

# Natural image clutter degrades overt search performance independently of set size

Yelda Semizer

Department of Psychology, Rutgers University,  
New Brunswick, NJ, USA



Melchi M. Michel

Department of Psychology, Rutgers University,  
New Brunswick, NJ, USA



**Although studies of visual search have repeatedly demonstrated that visual clutter impairs search performance in natural scenes, these studies have not attempted to disentangle the effects of search set size from those of clutter per se. Here, we investigate the effect of natural image clutter on performance in an overt search for categorical targets when the search set size is controlled. Observers completed a search task that required detecting and localizing common objects in a set of natural images. The images were sorted into high- and low-clutter conditions based on the clutter metric by Bravo and Farid (2008). The search set size was varied independently by fixing the number and positions of potential targets across set size conditions within a block of trials. Within each fixed set size condition, search times increased as a function of increasing clutter, suggesting that clutter degrades overt search performance independently of set size.**

result of imprecise specification of the likely target location. This type of position uncertainty increases with the number of potential target locations and is called *extrinsic* position uncertainty. However, even when the observer is familiar with the bookshelf and knows the order of its books, she or he might still have a hard time localizing the book in the visual periphery. This uncertainty is a result of the limitations intrinsic to the visual system, such as the limitations of peripheral vision and visual memory, and it is called *intrinsic* position uncertainty.

Regardless of whether it is extrinsic or intrinsic, position uncertainty impairs performance for detecting, discriminating, and localizing stimuli. This is indicated by decreases in detection and localization accuracy (Burgess & Ghandeharian, 1984; Eckstein, Thomas, Palmer, & Shimozaki, 2000), by increases in detection thresholds (Cohn & Wardlaw, 1985; Palmer, Verghese, & Pavel, 2000), and by increases in search times (Egeth, Atkinson, Gilmore, & Marcus, 1973; Treisman & Gelade, 1980). Although research on the effects of position uncertainty has typically focused on extrinsic sources of uncertainty (e.g., Bochud, Abbey, & Eckstein, 2004; Burgess & Ghandeharian, 1984; Swensson & Judy, 1981), a few studies have explicitly focused on intrinsic sources (e.g., Michel & Geisler, 2011; Pelli, 1985; Tanner, 1961). Evidence from these studies, and from studies of visual crowding (e.g., Bouma, 1970; Levi, 2008; Pelli et al., 2007; Pelli, Palomares, & Majaj, 2004) suggests that the ability to identify and localize features declines systematically in the periphery. Indeed, position uncertainty has been repeatedly implicated as a primary contributor to crowding (Krumhansl & Thomas, 1977; Popple & Levi, 2005; Wolford, 1975). For example, similar to crowding (Bouma, 1970; Levi, 2008; Levi, Hariharan, & Klein, 2002), intrinsic position uncertainty also increases approximately linearly with eccentricity (Michel & Geisler, 2011). Moreover, the eccentricity-dependent

## Introduction

Interacting with the world involves, as frequent and ubiquitous subtasks, the detection and localization of objects in our visual environment. These subtasks are called *visual searches*. One fundamental property common to all visual searches is uncertainty regarding the positions of target objects. This study examines the properties of the visual environment and of the visual system that contribute to this position uncertainty. In particular, our goal is to investigate how visual clutter affects performance when observers search natural images for categorical targets.

Position uncertainty can be due to either extrinsic or intrinsic sources. For example, an observer searching an unfamiliar bookshelf for a particular book will probably have some uncertainty about the location of the book. In this case, (i.e., when the observer does not know the book's location a priori), the uncertainty is a

Citation: Semizer, Y., & Michel, M. M. (2019). Natural image clutter degrades overt search performance independently of set size. *Journal of Vision*, 19(4):1, 1–16, <https://doi.org/10.1167/19.4.1>.



effects of position uncertainty seem to persist in search tasks requiring eye movements (i.e., overt search tasks; Semizer & Michel, 2017).

As an inherent property of the observer's visual system, intrinsic position uncertainty cannot be experimentally controlled. However, its effect on performance can be observed by manipulating the visual environment. In a recent study, Semizer and Michel (2017) introduced an experimental technique that modulates the effects of intrinsic uncertainty independently of extrinsic uncertainty by manipulating the distribution of clutter in synthetic noise displays. Using this technique, the authors showed that intrinsic position uncertainty substantially limits overt search performance and that its effects are especially evident when the amount of extrinsic uncertainty is controlled. Does this result generalize to real-world searches?

In many ways, synthetic visual stimuli have been incredibly useful for vision research. Synthetic stimuli provide researchers with a great deal of flexibility and control, enabling them to manipulate individual stimulus features and to determine how these contribute to performance in a variety of tasks. In visual search, for example, measuring performance in synthetic search displays has allowed researchers to discover how observers use information about peripheral target visibility to select fixations (Geisler, Perry, & Najemnik, 2006; Najemnik & Geisler, 2005; Najemnik & Geisler, 2008; Michel & Geisler, 2009; Verghese, 2012; Zhang & Eckstein, 2010), how intrinsic position uncertainty and clutter in the periphery degrade performance (Michel & Geisler, 2011; Rosenholtz, Huang, Raj, Balas, & Ilie, 2012; Semizer & Michel, 2017), how the template for known search targets is structured (Eckstein, Beutter, Pham, Shimozaki, & Stone, 2007), and how observers integrate information about the target across fixations (Caspi, Beutter, & Eckstein, 2004; Kleene & Michel, 2018), all while controlling extraneous properties of the search display (e.g., spectral spatial frequency statistics, environmental contingencies, target location probabilities, etc.) in ways that would be difficult or impossible with natural scenes. However, their highly controlled nature means that synthetic displays may provide only limited insight into how observers search in naturalistic settings.

For example, the search targets used in synthetic displays typically exhibit very little variability across trials, and observers are therefore assumed to represent them with little uncertainty. In contrast, the targets of natural searches typically exhibit many sources of variability. Objects in natural scenes appear in various positions and orientations, occlude one another, and change appearance depending on the lighting conditions. Moreover, individual exemplars may vary considerably within a natural object category. These sources of variability introduce additional uncertainty

that might overwhelm any effects of intrinsic uncertainty on search performance. Thus, it is important to verify that the factors that explain search performance in synthetic displays generalize to account for searches in more naturalistic displays.

One of the major challenges associated with naturalistic tasks in the context of visual search is to quantify the amount of clutter in natural images. Unlike in artificial displays, clutter cannot be directly manipulated in natural images. However, a variety of models have been proposed to quantify scene clutter. These include edge density (Mack & Oliva, 2004), feature congestion (Rosenholtz, Li, Mansfield, & Jin, 2005; Rosenholtz, Li, & Nakano, 2007), subband entropy (Rosenholtz et al., 2007), the scale invariant clutter measure (Bravo & Farid, 2008), the color-clustering clutter (C3) model (Lohrenz, Trafton, Beck, & Gendron, 2009), and the proto-object model (Yu, Samaras, & Zelinsky, 2014). Using these measures, several studies have shown that clutter degrades performance for search in various types of naturalistic displays including geographic maps (Rosenholtz et al., 2007; Lohrenz et al., 2009), quasi-realistic scenes (Neider & Zelinsky, 2011), natural scenes (Henderson, Chanceaux, & Smith, 2009, but also see Asher, Tolhurst, Troscianko, & Gilchrist, 2013), images displaying contents of bags (Bravo & Farid, 2008), and photo collages of objects (Bravo & Farid, 2004; Bravo & Farid, 2008).

However, these findings confound different potential sources of position uncertainty. As a scene becomes increasingly cluttered, the number of possible target locations (i.e., set size) also increases. This increase in set size augments the position uncertainty because of extrinsic sources. At the same time, because of intrinsic sources of position uncertainty, the ability to exclude irrelevant signals in the periphery decreases in highly cluttered scenes (Michel & Geisler, 2011; Semizer & Michel, 2017). These two concurrent effects of clutter make it challenging to separate the contributions of extrinsic versus intrinsic uncertainty on performance in highly cluttered images.

The goal of the current study was to separate out the contributions of intrinsic versus extrinsic sources of position uncertainty and to characterize them in a naturalistic search task that requires searching natural images for categorical targets. We approached this goal by controlling and manipulating set size independently of clutter, as in Semizer and Michel (2017). Instead of imposing synthetic clutter, we used an existing clutter measure (Bravo & Farid, 2008), chosen for its efficiency and its demonstrated correlation with search performance, to quantify the existing clutter in a set of natural images. The images were sorted into high- and low-clutter conditions based on this clutter measure. The "relevant set size" (Palmer, 1994; Palmer, 1995),

which governed the extrinsic position uncertainty, was varied independently by manipulating the number and positions of cues indicating potential target locations. Within each fixed set size condition, search times increased as a function of increasing clutter, suggesting that clutter degrades overt search performance independently of set size.

## Methods

### Observers

A total of 25 observers participated in the study. One of the observers was an author; the remaining observers were naïve to the purpose of the experiment and received compensation for their participation. All observers had normal or corrected-to-normal vision.

### Apparatus

Stimuli were presented on a 22-in. Philips 202P4 CRT monitor at 100 Hz. The resolution was set to  $1,280 \times 1,024$  pixels. Observers were seated 70 cm away from the display so that the display subtended  $15.8^\circ \times 21.1^\circ$  of visual angle. The stimuli displays were programmed using MATLAB software (MathWorks, Natick, MA) and the Psychophysics Toolbox extensions (Brainard, 1997). Observers' eye movement signals were monitored and recorded using an Eyelink 1000 infrared eye tracker (SR Research, Kanata, Ontario, Canada) at 1000 Hz. Head position was stabilized using a forehead and chin rest.

### Stimuli

Images of natural scenes often contain contextual information that effectively reduces the search set size (Castelhano & Heaven, 2011; Neider & Zelinsky, 2006; Oliva & Torralba, 2006; Torralba, Oliva, Castelhano, & Henderson, 2006). To minimize this contextual information, we chose a set of images displaying the contents of bags in arbitrary arrangements (see Figure 1). These images were retrieved from the “What’s in your bag?” group on Flickr.<sup>1</sup> We selected five of the most common objects in the image set (cellphones, glasses, iPods, keys, and pens/pencils) to serve as the categorical search targets. If a target object was present in the image, it was either present as a single instance or, in the case of collective objects, as a single group of instances in close proximity (e.g., keys attached to a keychain). There was never more than one instance or group of the target object present in the image.

### Creating the image data set

The image data set was created by processing raw images in four separate stages: initial filtering, transformation, labeling, and selection. Each stage is described in detail next.

*Initial filtering stage:* Images were downloaded and subsequently checked for duplicates and quality (e.g., blurs, artifacts, etc.). We avoided scaling the size of small images up to preserve image quality. Therefore, images whose maximum dimension was smaller than the height of the stimulus window (1,024 pixels) were excluded.

*Transformation stage:* The clutter measure used in our experiment is sensitive to the image size (see the Measuring Clutter section). To control for any potential effects of image size on quantifying clutter, we resized the minimum dimension to 1,024 pixels.

Next, we considered the variability in color across images. To control for the effects of color on performance, colored images were converted to gray-scale intensity images by removing the hue and saturation information while keeping the luminance information. RGB values were converted to gray-scale values by computing a weighted sum of the channels using the intensity transformation:

$$I = 0.299R + 0.587G + 0.114B, \quad (1)$$

where  $I$  represents the gray-scale intensity and  $R$ ,  $G$ , and  $B$  corresponds to red, blue, and green channels, respectively.<sup>2</sup> Also, to control the variability in luminance and contrast levels across images, the average luminance of each image was set to  $40 \text{ cd/m}^2$ , and its contrast level (root mean square) was adjusted to 0.4. Then, the clutter was computed for each image (see the Measuring Clutter section). The distribution of clutter was similar across search images containing different target object categories (see Figure 2, left panel).

*Labeling stage:* Images were annotated by labeling the category of potential target objects present in them. Then, target locations were marked by drawing circumscribing polygons around the target objects. The vertices of these polygons were recorded. At the end of this stage, each image was associated with an annotation consisting of a list of target objects within the image, a list of vertices describing the circumscribing polygon for each target object, and the clutter value for the image.

*Selection stage:* For each of five target categories, 800 test images were selected. The target object was present in only half of these images. Test images were chosen based on the following criteria.

First, to label images as high and low clutter, we computed the median of the clutter distribution. Images with clutter values higher than the median were marked as high clutter, whereas images with clutter



Figure 1. Example displays for the low-clutter (left) and the high-clutter (right) conditions with keys as the search target. Keys are located near the center in both images. Images are retrieved from the “What’s in your bag?” group on <https://www.flickr.com>.

values lower than the median were marked as low clutter.

Next, we expected that target size might affect search performance in target present images. To control for any size effects, we first measured the size of each target object by computing the area of its circumscribing polygon. Target size varied depending on the target category (see Figure 2, right panel). For example, on average, cellphones were larger than keys. To limit the effects of unusually sized objects, we restricted the variability in target size by including images only if  $t \in [1/4m, 4m]$ , where  $t$  is the target size and  $m$  is the median target size.

A final inclusion criterion considered the variants of targets. If we suspected that observers might not be familiar with a particular variant of target object, images displaying that variant were not selected. For example, in the case of cellphones, images did not include any flip phones. Similarly, in the case of iPods, only images with iPods with a particular shape, a rectangular screen at the top, and a circular area at the bottom were included. Further, images with objects that looked highly similar to targets were also excluded. For example, images containing an iPod touch (which might look like an iPhone to the observer) were

excluded. Similarly, in the case of pens, we excluded images that included makeup pencils.

At the end of this process, 800 images were selected for each target category. Four hundred test images were selected for the target-present trials by prioritizing the amount of clutter and checking for the criteria listed above, and another 400 images without the target object were selected for the target-absent trials.

### Preparing the search stimuli

For their presentation in the search task, individual images in each of the two clutter conditions were randomly assigned to either the low (5 locations) or high (13 locations) set size conditions. Potential target locations were marked by small circular cues overlaid on the image. To minimize uncertainty about cue locations, we used large cues ( $0.25^\circ$  in diameter) that were red on a gray-scale image; placed the cues on a regular hexagonal grid, with spacings of  $8.0^\circ$  and  $5.9^\circ$  for set sizes 5 and 13, respectively; and made them continuously visible on the screen across each block. The average distance of the cue locations from the origin was set to be two-thirds of the radius of the stimulus circle. Images were shifted and rotated so that

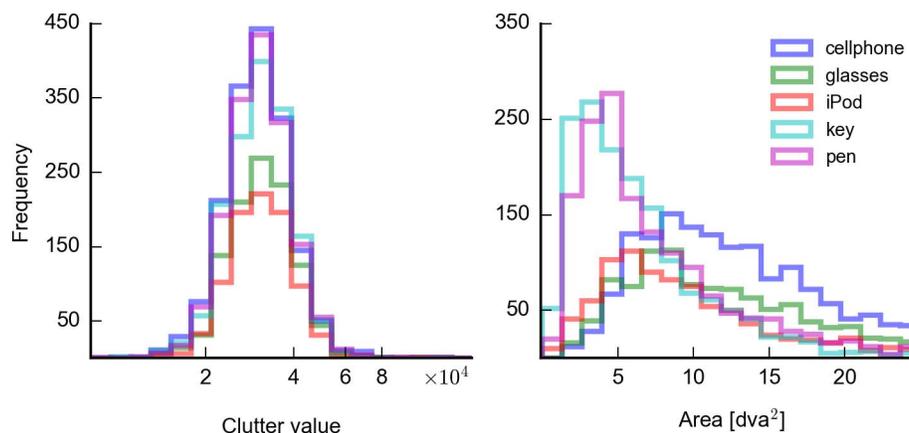


Figure 2. Distribution of clutter values (left) and target size (right) for each target category.

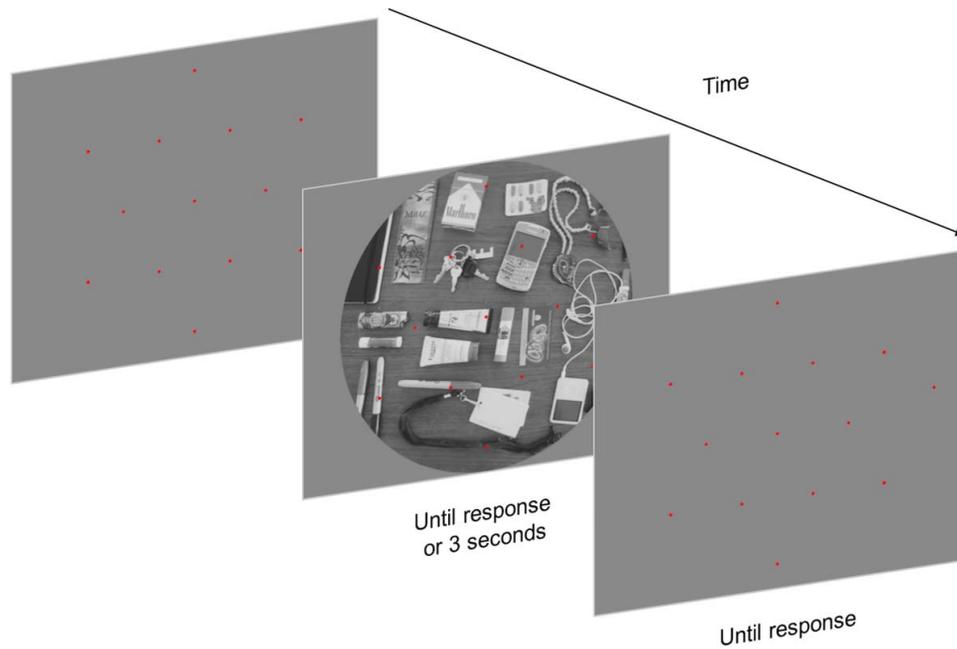


Figure 3. Search task sequence for a trial with keys as the search target. The small red cue markers, which were continuously visible, represent the potential target locations ( $N = 13$ ). The keys are located within the top left quadrant of the image.

only one of these cues appeared within the circumscribing polygon associated with the correct target location. Images were presented in a circular region,  $24^\circ$  in diameter. This region was chosen with the constraint that it contained the target object. Finally, the area around the circular region was set to uniform gray. The final form of images used as stimuli in the search task is shown in Figure 3.

### Measuring clutter

We quantified image clutter using a modified version of the clutter measure described in Bravo and Farid (2008). We chose this clutter measure because it has been shown to successfully predict search times in a similar set of images. In addition, this measure is computationally efficient and scale invariant. Briefly, this measure estimates the amount of clutter in an image as a function of the relationship between the number of “segments” in an image and the scale of segmentation. The details of the segmentation procedure and our implementation of the clutter measure are described below.

**Segmentation algorithm:** To count the number of segments in each image, we used the graph-based segmentation algorithm introduced by Felzenszwalb and Huttenlocher (2004). This algorithm segments the image by considering the variability of nearby regions. In particular, it draws boundaries between regions based on pairwise comparisons of the intensities within and across regions. The threshold for drawing these boundaries is controlled by a scale parameter  $k$ . Larger

$k$  leads the algorithm to favor larger regions and results in a smaller number of segments. The algorithm produces perceptually reasonable segments (e.g., see Figure 4), and it runs at a high speed in practice.

Our search stimuli consisted of different cropped sections of the images for different target categories. We wanted the clutter estimates to be robust to minor changes of the position in the image. To get more stable estimates of clutter, we created random sections from the images, counted segments for each section at multiple scales, and then computed the geometric mean of segment counts across sections at each scale. At the end of this process, each image was associated with a segment count for each scale.

**Clutter measure:** We measured the clutter in each image by characterizing the relationship between the scale of segmentation  $k$  and the number of segments for that scale  $y(k)$ . We determined this relationship empirically by varying the scale parameter across a range of values, applying the segmentation algorithm, and counting the resulting number of segments. Figure 4 shows examples of segmented images and the number of segments at several scales of segmentation. For any given image, the number of segments is log-linearly related to the scale of segmentation, such that

$$\ln y(k) = \alpha + \beta \ln k, \quad (2)$$

where  $\ln$  represents the natural logarithm.

The slope of this relationship is approximately constant ( $\beta \approx -0.69$ ), but the intercept  $\alpha$  varies across images. In particular, for any setting of the scale parameter, highly cluttered images tend to have more

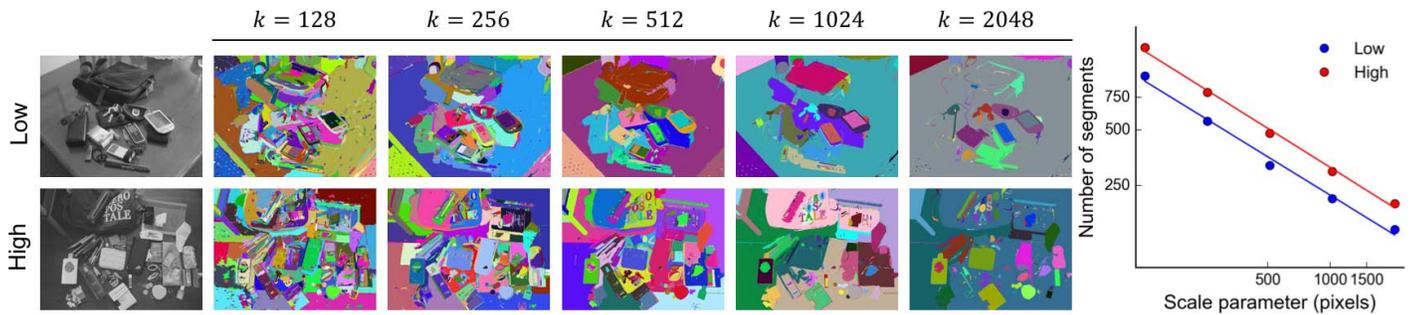


Figure 4. Example of segmented images of a low-clutter (top) and high-clutter (bottom) image at five of the 12 possible values of the scale parameter ( $k \in [90, 4096]$ ). Color is used only to show the segmented regions in the image. The plot on the right shows the number of segments as a function of the scale parameter for each image. Points represent the raw number of segments, whereas the lines represent the log-linear fits.

segments than the minimally cluttered images. Therefore, we used the intercept of each image to quantify its clutter.

To obtain robust estimates of these log-linear relationships for each image, we (a) randomly sampled ten  $1,024 \times 1,024$  sections of the image, (b) computed the segment counts for each of these sections across a range of scales ( $k \in \{90, 128, 181, 256, 362, 512, 724, 1024, 1448, 2048, 2896, 4096\}$ ), and (c) computed the intercept of the log-linear fit using a least-squares procedure. The slope was computed as the average least-squares slope for all of the images in the data set ( $N = 4,953$ ), and the intercepts for individual images were fitted with this average slope held constant.

Our implementation of this clutter metric differed from the clutter metric on which it is based (Bravo & Farid, 2008) in two ways: First, we evaluated the least-squares fit in log units, in which the power law functions are linear. This was done to make the model residuals more homoscedastic and thereby make the fitting more robust to outliers. Second, we sampled multiple sections from each image before segmenting them and computed the fit using the set of segment counts obtained for all of these sections. This resulted in fits that were robust to the small changes in cropping boundaries that occurred when the images were repositioned to align the target location with the grid of cue positions.

To evaluate the generalizability/robustness of our clutter measurements, we also quantified clutter using alternative clutter measures including edge density (Mack & Oliva, 2004), feature congestion (Rosenholtz et al., 2005; Rosenholtz et al., 2007), and subband entropy (Rosenholtz et al., 2007) for the images used in our experiment. The clutter measures were all significantly correlated (see Table 1), suggesting that the particular choice of clutter metric is not important.

The code for implementation of the segmentation algorithm is made publicly available by its authors. A MATLAB implementation of the clutter measure using this algorithm as described above can be downloaded

from our lab Github page ([https://github.com/mmmlab/clutter\\_metric\\_code](https://github.com/mmmlab/clutter_metric_code)).

### Procedure

The design of the experiment was 5 (target object category: cellphones, glasses, iPods, keys, or pens/pencils)  $\times$  2 (relevant set size: 5 or 13)  $\times$  2 (clutter level: low or high)  $\times$  2 (target presence: target present or target absent), with one between-subjects variable (target object category) and three within-subjects variables (relevant set size, clutter level, and target presence). At the start of the search experiment, observers were randomly assigned to one of five search target categories (cellphones, glasses, iPods, keys, or pens/pencils). Observers were instructed to detect and locate the target object within an image as quickly and accurately as possible. In addition, they were told that if the search target was present in an image, there was only one single item or a group of items in close proximity from the search category, and the item was visible.

Before the start of each trial, observers fixated a point at the center of the display while a set of circular cues indicated the potential target locations (see Figure 3). Observers began the trial by pressing a start key. After the trial was initiated, the search display appeared and observers freely searched for the target. Observers were allowed 3 s to search. After either 3 s

	Segmentation	Edge density	Feature congestion	Subband entropy
Segmentation	—			
Edge density	0.67	—		
Feature congestion	0.64	0.65	—	
Subband entropy	0.57	0.70	0.60	—

Table 1. Pearson correlation coefficients among clutter measures. Notes: All  $p < 0.001$ .

had elapsed or the observer pressed a key to complete the search early, the search image disappeared while the set of circular cues indicating the potential target locations remained on the screen. An additional cue appeared at a random location  $1^\circ$  outside the search region. Observers were instructed to make a localization decision. They fixated either the cue corresponding to the perceived location of the target (if target was present) or the additional cue (if target was absent). The cue corresponding to the current fixation was highlighted in real time to ensure that observers knew which locations they were selecting. Observers could correct their gaze if the wrong location was highlighted. When they were satisfied with their selections, observers logged their responses with a keypress.

The amount of time spent inspecting each image was recorded as the search time and was the primary measure of performance. In target-present trials, a response was registered as “correct” only if the selected cue corresponded to the target location. In the target-absent trials, a response was registered as “correct” only if the absent cue location was selected. All other responses were registered as errors. Observers received auditory feedback indicating the accuracy of their responses.

Trials were blocked by the relevant set size. Each block consisted of 50 experimental trials. At the start of each block, observers completed a 13-point calibration routine covering the central  $22^\circ$  of gaze angle. The calibration was repeated until the average test-retest calibration error across gaze points fell below  $0.25^\circ$ . The calibration routine could be repeated if necessary during a block. If a blink was detected during the search phase of the trial (when the image was present on the screen), the trial was aborted, and the observer was notified. Data from aborted trials were discarded, but the image from the discarded trial was repeated later in the experiment. Observers very rarely broke fixation, so fewer than 1% of trials were aborted.

Observers completed the study in two 1-hr sessions on separate days. Each session contained eight blocks, resulting in a total of 800 trials. The block order was randomized across sessions and observers.

Observers were trained and refamiliarized with the task by completing eight practice trials at the start of the experiment and a single practice trial at the start of each block. Data from the practice trials were excluded from the analysis.

represents data from five observers searching for one type of target (shapes) in either high-clutter (red lines) or low-clutter (blue lines) condition as a function of relevant set size. Each heavy line represents the average search times across target categories. Two main trends are evident: (a) search times tend to increase as the relevant set size increases and (b) search times tend to increase as the amount of clutter increases.

Search times were analyzed by conducting a 5 (target object category: cellphones, glasses, iPods, keys, or pens/pencils)  $\times$  2 (relevant set size: 5 or 13)  $\times$  2 (clutter level: low or high)  $\times$  2 (target presence: target present or target absent) mixed design analysis of variance (ANOVA), with one between-subjects variable (target object category) and three within-subjects variables (relevant set size, clutter level, and target presence).

The ANOVA revealed main effects of clutter level,  $F(1, 20) = 260.37, p < 0.001$ ; of relevant set size,  $F(1, 20) = 48.69, p < 0.001$ ; and of target presence,  $F(1, 20) = 173.69, p < 0.001$ . In particular, average search times were longer in the high-clutter condition ( $M = 1.37, SE = 0.01$ ) than in the low-clutter condition ( $M = 1.18, SE = 0.01$ ), suggesting that clutter degrades search performance. Search times were also longer in the set size 13 condition ( $M = 1.37, SE = 0.01$ ) than in the set size 5 condition ( $M = 1.19, SE = 0.01$ ), confirming our manipulation of set size. Finally, target-absent trials resulted in longer search times ( $M = 1.55, SE = 0.01$ ) than the target-present trials ( $M = 1.00, SE = 0.01$ ). The main effect of target category did not reach significance,  $F < 1, n.s.$

Our analysis also revealed several significant interaction effects. There was a significant clutter level  $\times$  relevant set size interaction,  $F(1, 20) = 4.93, p = 0.036$ , which suggests that the effect of clutter tends to be larger for larger set size. The clutter level  $\times$  target category interaction was also significant,  $F(4, 20) = 5.29, p = 0.005$ , suggesting that the effect of clutter was larger for some target categories than others. In addition, the clutter level  $\times$  target presence interaction reached significance,  $F(1, 20) = 48.65, p < 0.001$ , suggesting larger effects of clutter in the target-absent trials compared with the target-present trials. Moreover, set size  $\times$  target presence interaction was significant,  $F(1, 20) = 11.01, p = 0.003$ , which suggests larger set size effects in the target-absent trials compared with the target-present trials.

## Results

### Search times

Figure 5 shows average search times in the target-present and target-absent trials. Each faint line

### Fixation distributions

As a further check on our manipulation of set size, we examined observers' fixation distributions during search. If observers make use of the target location information provided by the cues when planning their

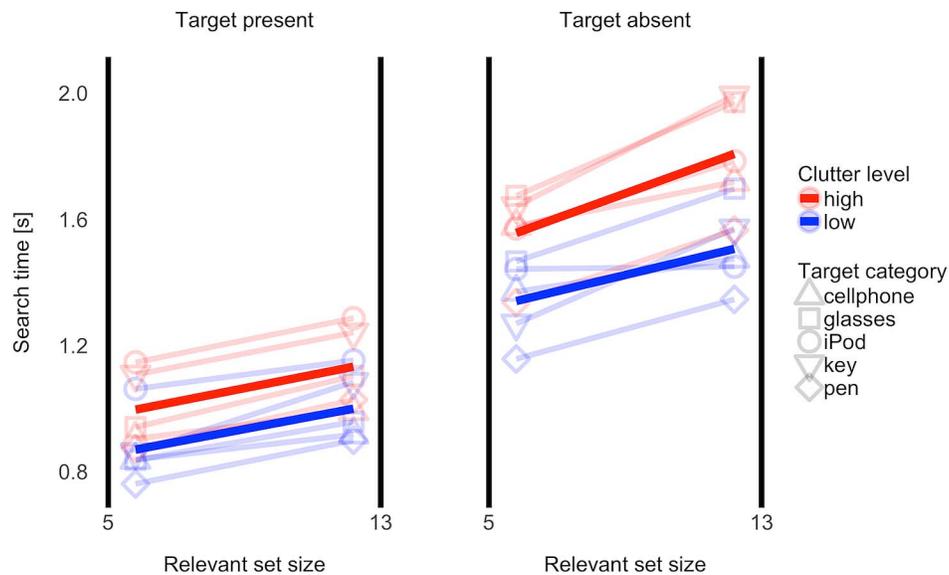


Figure 5. Average search times as a function of relevant set size in the target-present trials (left) and in the target-absent trials (right). Each combination of line and symbols represents data from five observers searching for one type of target (shapes) in either the high-clutter (red lines) or low-clutter (blue lines) condition. Average search times across target categories are represented by the heavy lines.

fixations, they should be more likely to fixate the cued locations than other locations in the display. Figure 6 shows the fixation distributions for each of the set size conditions, aggregated across all observers and trials, with the first and last fixations excluded. We excluded the first fixation from the analysis because observers always started the search by fixating at the center of the screen, and we excluded the last fixation from the analysis to avoid biasing the fixation distributions toward the target locations (i.e., because observers typically completed the search by fixating the target location). Our analysis of fixation distributions shows that observers indeed use cue location information when selecting their fixation locations, confirming the effectiveness of our set size manipulation.

### Search target sizes

We also examined how search time changed as a function of target size. Although we restricted the size of the targets to a limited range, there was still some degree of variability. Each target's size was quantified using either the area of its circumscribing polygon or the length of the longest axis of this polygon. To remedy the curvilinear relationship observed between the target area and the search times, the areas were transformed by taking their square root, which resulted in a more linear relationship. Figure 7 shows that (a) search times tend to decrease as the search target gets larger in size and (b) some targets are larger, on average, than others. The analysis showed that search times decrease significantly as target size increases, both when the size was measured as the area ( $r = -0.29$ ,  $p <$

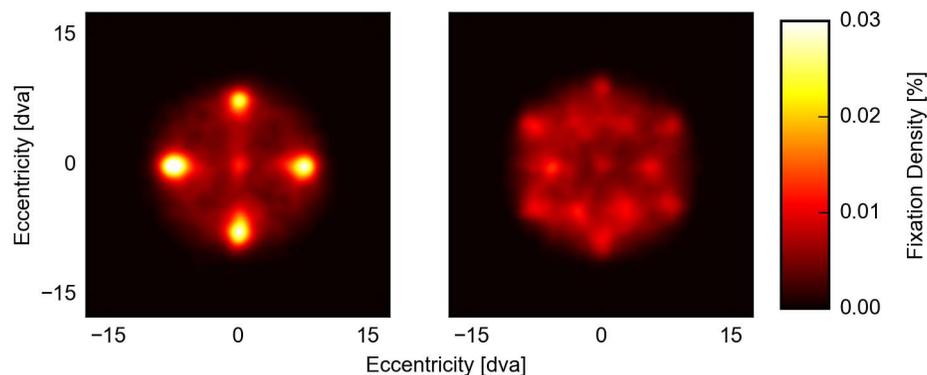


Figure 6. Aggregated fixation distributions across all of the observers, for set size 5 (left) and for set size 13 (right). The first and final fixations were excluded from the analysis.

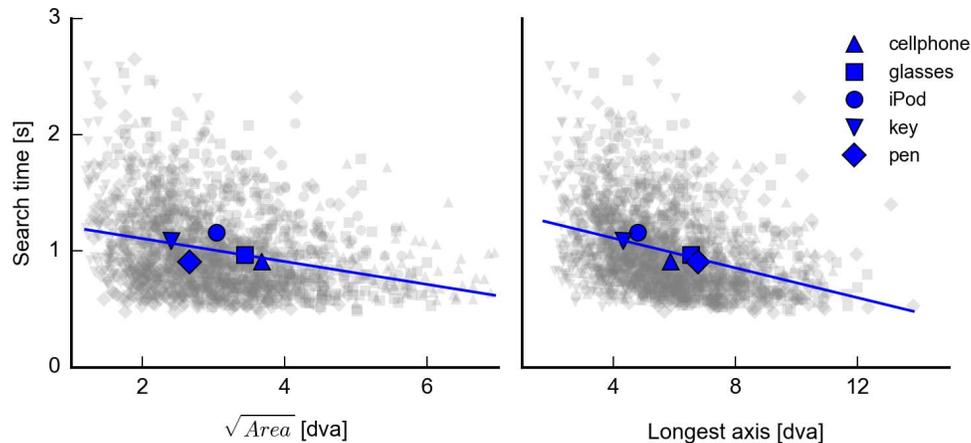


Figure 7. Search times as a function of target size represented by the square root of the area (left panel) or the longest axis (right panel) of its bounding polygon. Each gray dot represents the average search time across five observers for a particular target in an image. Shaped markers represent the average size for each target category. Blue lines represent the least-squares linear fits.

0.001) and when it was measured as the length of longest axis ( $r = -0.35$ ,  $p < 0.001$ ). These results suggest that target size may be one of the factors driving differences in search performance among target categories.

## Search target categories

Our stimulus set contained some common images across different target categories. That is, in some cases, different observers searched for different targets in the same image. These cases gave us the ability to dissociate effects of the search image from those of the search target and to directly examine the effect of target category on search performance. Figure 8 shows the average search times while searching for different targets in the same image. For example, the first plot shows the search time while looking for a cellphone compared with the search time while looking for the other targets in the same image. If the search performance was determined only by the amount of clutter or the relevant set size, then all points would line up on the diagonal. However, these results show that some targets were more difficult to find than others. For example, on average, observers seem to be faster at locating cellphones than other targets. We discuss potential implications of these results in the Discussion section.

## Error rates

Error rates were defined as the proportion of trials in which observers did not fixate the correct cue location (see the Procedure section). The probability of choosing any location other than the correct location (i.e., the

baseline error rate) was computed as 0.83 and 0.93 for set sizes of 5 and 13, respectively. Table 2 shows error rates across conditions. In general, error rates were well below baseline rates, suggesting that observers were extremely accurate in their judgments.

## Discussion

The purpose of the current study was to determine how clutter affects the search for categorical targets in real-world scenes. In particular, we sought to disentangle the effects of extrinsic position uncertainty (i.e., search set size) from those due, through the modulating effect of clutter, to intrinsic position uncertainty (Semizer & Michel, 2017). Our results exhibited several trends:

First, search times increased significantly as a function of increasing clutter. This pattern was evident across target categories, but the effect was larger for some targets than others. Second, search times increased significantly as the number of possible target locations increased, revealing the classic set size effect. This finding provided evidence that our manipulation of extrinsic uncertainty, via the relevant set size, was successful. Third, when all other manipulated factors (i.e., clutter level and set size) were fixed, search times changed as a function of target category. Finally, search times decreased significantly as a function of target size, both when the size was measured as the area of the circumscribing polygon and when it was measured as the length of the longest axis of this polygon. We discuss potential implications of these findings below.

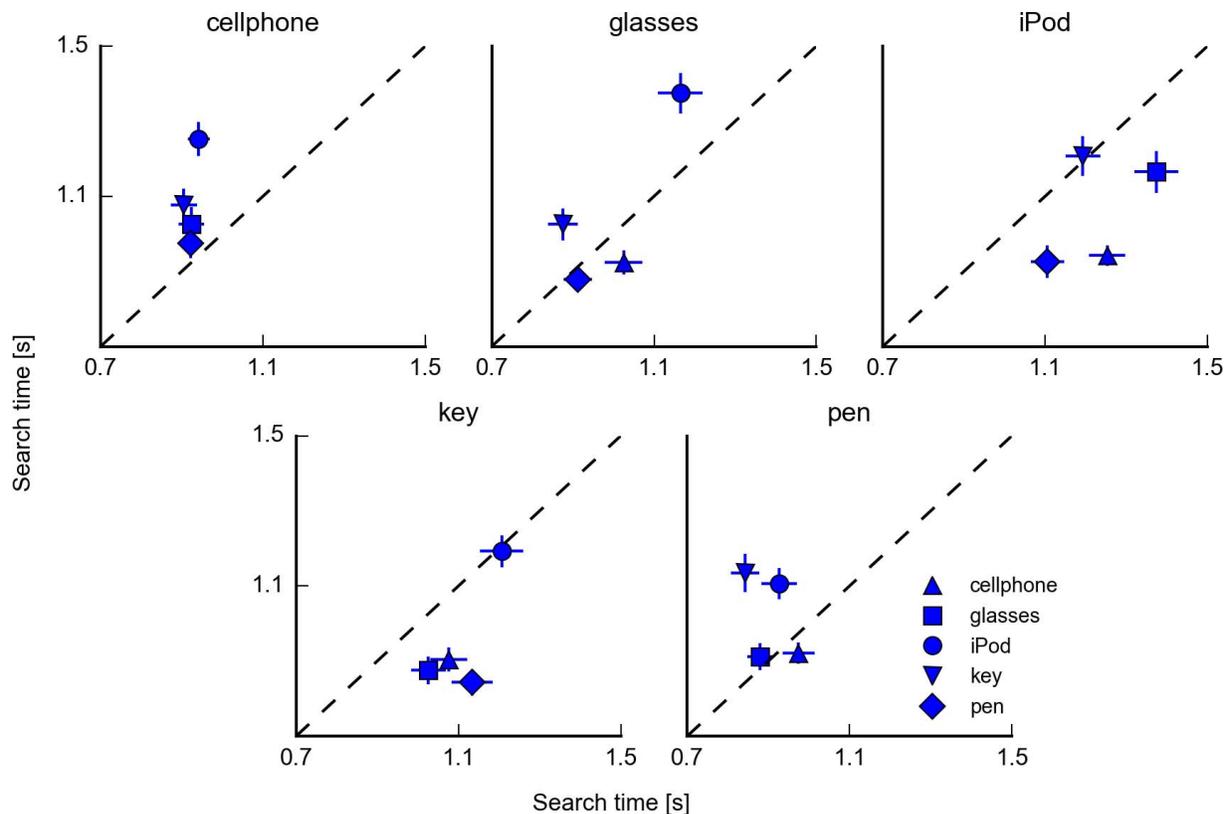


Figure 8. Average search times for different targets in common images. Each panel compares search time for a particular target category (on the x-axis) to search time for other targets (on the y-axis) in the same image. Each point represents average search times across all images that contained the indicated pair of targets. Error bars indicate standard error.

### Clutter-specific effects

Several studies have shown that clutter degrades search performance in naturalistic stimuli (e.g., Bravo & Farid, 2004; Bravo & Farid, 2008; Henderson et al., 2009; Neider & Zelinsky, 2011; Rosenholtz et al., 2007). However, there are various ways in which clutter can lead to the observed performance impairments. For example, clutter has been used as a proxy for set size in natural scenes because, as scene clutter increases, the (implicit) number of potential target locations also tends to increase (Rosenholtz et al., 2005; Rosenholtz et al., 2007). In addition, clutter can force observers to consider features at irrelevant locations during search, exacerbating the effects of intrinsic position uncertainty (Michel & Geisler, 2011; Semizer & Michel, 2017). As a

result, localizing the source of peripherally perceived stimuli becomes more difficult, which degrades search performance. Finally, clutter can make search more difficult by obscuring search targets. Adding clutter to real-world scenes increases the probability that objects will partially or completely occlude one another. In the current study, we controlled for set size and for occlusions of the search target to isolate those effects of clutter that are due to intrinsic position uncertainty.

The results of the current study, obtained using real-world images, are in broad agreement with those of a related study that showed how clutter degrades search performance in synthetic noise displays (Semizer & Michel, 2017). However, the results of the current study differ in one notable respect. Semizer and Michel (2017) reported that the effect of extrinsic position uncertainty

Clutter level	Relevant set size	Target category					Average
		Cellphone	Glasses	iPod	Key	Pen	
Low	5	0.04	0.03	0.09	0.03	0.06	0.05
	13	0.05	0.06	0.11	0.04	0.09	0.07
High	5	0.03	0.05	0.09	0.06	0.09	0.06
	13	0.05	0.07	0.12	0.09	0.11	0.09

Table 2. Error rates across conditions.

diminished at larger set sizes when the searcher was limited by intrinsic position uncertainty. As a result, search performance was similar across cluttered and uncluttered conditions when the relevant set size was large. However, our results showed that search performance was worse in the high-clutter condition than in the low-clutter condition regardless of the relevant set size. This difference might be due to either of two reasons: First, the images in our experiment were far less cluttered than the synthetic displays created in the lab. When measured using the same clutter metric, the synthetic stimuli from Semizer and Michel (2017) yielded clutter values of about  $\alpha = 4 \times 10^8$ , which was several orders of magnitude larger than the clutter values measured for our images (see Figure 2). Second, the relevant set sizes used in our study were much smaller than those in Semizer and Michel (2017). In the current study, the set sizes consisted of either five or 13, whereas the set sizes of Semizer and Michel (2017) ranged from a minimum of 37 to a maximum of 817 potential target locations. Indeed, our results are completely consistent with those of Semizer and Michel (2017) when we consider only the smaller set sizes used in that study.

### Characterizing set size

In many traditional search experiments, set size is defined as the number of items in a search display (see Wolfe, 1998, for a review). This type of set size is also called the “display set size.” A task-relevant subset of these items form the “relevant set size,” which can be manipulated independently of the display set size by cueing only the locations that might contain the target (Palmer, 1994; Palmer, 1995). Palmer (1994) introduced this distinction to characterize searches of displays comprising a small number of elements on a uniform background. However, the notion of relevant set size is especially critical in searches for which the number of elements is either very large or undefined. For example, researchers have used this method to define and manipulate set sizes for searches in synthetic noise displays (Burgess & Ghandeharian, 1984; Eckstein et al., 2007; Manjeshwar & Wilson, 2001; Najemnik & Geisler, 2005; Semizer & Michel, 2017; Swenson & Judy, 1981) and in structured medical images (Bochud et al., 2004; Eckstein & Whiting, 1996). In the current study, we likewise used Palmer’s (1994) method, cueing potential target locations to define the relevant set size in natural scenes.

In the context of natural scenes, there are several ways in which the effective set size might be made functionally smaller than the (nominal) relevant set size. First, if observers consider only the locations that contain “stuff” (i.e., objects, items, or feature elements)

and preferentially fixate only cues that fall on these locations, this would reduce the relevant set size. However, in the context of natural scenes, characterizing what “stuff” entails is problematic because labeling or counting every single item in a scene is an ill-defined problem. In particular, identifying what constitutes an object or a background is not clear (Neider & Zelinsky, 2008; Neider & Zelinsky, 2011; Rosenholtz et al., 2007; Wolfe, Vö, Evans, & Greene, 2011), especially when texture elements are involved. For example, in a kitchen scene, if objects are placed on a table covered with a patterned cloth or if objects are placed on other objects, it is not clear how to segment the scene into object or background. Similarly, if a scene contains a textbook (with text or illustrations on its cover), or a patch work quilt, or an articulated figure, it is ambiguous at what level objects should be segmented (should individual letters be considered objects? individual patches? individual parts?).

In the current study, if observers preferentially fixated cues that fall on objects, we would expect this to reduce the effective set size similarly for both clutter conditions. Because both set sizes would be reduced, this should not systematically influence the set size and clutter effects observed in our experiment. To test this relationship empirically (and based on a reviewer’s suggestion), we characterized the “effective” set size by counting the number of object cues (cues that fall on objects) in each trial of our experiment. The proportion of object cues were similar across low- and high-clutter conditions, 0.72 and 0.79, respectively, suggesting no substantial differences between the clutter conditions. Therefore, if having cues land on the “background” reduced the effective set size, then it did so similarly for both clutter conditions. Figure 9 shows search times as a function of the number of object cues, suggesting that the proportion of object cues is independent of clutter in this set of stimuli.

Set size might also be effectively reduced when targets are not distributed uniformly across cued locations. If the appearance of the target at a subset of the cued locations was much more probable than at other locations, then observers might restrict their searches to those locations within the probable subset, effectively reducing the set size. To investigate this possibility, we measured the likelihood that each cue location contained a target in our stimulus set. Target locations were well distributed across possible locations (Figure 10), with the exception of the center location, which was somewhat underrepresented.

To quantify the effect of this nonuniformity of the target location distributions in reducing the relevant set size, we computed the information entropy (Shannon, 1948) for each set size, given by

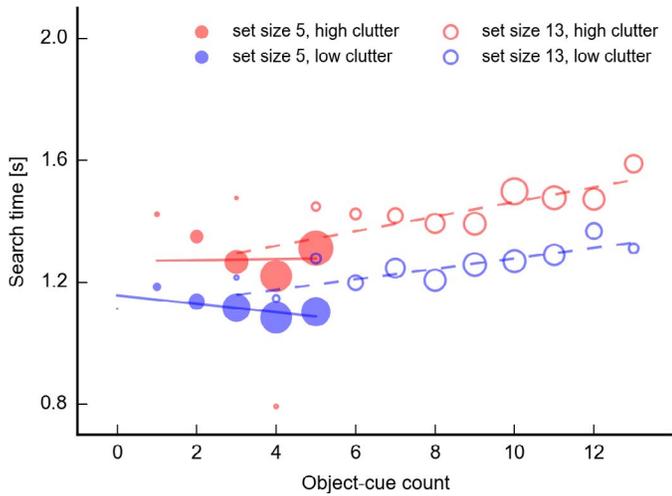


Figure 9. Search times as a function of the object-cue count (i.e., the number of cues that fall on objects) in the search image. The area of each marker is proportional to the number of trials exhibiting the corresponding object-cue count. Solid and dashed lines represent the least-squares linear fits for set sizes 5 and 13, respectively.

$$H = - \sum_i^n p_i \log_2 p_i, \quad (3)$$

where  $p$  is the probability that a cue location contains a target and  $n$  is the relevant set size. If the search targets were uniformly distributed across the cued locations, then the entropies associated with set sizes 5 and 13 would be 2.32 and 3.70 bits, respectively. The corresponding entropies computed for the empirical distributions measured in our experiment (Figure 10) were 2.21 and 3.59 bits, respectively. This reduction in entropy was small for both set sizes (approximately 5% for set size 5 and 3% for set size 13), suggesting that any reduction in effective set size caused by the nonuniformity of the target distributions should have a negligible effect on search performance.

### Target-specific effects

We used different types of target categories in the search experiment, and this allowed us to examine how a particular target category contributes to the effect of clutter on search performance. Our analysis of search times suggests that target category interacts with clutter. But what makes a target category more or less susceptible to clutter? Intuitively, it seems obvious that certain features of the target (e.g., target size, color, shape, etc.) might interact with features of clutter to determine search performance. Imagine, for example, a peripheral search task that requires the localization of a target object among green distractors. If the target object is red, then green objects will not provide effective clutter because their features are not confusable with features of the target. However, if the target is green, then the distractors should provide effective clutter.

More generally, the similarity of the target to features of the background might affect the susceptibility to clutter. Search times are longer when targets are similar to distractors or when distractors are dissimilar to other distractors (Duncan & Humphrey, 1989). Also, the similarity of targets to the search background affects search performance (Neider & Zelinsky, 2006). For example, when the search target and the background share a common spatial frequency band, search becomes more difficult (Semizer & Michel, 2017). Thus, when investigating the effects of clutter on search performance, it makes sense to expect performance differences depending on the similarity of target features to background features. For example, imagine a scene of leaves. The traditional models of clutter would consider this scene “highly cluttered” and predict poor search performance in this scene regardless of the type of the search target. However, if the search target, such as a cellphone, does not share a lot of similar features with the background, then the search

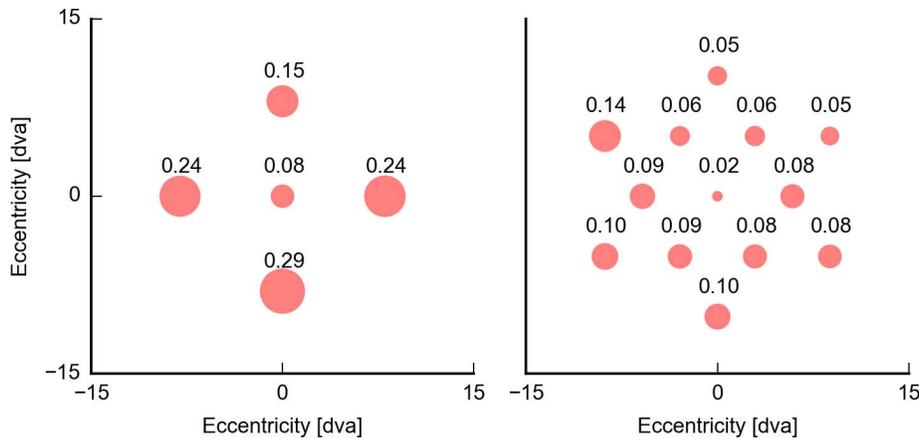


Figure 10. Distribution of target locations for set size 5 (left) and set size 13 (right) conditions. The area of each location marker is proportional to the empirical target probability for the corresponding location. The probabilities are also printed above each cue location marker.

should be pretty easy. In fact, there are models of clutter from the field of image optics, which quantify clutter in electro-optical images in terms of the target-background similarity (e.g., Chang & Zhang, 2006; Chu, Yang, & Qian, 2012; Moore, Camp, Moyer, & Halford, 2010; Schmieder & Weathersby, 1983; Silk, 1995; Tidhar, Reiter, Avital, & Hadar, 1994).

Another feature of the target that might affect its susceptibility to clutter is its size. Target size has been identified as one of the fundamental attributes in guiding attention (Wolfe & Horowitz, 2004). As one of the target-specific features, we measured size of the search targets in our stimuli and examined how search performance changes as a function of target size. The analysis suggests that larger targets are associated with shorter search times. This means that, within the context our study, larger targets are easier to find. Similarly, target size within a category might affect susceptibility to clutter, and the strength of the relationship between clutter and search performance might depend on details particular to different search targets. Revealing the nature of these specific target features remains an open question and one that we plan to pursue in future work.

## Conclusion

Overall, our results demonstrate that increased clutter reduces performance in searches of real-world scenes and does so independently of set size. When considered in the context of previous studies, Michel and Geisler (2011) and Semizer and Michel (2017), that explicitly modeled intrinsic position uncertainty, this study suggests that the intrinsic position uncertainty of peripheral vision significantly limits searches of real-world scenes in the same way it limits searches of synthetic scenes. Therefore, it is important to account for these effects of intrinsic position uncertainty when evaluating and modeling performance in search tasks.

*Keywords:* clutter, visual search, natural images, set size, intrinsic position uncertainty

## Acknowledgments

This work was supported by National Science Foundation Grant BCS-1456822.

Commercial relationships: none.  
Corresponding author: Yelda Semizer.  
Email: yelda.semizer@rutgers.edu.  
Address: Department of Psychology, Rutgers University, New Brunswick, NJ, USA.

## Footnotes

<sup>1</sup> A subset of images from this group was also used in a search task by Bravo and Farid (2008).

<sup>2</sup> The weights used in conversion of RGB values to gray-scale values were based on ITU-R Recommendation BT.601-7 standard for color video encoding.

## References

- Asher, M. F., Tolhurst, D. J., Troscianko, T., & Gilchrist, I. D. (2013). Regional effects of clutter on human target detection performance. *Journal of Vision*, *13*(5):25, 1–15, <https://doi.org/10.1167/13.5.25>. [PubMed] [Article]
- Bochud, F. O., Abbey, C. K., & Eckstein, M. P. (2004). Search for lesions in mammograms: Statistical characterization of observer responses. *Medical Physics*, *31*(1), 24–36, <https://doi.org/10.1118/1.1630493>.
- Bouma, H. (1970, April 11). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178, <https://doi.org/10.1038/226177a0>.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Bravo, M. J., & Farid, H. (2004). Search for a category target in clutter. *Perception*, *33*, 643–652, <https://doi.org/10.1068/p5244>.
- Bravo, M. J., & Farid, H. (2008). A scale invariant measure of clutter. *Journal of Vision*, *8*(1):23, 1–9, <https://doi.org/10.1167/8.1.23>. [PubMed] [Article]
- Burgess, A. E., & Ghandeharian, H. (1984). Visual signal detection: II. Signal-location identification. *Journal of the Optical Society of America A*, *1*, 906–910, <https://doi.org/10.1364/JOSAA.1.000906>.
- Caspi, A., Beutter, B. R., & Eckstein, M. P. (2004). The time course of visual information accrual guiding eye movement decisions. *Proceedings of the National Academy of Sciences, USA*, *101*, 13086–13090, <https://doi.org/10.1073/pnas.0305329101>.
- Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review*, *18*, 890–896, <https://doi.org/10.3758/s13423-011-0107-8>.
- Chang, H., & Zhang, J. (2006). New metrics for clutter affecting human target acquisition. *IEEE Transactions on Aerospace and Electronic Systems*, *42*, 361–368, <https://doi.org/10.1109/TAES.2006.1603429>.

- Chu, X.-q., Yang, C., & Qian, L. (2012). Contrast-sensitivity-function-based clutter metric. *Optical Engineering*, *51*, 067003, <https://doi.org/10.1117/1.OE.51.6.067003>.
- Cohn, T. E., & Wardlaw, J. C. (1985). Effect of large spatial uncertainty on foveal luminance increment detectability. *Journal of the Optical Society of America A*, *2*, 820–825, <https://doi.org/10.1364/JOSAA.2.000820>.
- Duncan, J., & Humphrey, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Eckstein, M. P., Beutter, B. R., Pham, B. T., Shimozaki, S. S., & Stone, L. S. (2007). Similar neural representations of the target for saccades and perception during search. *Journal of Neuroscience*, *27*, 1266–1270, <https://doi.org/10.1523/JNEUROSCI.3975-06.2007>.
- Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays. *Perception & Psychophysics*, *62*, 425–451, <https://doi.org/10.3758/BF03212096>.
- Eckstein, M. P., & Whiting, J. S. (1996). Visual signal detection in structured backgrounds. I. Effect of number of possible spatial locations and signal contrast. *Journal of the Optical Society of America A*, *13*, 1777–1787, <https://doi.org/10.1364/JOSAA.13.001777>.
- Egeth, H., Atkinson, J., Gilmore, G., & Marcus, N. (1973). Factors affecting processing mode in visual search. *Perception & Psychophysics*, *13*, 394–402, <https://doi.org/10.3758/BF03205792>.
- Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *International Journal of Computer Vision*, *59*, 167–181, <https://doi.org/10.1023/B:VISI.0000022288.19776.77>.
- Geisler, W. S., Perry, J. S., & Najemnik, J. (2006). Visual search: The role of peripheral information measured using gaze-contingent displays. *Journal of Vision*, *6*(9):1, 858–873, <https://doi.org/10.1167/6.9.1>. [PubMed] [Article]
- Henderson, J. M., Chanceaux, M., & Smith, T. J. (2009). The influence of clutter on real-world scene search: Evidence from search efficiency and eye movements. *Journal of Vision*, *9*(1):32, 1–8, <https://doi.org/10.1167/9.1.32>. [PubMed] [Article]
- Kleene, N., & Michel, M. (2018). The capacity of trans-saccadic memory in visual search. *Psychological Review*, *125*, 391–408, <https://doi.org/10.1037/rev0000099>.
- Krumhansl, C. L., & Thomas, E. A. C. (1977). Effect of level of confusability on reporting letters from briefly presented visual displays. *Perception & Psychophysics*, *21*, 269–279, <https://doi.org/10.3758/BF03214239>.
- Levi, D. M. (2008). Crowding—An essential bottleneck for object recognition: A mini-review. *Vision Research*, *48*, 635–654, <https://doi.org/10.1016/j.visres.2007.12.009>.
- Levi, D. M., Hariharan, S., & Klein, S. A. (2002). Suppressive and facilitatory spatial interactions in peripheral vision: Peripheral crowding is neither size invariant nor simple contrast masking. *Journal of Vision*, *2*(2):3, 167–177, <https://doi.org/10.1167/2.2.3>. [PubMed] [Article]
- Lohrenz, M. C., Trafton, J. G., Beck, M. R., & Gendron, M. L. (2009). A model of clutter for complex, multivariate geospatial displays. *Human Factors*, *51*, 90–101, <https://doi.org/10.1177/0018720809333518>.
- Mack, M. L., & Oliva, A. (2004). *Computational estimation of visual complexity*. Poster presented at the 12th Annual Object, Perception, Attention, and Memory Conference, Minneapolis, MN.
- Manjeshwar, R. M., & Wilson, D. L. (2001). Hyper-efficient detection of targets in noisy images. *Journal of the Optical Society of America A*, *18*, 507–513.
- Michel, M., & Geisler, W. (2009). Gaze contingent displays: Analysis of saccadic plasticity in visual search. *Society for Information Display Technical Digest*, *40*, 911–914, <https://doi.org/10.1889/1.3256945>.
- Michel, M., & Geisler, W. (2011). Intrinsic position uncertainty explains detection and localization performance in peripheral vision. *Journal of Vision*, *11*(1):18, 1–18, <https://doi.org/10.1167/11.1.18>. [PubMed] [Article]
- Moore, R. K., Camp, H. A., Moyer, S., & Halford, C. E. (2010). Masked target transform volume clutter metric applied to vehicle search. *Proceedings of SPIE*, *7662*, 76620M-1-11, <https://doi.org/10.1117/12.850429>.
- Najemnik, J., & Geisler, W. S. (2005, March 17). Optimal eye movement strategies in visual search. *Nature*, *434*, 387–391, <https://doi.org/10.1038/nature03390>.
- Najemnik, J., & Geisler, W. S. (2008). Eye movement statistics in humans are consistent with an optimal search strategy. *Journal of Vision*, *8*(3):4, 1–14, <https://doi.org/10.1167/8.3.4>. [PubMed] [Article]

- Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research*, *46*, 614–621, <https://doi.org/10.1016/j.visres.2005.08.025>.
- Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition*, *16*, 1–10, <https://doi.org/10.1080/13506280701381691>.
- Neider, M. B., & Zelinsky, G. J. (2011). Cutting through the clutter: Searching for targets in evolving complex scenes. *Journal of Vision*, *11*(14):7, 1–16, <https://doi.org/10.1167/11.14.7>. [PubMed] [Article]
- Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research*, *155*, 23–36, [https://doi.org/10.1016/S0079-6123\(06\)55002-2](https://doi.org/10.1016/S0079-6123(06)55002-2).
- Palmer, J. (1994). Set-size effects in visual search: The effect of attention is independent of the stimulus for simple tasks. *Vision Research*, *34*, 1703–1721, [https://doi.org/10.1016/0042-6989\(94\)90128-7](https://doi.org/10.1016/0042-6989(94)90128-7).
- Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, *4*, 118–123, <https://doi.org/10.1111/1467-8721.ep10772534>.
- Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*, 1227–1268, [https://doi.org/10.1016/S0042-6989\(99\)00244-8](https://doi.org/10.1016/S0042-6989(99)00244-8).
- Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, *2*, 1508–1532, <https://doi.org/10.1364/JOSAA.2.001508>.
- Pelli, D. G., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12):12, 1136–1169, <https://doi.org/10.1167/4.12.12>. [PubMed] [Article]
- Pelli, D. G., Tillman, K. A., Freeman, J., Su, M., Berger, T. D., & Majaj, N. J. (2007). Crowding and eccentricity determine reading rate. *Journal of Vision*, *7*(2):20, 1–36, <https://doi.org/10.1167/7.2.20>. [PubMed] [Article]
- Popple, A. V., & Levi, D. M. (2005). The perception of spatial order at a glance. *Vision Research*, *45*, 1085–1090, <https://doi.org/10.1016/j.visres.2004.11.008>.
- Rosenholtz, R., Huang, J., Raj, A., Balas, B., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, *12*(4):14, 1–17, <https://doi.org/10.1167/12.4.14>. [PubMed] [Article]
- Rosenholtz, R., Li, Y., Mansfield, J., & Jin, Z. (2005, April). *Feature congestion, a measure of display clutter*. Paper presented at CHI 2005, Portland, OR.
- Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of Vision*, *7*(2):17, 17.1–22, <https://doi.org/10.1167/7.2.17>. [PubMed] [Article]
- Schmieder, D. E., & Weathersby, M. R. (1983). Detection performance in clutter with variable resolution. *IEEE Transactions on Aerospace and Electronic Systems*, *AES-19*, 622–630, <https://doi.org/10.1109/TAES.1983.309351>.
- Semizer, Y., & Michel, M. (2017). Intrinsic position uncertainty impairs overt search performance. *Journal of Vision*, *17*(9):13, 1–17, <https://doi.org/10.1167/17.9.13>. [PubMed] [Article]
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 623–656, <https://doi.org/10.1145/584091.584093>.
- Silk, J. (1995). *Statistical variance analysis of clutter scenes and applications to a target acquisition test*. IDA Paper P-2950. Alexandria, VA: Institute for Defense Analysis.
- Swensson, R. G., & Judy, P. F. (1981). Detection of noisy visual targets: Models for the effects of spatial uncertainty and signal-to-noise ratio. *Perception & Psychophysics*, *29*, 521–534.
- Tanner, W. P. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences*, *89*, 752–765, <https://doi.org/10.1111/j.1749-6632.1961.tb20176.x>.
- Tidhar, G., Reiter, G., Avital, Z., & Hadar, Y. (1994). Modeling human search and target acquisition performance: IV. Detection probability in the cluttered environment. *Optical Engineering*, *33*, 801–808.
- Torralba, A., Oliva, A., Castelano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review*, *113*, 766–786, <https://doi.org/10.1037/0033-295X.113.4.766>.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136, [https://doi.org/10.1016/0010-0285\(80\)90005-5](https://doi.org/10.1016/0010-0285(80)90005-5).
- Verghese, P. (2012). Active search for multiple targets is inefficient. *Vision Research*, *74*, 61–71, <https://doi.org/10.1016/j.visres.2012.08.008>.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.),

- Attention* (pp. 13–71). London: University College London Press.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*, 495–501, <https://doi.org/10.1038/nrn1411>.
- Wolfe, J. M., Võ, M. L. H., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences*, *15*, 77–84, <https://doi.org/10.1016/j.tics.2010.12.001>.
- Wolford, G. (1975). Perturbation model for letter identification. *Psychological Review*, *82*, 184–199, <https://doi.org/10.1037/0033-295X.82.3.184>.
- Yu, C.-P., Samaras, D., & Zelinsky, G. J. (2014). Modeling visual clutter perception using proto-object segmentation. *Journal of Vision*, *14*(7):4, 1–16, <https://doi.org/10.1167/14.7.4>. [PubMed] [Article]
- Zhang, S., & Eckstein, M. P. (2010). Evolution and optimality of similar neural mechanisms for perception and action during search. *PLoS Computational Biology*, *6*, <https://doi.org/10.1371/journal.pcbi.1000930>.