For search initiation time, there were no significant effects (Table 1). Importantly, scanning time was not prolonged when searching with a foveal scotoma, *b* = -21.93, *SE* = 76.84, *t* = -0.29. However, scanning times became progressively faster for larger targets, with all three target-size contrasts yielding statistically significant differences (Table 1). Scotoma and target size did not interact (Table 1).

Interestingly, verification time was significantly prolonged when searching with a foveal scotoma, *b* = 167.81, *SE* = 43.87, *t* = 3.82. For larger targets, target verification was completed faster. Specifically, the contrasts testing S-targets and M-targets against respective larger targets were significant (Table 1). Moreover, the effect of scotoma was significantly stronger for S-targets compared to the mean effect of scotoma for M- through XL-targets, *b* = 285.18, *SE* = 141.38, *t* = 2.02. For the transformed data, however, this interaction was not significant, *b* = 0.092, *SE* = 0.084, *t* = 1.09.

3.2.2. Experiment 2

Search times were significantly longer with a foveal scotoma than without, *b* = 299.94, *SE* = 87.76, *t* = 3.42. Moreover, search times were faster for larger targets, with all three target-size contrasts yielding statistically significant differences (Table 2). The effect of scotoma was significantly stronger for I-targets compared to the mean effect for M- and L-targets, *b* = 300.99, *SE* = 143.18, *t* = 2.1. There were no other significant interaction effects for search time (Table 2).

In contrast to Experiment 1, the effect of scotoma on search initiation time was statistically significant, *b* = 25.6, *SE* = 13, *t* = 1.967. Moreover, the target-size contrast comparing M-targets with L-targets was significant, *b* = 11.46, *SE* = 4.96, *t* = 2.31. As in Experiment 1, scotoma and target size did not interact (Table 2).

For scanning time, there was no significant effect of scotoma, *b* = -57.5, *SE* = 55.55, *t* = -1.03. All three target-size contrasts were significant, the larger the target the faster the search (Table 2). Scotoma and target size did not interact (Table 2).

As in Experiment 1, verification time was significantly prolonged when searching with a foveal scotoma, *b* = 331.23, *SE* = 60.4, *t* = 5.48. Verification times were shorter for larger targets; specifically, the contrasts testing S-targets and I-targets against respective larger targets were significant (S-targets vs. mean for I-, M-, and L-targets: *b* = 401.5, *SE* = 81.37, *t* = 4.93; I-targets vs. mean for M- and L-targets: *b* = 196.24, *SE* = 52.33, *t* = 3.75). What about the theoretically salient interaction between scotoma and target size? The first interaction term tested whether the effect of scotoma was significantly different for S-targets compared to the mean effect of scotoma for I- through L-targets; for the untransformed data, the interaction was not significant, *b* = 214.12, *SE* = 141.31, *t* = 1.52, but for the transformed data it was, *b* = 0.022, *SE* = 0.011, *t* = 2.09. The second interaction term compared the effect of scotoma for I-targets to the mean effect of scotoma for M- and L-targets. For the untransformed data, the effect of scotoma was significantly stronger for I-targets compared to the mean effect of scotoma for M- and L-targets, *b* = 267.63, *SE* = 96.25, *t* = 2.78; for the transformed data, however, this interaction was not significant, *b* = 0.018, *SE* = 0.010, *t* = 1.77. The third interaction, comparing the effect of scotoma for M-targets to the effect of scotoma for L-targets, was not significant (Table 2).

3.2.3. Where are the eyes during the verification epoch?

Two more questions arise regarding the last component of search. Why are verification times longer for smaller targets? And what are the eyes doing when foveal analysis of the search target is not possible during fixation? In the scotoma conditions of our previous study (Nuthmann, 2014), observers had no problem selecting the target in parafoveal vision and fixating their gaze on it. Within-object fixation positions showed a central *Preferred Viewing Location* (PVL) such that most initial fixations were placed in proximity to object center (Pajak & Nuthmann, 2013). Moreover, prolonged verification times in the central-scotoma condition were due to an increased number of off-target fixations to unmask the object and to further analyze it in peripheral vision. Here, we used one common AOI for all target sizes; thus, the margin around the actual target was larger for smaller targets. Therefore, differences in oculomotor behavior for the different target-size conditions are not well captured by a binary distinction between on-target and off-target fixations. Instead, we explored fixation positions during the verification epoch through two-dimensional scatter and density plots. Since the AOI was used for data scoring, we still refer to fixations within the AOI as on-target fixations and fixations outside the AOI as off-target fixations. We summarize important aspects of a complex data pattern by comparing extreme target sizes, that is S-targets and XL-targets from Experiment 1 (Figure 5).

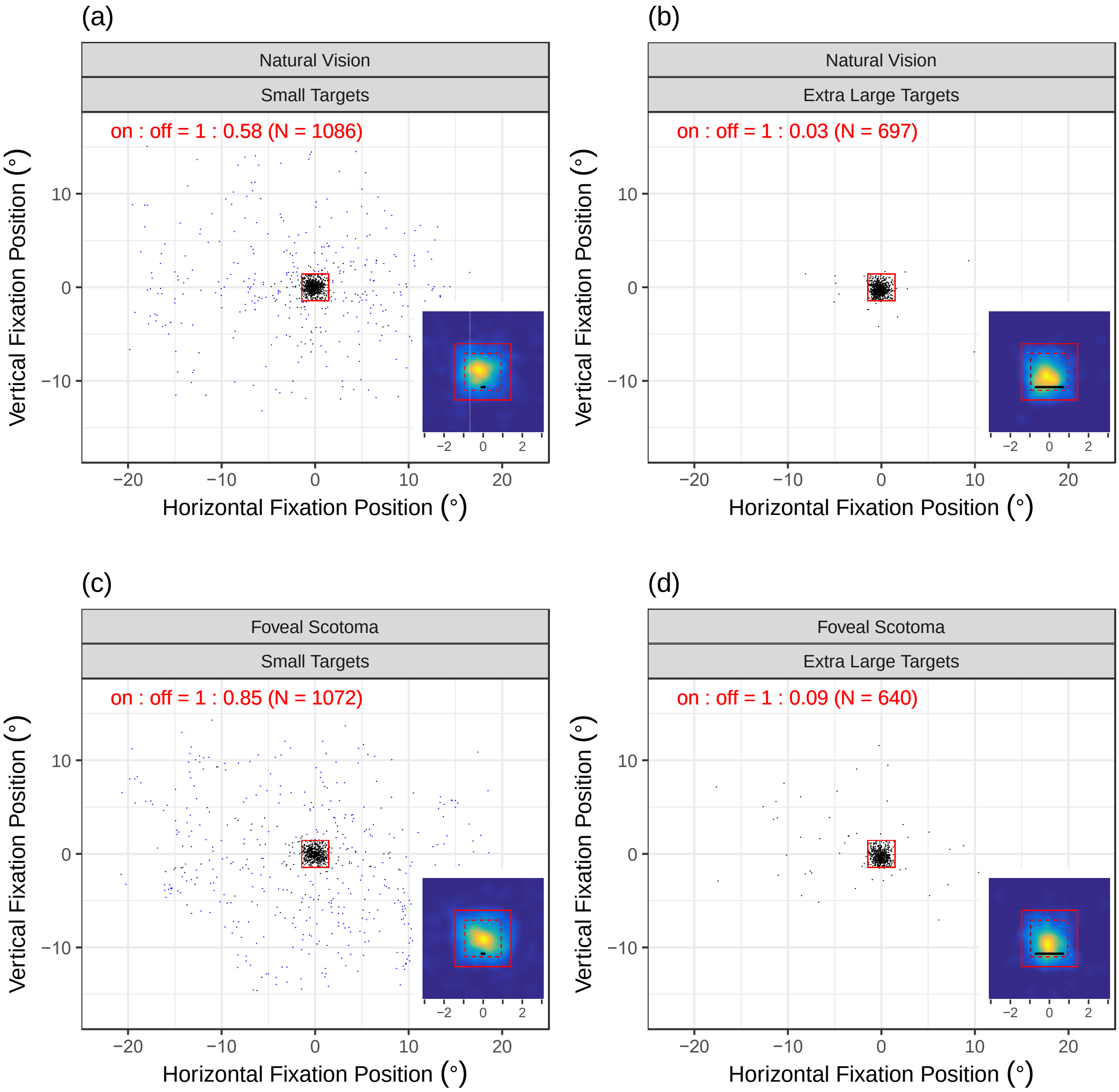

Figure 5. Analysis of fixation positions during the verification epoch. The four panels show data for small (left) and extra-large (right) targets in the natural-vision (top) and foveal-scotoma (bottom) conditions from Experiment 1. The red square with solid lines represents the area of interest (AOI) used for distinguishing between on-target fixations (within the AOI) and off-target fixations (outside the AOI). Fixation positions are expressed relative to the center of the AOI. The scatter plots show all fixations made during the verification epoch. Fixations belonging to sequences with five or more successive off-target fixations are depicted in blue rather than black. Also presented is the ratio of on-target to off-target

fixations, along with the absolute number of fixations. The inset plots zoom into the AOI region (6° × 6°) and display fixation positions as two-dimensional density plots. The frequency information is displayed as variations in color, with colors ranging from blue (few fixations), through the parula colormap to yellow (many fixations). The red square with dashed lines is the AOI encompassing the extra-large letter. The bold black line depicts the actual width of the target letter.

To unmask the target, the best strategy would be to move the eyes outside the target AOI. However, the scatter plots for XL-targets show very few off-target fixations, both with a foveal scotoma (Figure 5d) and without (Figure 5b). To overcome overplotting for on-target fixations, the inset plots zoom into the region where the AOI was situated and display density heatmaps of fixations. The fixation positions within the AOI comprise initial fixations, immediate refixations, and later revisits. The data for XL-targets show a central “hot spot”, replicating the finding of a PVL, whether foveal vision was available or not. Collectively, the data suggest that extrafoveal information from the last scanning fixation was oftentimes sufficient to identify extra-large targets when foveal vision was not available.

For the smaller target sizes, a different pattern of results emerged. There were still many more on-target than off-target fixations, but off-target fixations were much more frequent than for XL-targets. For S-targets (Figure 5a and c), off-target fixations were widely spread around the target AOI. There were also more off-target fixations with a foveal scotoma than without, as reflected by the on : off ratios. In the foveal-scotoma condition, some fixations close to the AOI may have been placed there purposely to unmask the letter and to process it in extrafoveal vision. More generally, off-target fixations are thought to be double-checking fixations to ensure that other scene regions did not contain the target (cf. Castelhano et al., 2008). It is also possible that observers did not actually attend to the target when they first encountered it and therefore kept exploring other scene regions. We cannot reliably distinguish between these alternatives. In any case, fixations far away from the AOI tended to come from trials in which longer sequences of successive off-target fixations were made before the eyes returned to the target. To highlight this, in the scatter plots all fixations that come from sequences with five or more successive off-target fixations are presented in blue rather than black (the number 5 was arbitrarily chosen). Fixation positions within the AOI showed a central PVL both in the presence and absence of foveal vision (inset plots in Figure 5a and c).

3.3. Saccade amplitudes and fixation durations

Saccade amplitudes and fixation durations were analyzed to characterize eye-movement behavior during visual search (Figure 6). In the presence of a simulated scotoma, we should observe somewhat larger saccade amplitudes and longer fixation durations (Bertera & Rayner, 2000; F. W. Cornelissen, Bruin, & Kooijman, 2005; Miellet et al., 2010; Nuthmann, 2014). We had no a priori hypotheses regarding the relationship between target size and saccade amplitudes and/or fixation durations.

For both experiments, results for mean saccade amplitudes showed a significant effect of scotoma, with larger saccades when searching with a foveal scotoma than without (Experiment 1: b = 0.49, SE = 0.07, t = 6.74, Figure 6a; Experiment 2: b = 0.74, SE = 0.1, t = 7.14, Figure 6c). In both experiments, an increase in target size was associated with shorter saccade amplitudes (Experiment 1: Table 1, Experiment 2: Table 2). For Experiment 1, the two contrasts testing S-targets and M-targets against respective larger targets were significant. For Experiment 2, the effect of target size on saccade amplitudes was driven by S-targets only. For M-Targets in Experiment 1, the effect of scotoma was significantly stronger than the mean effect of scotoma for any larger targets (Table 1). In Experiment 2, scotoma and target size did not interact (Table 2).

Fixation durations also showed a significant effect of scotoma, with longer fixation durations when searching with a foveal scotoma than without (Experiment 1: *b* = 16.57, *SE* = 3.83, *t* = 4.33, Figure 6b; Experiment 2: *b* = 18.12, *SE* = 4.73, *t* = 3.83, Figure 6d). Moreover, fixation durations tended to be shorter for larger targets (Tables 1 and 2). For Experiment 1, the contrasts testing S-targets and M-targets against respective larger targets were significant (S-targets vs. mean for M-, L-, and XL-targets: *b* = 20.48, *SE* = 2.59, *t* = 7.91; M-targets vs. mean for L- and XL-targets: *b* = 6.25, *SE* = 2.9, *t* = 2.15). For Experiment 2, mean fixation duration was significantly increased for S-Targets compared to the mean for I- through L-targets, *b* = 9.54, *SE* = 2.46, *t* = 3.88. Scotoma and target size did not interact (Tables 1 and 2).

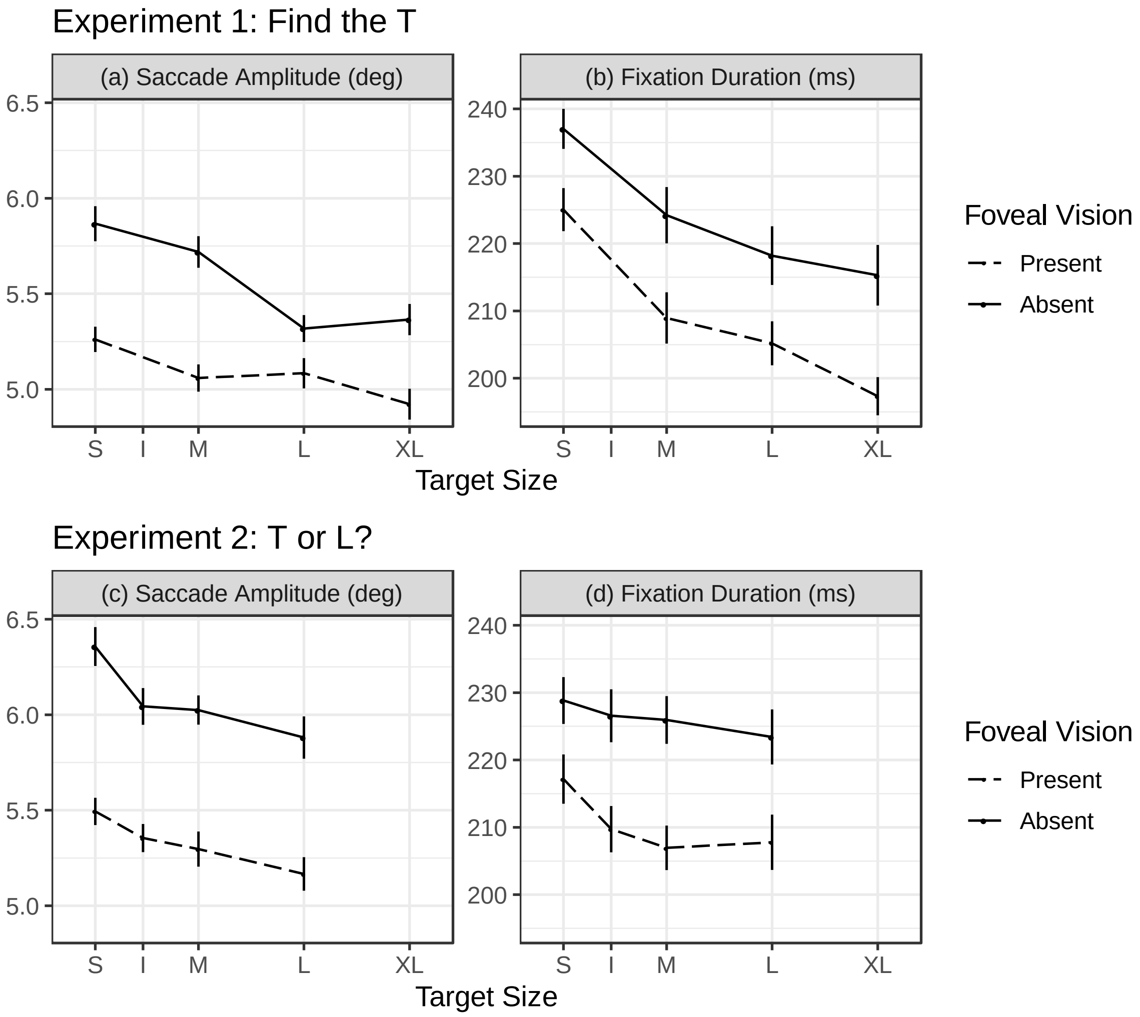

Figure 6. Mean saccade amplitudes and fixation durations for both experiments. Solid bold lines represent the scotoma condition in which foveal vision was absent; dashed lines represent the normal-vision control condition. Target sizes on the *x*-axis are described by letters (S: Small, I: Intermediate - Experiment 2 only, M: Medium, L: Large, XL: Extra Large - Experiment 1 only). The *x*-axis is scaled to show all target sizes across both experiments; the spacing on the *x*-axis preserves the relative distances between target sizes. Error bars are within-subjects standard errors, using the Cousineau-Morey method (Cousineau, 2005; Morey, 2008).

4. General Discussion

Two experiments were conducted to test the degree to which foveal vision was necessary to find context-free target letters in naturalistic scenes. A gaze-contingent moving mask (Rayner & Bertera, 1979) was used to simulate the absence of foveal vision. In Experiment 1, observers searched for the letter "T" which could occur at four different sizes. In Experiment 2, the target was either a "T" or an "L", and participants had to indicate which letter it was. If foveal vision was necessary to achieve normal search performance, the time taken to find the target should be significantly longer without foveal vision than with. Moreover, we reasoned that the importance of foveal vision may depend on the size of the search target such that foveal vision loss may be more detrimental for smaller targets.

While searching for the target without foveal vision, observers were significantly less likely to find the target than with normal vision. Our main analyses considered all correct trials ("hits"), for which we analyzed search times along with three sub-processes of search (cf. Nuthmann, 2014). With a foveal scotoma, search initiation times were significantly prolonged in Experiment 2, but not in Experiment 1. Thus, when foveal vision is not available it may take a little longer to launch the very first saccade, but this is not always the case. In both experiments, without foveal vision participants were not impaired in locating the search target in the scene (indexed by scanning time), but the process of accepting the target and responding was delayed (indexed by verification time).

Button-press search times are the sum of search initiation, scanning, and verification times. Average verification times are typically shorter than scanning times, and initiation times are shorter still. The question then arises whether small effects on faster sub-processes of search are large enough to affect total search time. For Experiment 2, search times were significantly prolonged when searching with a foveal scotoma. For Experiment 1, the effect of scotoma was significant for the untransformed data (Table 1) but not for the transformed data. Moreover, when analyzing the search-time data from a given experiment with (less appropriate) two-way repeated measures analyses of variance (*F*1 test with subject as random effect), no significant effect of scotoma was detected. In summary, the effect of a foveal scotoma on search times was fairly small and not very stable (Experiment 1).

The experiments also tested whether target size was a mediating factor for previous findings on the (un)importance of foveal vision during scene search (McIlreavy et al., 2012; Nuthmann, 2014). Not surprisingly, the data from both experiments were indicative of better search performance for larger targets, in keeping with previous research (Miellet et al., 2010). Searching the scenes for small letters proved to be a difficult task, with timed out trials and fairly long search times on successful trials. Button-press search times for medium-sized letters were similar to the ones for contextually relevant objects in our previous study (Nuthmann, 2014). The critical question was whether the size of the search target would affect the importance of foveal vision to the task (cf. Miellet et al., 2010). Specifically, we hypothesized that any detrimental effect of the foveal scotoma on the target verification process may only occur for smaller targets, or may be more pronounced for smaller than for larger targets. Significant interactions between scotoma and target size would lend support to this hypothesis. For both experiments, we found that the presence of significant interaction terms was scale dependent. In Experiment 1, the scotoma × target size 1 interaction was only significant for the untransformed data. In Experiment 2, the scotoma × target size 1 interaction was only significant for the transformed data, whereas the scotoma × target size 2 interaction was only significant for the untransformed data. Given the discrepant results for untransformed and transformed data, we do not place much interpretative weight on the interaction effects (Brysbaert & Stevens, 2018; Loftus, 1978). It is clear that any effects are small, suggesting the value of a replication study to support these conclusions.

Recent results regarding the unimportance of foveal vision when searching for spatial distortions (McIlreavy et al., 2012) or real-world objects (Nuthmann, 2014) in naturalistic scenes were surprising, given the importance of foveal vision in both reading (Rayner & Bertera, 1979) and visual search within alphanumeric displays (Bertera & Rayner, 2000). To better understand these task differences, we combined design features from letter search and scene search tasks by embedding letters into images of real-world scenes. In the following, we discuss the present results in the context of existing literature.

Foveal vision appeared to be more important in the present letter-in-scene search tasks than during object-in-scene search (Nuthmann, 2014). Neither search accuracy, nor search time or any of its components were affected by a simulated foveal scotoma in Nuthmann (2014). In contrast, search accuracy was significantly lower, and target verification time significantly prolonged in the present experiments, in which the target was a context-free letter rather than a contextually relevant object. We note that the objects used in Nuthmann (2014) were, on average, larger in size than the largest letters used here. In the present experiments, the simulated scotoma completely masked the target when observers directed their gaze to the geometrical center of the letter target, regardless of its size. Thus, the foveal scotoma could occlude the entire letter. The fact that target verification was still possible demonstrates that it could be done on the basis of extrafoveal information alone. For one, there was extrafoveal information about the target from the last scanning fixation. Moreover, during the subsequent verification epoch—and for all but the XL-targets—observers had an increased tendency to make additional off-target fixations, which increased verification time.

Foveal vision appeared to be less important in the present letter-in-scene search tasks compared to letter search in alphanumeric displays for which quite dramatic search-time costs were observed (Bertera & Rayner, 2000). In the experiment by Bertera and Rayner (2000), each array consisted of 26 letters (with 4 letters repeated) and 9 digits. Even though alphanumeric characters are overlearned stimuli, searching such displays for a designated target letter is bound to be relatively inefficient because the distractor items consisted of a large and heterogeneous set of other letters, as well as numbers. In this case, the extraction of fine detail via foveal analysis was found to be beneficial to the task (Bertera & Rayner, 2000). Compared to such unguided letter search, the availability of foveal vision may be less important for search displays in which the distractor letters are relatively similar to each other and relatively different from the target letter. For covert search[3], it has been shown that search efficiency increases as distractor-distractor similarity increases and target-distractor similarity decreases (Duncan & Humphreys, 1989). During overt search, distractors which are similar to the target receive more fixations than dissimilar distractors (Reingold & Glaholt, 2014, for review). The fact that such saccadic selectivity exists indicates that extrafoveal processing and top-down factors influence the decision about where to look next. Moreover, extrafoveal processing during the scanning epoch may also facilitate later target

---

[3] In this research, displays are smaller than the observer’s visual span such that eye movements are not essential (Findlay & Gilchrist, 2003). At the same time, unless the target was located in foveal vision, search success implies that the target has been discriminated outside foveal vision.

verification. Thus, it is an open question for future research to determine whether a stronger reliance on guidance mechanisms may render foveal vision less important.

When search takes place in real-world scenes, basic feature guidance by object properties is complemented by different types of scene guidance, in particular syntactic, semantic, and episodic guidance (Henderson & Ferreira, 2004, for review). Studying visual search in scenes poses some methodological challenges. It is unclear what to count as an "object" in a real-world scene (Neider & Zelinsky, 2008). Thus, there is no clear separation between targets and distractors. Distractor features tend to be heterogeneous (Wolfe et al., 2011) and the degree to which visual similarity relationships between objects in scenes affect guidance of gaze to search targets is hard to assess (Alexander & Zelinsky, 2012). Here, we used simple targets that were precisely specified (but varied in size) and attempted to control for their local salience. Our naturalistic scenes contained exactly one target letter to be analyzed against the scene background (Experiment 1: T, Experiment 2: T or L). The scenes in which the letter targets were embedded showed natural variation in (a) the number of elements that shared some similarity with the target (e.g., a chair leg), (b) overall target-background similarity (De Vries, Hooge, Wertheim, & Verstraten, 2013, for review), and (c) scene clutter (Rosenholtz, Li, & Nakano, 2007).

According to contemporary search theories like the target acquisition model (Zelinsky, 2008), observers compare their target representation to the search scene to obtain a map of evidence for the target at each image location. This map is then used to guide eye movements to target-like patterns in the scene (peripheral selection task). Upon fixation, incoming visual information is analyzed to decide whether this pattern is a target or a distractor (central discrimination task). The cycle of selection (guidance) and discrimination repeats until the target is found (Reingold & Glaholt, 2014; Zelinsky, Peng, Berg, & Samaras, 2013). By simulating a foveal scotoma, we selectively masked information that would otherwise be used for the central discrimination task. As a result, individual fixation durations were increased, a common finding. The foveal scotoma did not slow down the process of target localization, as measured by scanning time. This particular result highlights the importance of extrafoveal vision for target localization. It also implies that the reject decision during scanning epoch fixations was not substantially impaired, the logical conclusion being that the resolution of extrafoveal vision was sufficient to make that decision. However, making foveal vision unavailable increased the difficulty of the verification task. The accept decision during verification epoch fixations is thought to require a more complete analysis of the target candidate than the reject decision during scanning fixations (Malcolm & Henderson, 2009). In agreement with this view, in the scotoma condition observers spent more time making the accept decision, during on-target and off-target fixations.

As outlined above, there were various reasons for using letter targets. Methodologically, this design choice ensured that the effects of interest could not be mediated by other variables such as contextual constraints, target salience, or eccentricity. Importantly, when searching for a context-free letter target the scene is more than just a patterned background. Processing of scene and object relationships appears to be obligatory, in a sense that it is hard to suppress (T. H. W. Cornelissen & Võ, 2017). To extend the present findings, it would be useful to systematically explore the role played by various forms of scene guidance, using manipulations like scene inversion (Foulsham & Underwood, 2011), scene scrambling (Foulsham, Alan, & Kingstone, 2011), or pseudo-scene viewing (Luke & Henderson, 2016).

The present results replicate the finding that fixation durations and saccade amplitudes are both elevated in the presence of an artificial scotoma (Bertera & Rayner, 2000; F. W. Cornelissen et al., 2005; Miellet et al., 2010; Nuthmann, 2014; but see McIlreavy et al., 2012). The saccade amplitude adjustment reflects a tendency to fixate more locations in the non-degraded scene area than the degraded area (Nuthmann, 2014). Both global eye-movement parameters were also affected by target size; a reduction in target size was associated with both larger saccade amplitudes (see also Miellet et al., 2010) as well as longer fixation durations. These findings were unexpected, because participants had no way of knowing which target size would be displayed next, due to the randomized presentation of scenes. Over the course of scene viewing, there is a tendency for fixation durations to increase and saccade amplitudes to decrease (Pannasch, Helmert, Roth, Herbold, & Walter, 2008; Unema, Pannasch, Joos, & Velichkovsky, 2005). In our experiments, search time equates to viewing time, such that the longer search times for small targets could potentially explain the longer fixation durations (but not the larger saccade amplitudes). However, time course analyses (not reported here) provided no evidence for this. Without further research, any account of why or how observers adjust their fixation durations and saccade amplitudes in response to different target sizes (in otherwise identical scenes) remains speculative. A first step toward explaining this counterintuitive finding is to directly compare randomized and blocked presentations of different target sizes (cf. Rothkegel, Schütt, Trukenbrod, Wichmann, & Engbert, 2019).

Theories of visual search have largely been built on search for targets in arbitrary 2D arrays of items which observers searched without moving their eyes (Wolfe & Horowitz, 2017, for review). However, most real-world search takes place in structured scenes which observers explore through eye movements. The adoption of more ecologically valid stimuli has led to a new brand of image-based search theory (Eckstein, 2011, for review). Most of these models ignore that visual acuity declines systematically from the central fovea into the periphery (Nuthmann, 2014, for discussion). Moreover, visual search models usually aim at explaining the nature of peripheral selection (guidance) rather than central discrimination (Zelinsky et al., 2013). We analyzed both components and found that extrafoveal processing is not only important for selection but also for discrimination (cf. Reingold & Glaholt, 2014). The present results, together with our previous findings, inform future model building by specifying how (un)important the different regions of the visual field are for different sub-processes of search.

References

Alexander, R. G., & Zelinsky, G. J. (2012). Effects of part-based similarity on visual search: The Frankenbear experiment. *Vision Research*, *54*, 20–30. https://doi.org/10.1016/j.visres.2011.12.004

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Bertera, J. H., & Rayner, K. (2000). Eye movements and the span of the effective stimulus in visual search. *Perception & Psychophysics*, *62*(3), 576–585. https://doi.org/10.3758/BF03212109

Bex, P. J. (2010). (In) Sensitivity to spatial distortion in natural scenes. *Journal of Vision*, *10(2):23*, 1–15. https://doi.org/10.1167/10.2.23

Bex, P. J., & Makous, W. (2002). Spatial frequency, phase, and the contrast of natural images. *Journal of the Optical Society of America A-Optics Image Science and Vision*, *19*(6), 1096–1106. https://doi.org/10.1364/JOSAA.19.001096

Borji, A., Sihite, D. N., & Itti, L. (2013). Quantitative analysis of human-model agreement in visual saliency modeling: a comparative study. *IEEE Transactions on Image Processing*, *22*(1), 55–69. https://doi.org/10.1109/TIP.2012.2210727

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal*

Statistical Society Series B-Statistical Methodology, *26*(2), 211–252. https://doi.org/10.1111/j.2517-6161.1964.tb00553.x

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. https://doi.org/10.1163/156856897X00357

Brysbaert, M., & Stevens, M. (2018). Power analysis and effect size in mixed effects models: a tutorial. *Journal of Cognition*, *1*(1), 9. https://doi.org/10.5334/joc.10

Carrasco, M., Evert, D. L., Chang, I., & Katz, S. M. (1995). The eccentricity effect: Target eccentricity affects performance on conjunction searches. *Perception & Psychophysics*, *57*(8), 1241–1261. https://doi.org/10.3758/BF03208380

Carrasco, M., & Frieder, K. S. (1997). Cortical magnification neutralizes the eccentricity effect in visual search. *Vision Research*, *37*(1), 63–82. https://doi.org/10.1016/S0042-6989(96)00102-2

Carrasco, M., McLean, T. L., Katz, S. M., & Frieder, K. S. (1998). Feature asymmetries in visual search: Effects of display duration, target eccentricity, orientation and spatial frequency. *Vision Research*, *38*(3), 347–374. https://doi.org/10.1016/S0042-6989(97)00152-1

Castelhano, M. S., Pollatsek, A., & Cave, K. R. (2008). Typicality aids search for an unspecified target, but only in identification and not in attentional guidance. *Psychnomic Bulletin & Review*, *15*(4), 795–801. https://doi.org/10.3758/PBR.15.4.795

Cornelissen, F. W., Bruin, K. J., & Kooijman, A. C. (2005). The influence of artificial scotomas on eye movements during visual search. *Optometry and Vision Science*, *82*(1), 27–35.

Cornelissen, F. W., Peters, E. M., & Palmer, J. (2002). The Eyelink Toolbox: Eye tracking with MATLAB and the Psychophysics Toolbox. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 613–617. https://doi.org/10.3758/BF03195489

Cornelissen, T. H. W., & Võ, M. L.-H. (2017). Stuck on semantics: Processing of irrelevant object-scene inconsistencies modulates ongoing gaze behavior. *Attention Perception & Psychophysics*, *79*(1), 154–168. https://doi.org/10.3758/s13414-016-1203-7

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

De Vries, J. P., Hooge, I. T. C., Wertheim, A. H., & Verstraten, F. A. J. (2013). Background, an important factor in visual search. *Vision Research*, *86*, 128–138. https://doi.org/10.1016/j.visres.2013.04.010

Duchowski, A. T., & Çöltekin, A. (2007). Foveated gaze-contingent displays for peripheral LOD management, 3D visualization, and stereo Imaging. *ACM Transactions on Multimedia Computing Communications and Applications*, *3(4):24*, 1–18. https://doi.org/10.1145/1314303.1314309

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*(3), 433–458. https://doi.org/10.1037//0033-295X.96.3.433

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11(5):14*, 1–36. https://doi.org/10.1167/11.5.14

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford: University Press.

Foulsham, T., Alan, R., & Kingstone, A. (2011). Scrambled eyes? Disrupting scene structure impedes focal processing and increases bottom-up guidance. *Attention Perception & Psychophysics*, *73*(7), 2008–2025. https://doi.org/10.3758/s13414-011-0158-y

Foulsham, T., & Underwood, G. (2011). If visual saliency predicts search, then why? Evidence from normal and gaze-contingent search tasks in natural scenes. *Cognitive Computation*, *3*(1), 48–63. https://doi.org/10.1007/s12559-010-9069-9

Geisler, W. S., & Chou, K. L. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review, 102*(2), 356–378. https://doi.org/10.1037/0033-295X.102.2.356

Geringswald, F., Baumgartner, F., & Pollmann, S. (2012). Simulated loss of foveal vision eliminates visual search advantage in repeated displays. *Frontiers in Human Neuroscience*, *6*. https://doi.org/10.3389/fnhum.2012.00134

Geringswald, F., & Pollmann, S. (2015). Central and peripheral vision loss differentially affects contextual cueing in visual search. *Journal of Experimental Psychology: Learning Memory and Cognition*, *41*(5), 1485–1496. https://doi.org/10.1037/xlm0000117

Glaholt, M. G., Rayner, K., & Reingold, E. M. (2012). The mask-onset delay paradigm and the availability of central and peripheral visual information during scene viewing. *Journal of Vision*, *12(1):9*, 1–19. https://doi.org/10.1167/12.1.9

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: Psychology Press.

Holmqvist, K., & Andersson, R. (2017). Eye tracking: A comprehensive guide to methods, paradigms and measures. Lund, Sweden: Lund Eye-Tracking Research Institute.

Hooge, I. T. C., & Erkelens, C. J. (1999). Peripheral vision and oculomotor control during visual search. *Vision Research*, *39*(8), 1567–1575. https://doi.org/10.1016/S0042-6989(98)00213-2

Inhoff, A. W., & Radach, R. (1998). Definition and computation of oculomotor measures in the study of cognitive processes. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 29–53). Oxford: Elsevier Science Ltd. https://doi.org/10.1016/B978-008043361-5/50003-1

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, *40*(10–12), 1489–1506. https://doi.org/10.1016/S0042-6989(99)00163-7

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*, 14.

Kliegl, R., Masson, M. E. J., & Richter, E. M. (2010). A linear mixed model analysis of masked repetition priming. *Visual Cognition*, *18*(5), 655–681. https://doi.org/10.1080/13506280902986058

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, *6*(3), 312–319. https://doi.org/10.3758/BF03197461

Luke, S. G., & Henderson, J. M. (2016). The influence of content meaningfulness on eye movements across tasks: Evidence from scene viewing and reading. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2015.00257

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision, 9(11):8*, 1–13. https://doi.org/10.1167/9.11.8

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing Type I error and power in linear mixed models. *Journal of Memory and Language*, *94,* 305–

315. https://doi.org/10.1016/j.jml.2017.01.001

McIlreavy, L., Fiser, J., & Bex, P. J. (2012). Impact of simulated central scotomas on visual search in natural scenes. *Optometry and Vision Science*, *89*(9), 1385–1394. https://doi.org/10.1097/OPX.0b013e318267a914

Miellet, S., Zhou, X., He, L., Rodger, H., & Caldara, R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *Journal of Vision*, *10(6):21*, 1–18. https://doi.org/10.1167/10.6.21

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, *4*, 61–64.

Moulden, B., Kingdom, F., & Gatley, L. F. (1990). The standard deviation of luminance as a metric for contrast in random-dot images. *Perception*, *19*(1), 79–101. https://doi.org/10.1068/p190079

Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition*, *16*(1), 1–10. https://doi.org/10.1080/13506280701381691

Nuthmann, A. (2013). On the visual span during object search in real-world scenes. *Visual Cognition*, *21*(7), 803–837. https://doi.org/10.1080/13506285.2013.832449

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*(1), 342–360. https://doi.org/10.1037/a0033854

Nuthmann, A., & Einhäuser, W. (2015). A new approach to modeling the influence of image features on fixation selection in scenes. *Annals of the New York Academy of Sciences*, *1339*(1), 82–96. https://doi.org/10.1111/nyas.12705

Nuthmann, A., & Malcolm, G. L. (2016). Eye guidance during real-world scene search: The role color plays in central and peripheral vision. *Journal of Vision*, *16(2):3*, 1–16. https://doi.org/10.1167/16.2.3

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. https://doi.org/10.1167/13.5.2

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10–12), 1227–1268. https://doi.org/10.1016/S0042-6989(99)00244-8

Pannasch, S., Helmert, J. R., Roth, K., Herbold, A.-K., & Walter, H. (2008). Visual fixation durations and saccade amplitudes: Shifting relationship in a variety of conditions. *Journal of Eye Movement Research*, *2(2):4*, 1–19. https://doi.org/10.16910/jemr.2.2.4

Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*(4417), 468–469. https://doi.org/10.1126/science.504987

Reichle, E. D., & Reingold, E. M. (2013). Neurophysiological constraints on the eye-mind link. *Frontiers in Human Neuroscience*, *7*, 361. https://doi.org/10.3389/fnhum.2013.00361

Reinagel, P., & Zador, A. M. (1999). Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, *10*(4), 341–350. https://doi.org/10.1088/0954-898X/10/4/304

Reingold, E. M., & Glaholt, M. G. (2014). Cognitive control of fixation duration in visual search: The role of extrafoveal processing. *Visual Cognition*, *22*(3–4), 610–634. https://doi.org/10.1080/13506285.2014.881443

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, *7(2):17*, 1–22. https://doi.org/10.1167/7.2.17

Rothkegel, L. O. M., Schütt, H. H., Trukenbrod, H. A., Wichmann, F. A., & Engbert, R. (2019). Searchers adjust their eye-movement dynamics to target characteristics in natural scenes. *Scientific Reports*, *9*, 1635. https://doi.org/10.1038/s41598-018-37548-w

Rovamo, J., & Virsu, V. (1979). An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, *37*(3), 495–510. https://doi.org/10.1007/BF00236819

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*(6), 1067–1082. https://doi.org/10.3758/BF03211940

Seedorff, M., Oleson, J., & McMurray, B. (2019). Maybe maximal: Good enough mixed models optimize power while controlling Type I error. *PsyArXiv*. https://doi.org/10.31234/osf.io/xmhfr

Shen, J., Reingold, E. M., Pomplun, M., & Williams, D. E. (2003). Saccadic selectivity during visual search: The influence of central processing difficulty. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 65–88). Amsterdam: Elsevier. https://doi.org/10.1016/B978-044451020-4/50005-0

Spotorno, S., Malcolm, G. L., & Tatler, B. W. (2015). Disentangling the effects of spatial inconsistency of targets and distractors when searching in realistic scenes. *Journal of Vision*, *15(2):12*, 1–21. https://doi.org/10.1167/15.2.12

Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, *103*(2684), 677–680. https://doi.org/10.1126/science.103.2684.677

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, *11(5):13*, 1–82. https://doi.org/10.1167/11.5.13

Tanner, W. P., & Swets, J. A. (1954). A decision-making theory of visual detection. *Psychological Review*, *61*(6), 401–409. https://doi.org/10.1037/h0058700

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*(4), 766–786. https://doi.org/10.1037/0033-295X.113.4.766

Unema, P. J. A., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition*, *12*(3), 473–494. https://doi.org/10.1080/13506280444000409

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York: Springer. https://doi.org/10.1007/978-0-387-21706-2

Wagenmakers, E.-J., Krypotos, A.-M., Criss, A. H., & Iverson, G. (2012). On the interpretation of removable interactions: A survey of the field 33 years after Loftus. *Memory & Cognition*, *40*(2), 145–160. https://doi.org/10.3758/s13421-011-0158-0

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis* (2d ed.). New York: Springer.

Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, *22*(158), 209–212. https://doi.org/10.1080/01621459.1927.10502953

Wolfe, J. M. (2014). Approaches to visual search: Feature integration theory and guided search. In A. C. Nobre & S. Kastner (Eds.), *Oxford Handbook of Attention* (pp. 11–55). New York: Oxford University Press.

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*(6), 1650–1671. https://doi.org/10.3758/s13414-011-0153-3

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search.

Nature Human Behaviour, *1*, article number 0058. https://doi.org/10.1038/s41562-017-0058

Wolfe, J. M., O'Neill, P., & Bennett, S. C. (1998). Why are there eccentricity effects in visual search? Visual and attentional hypotheses. *Perception & Psychophysics*, *60*(1), 140–156. https://doi.org/10.3758/BF03211924

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787–835. https://doi.org/10.1037/a0013118

Zelinsky, G. J., Peng, Y., Berg, A. C., & Samaras, D. (2013). Modeling guidance and recognition in categorical search: Bridging human and computer object detection. *Journal of Vision*, *13(3):30*, 1–20. https://doi.org/10.1167/13.3.30