Intrinsic position uncertainty impairs overt search performance

Yelda Semizer

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

Melchi M. Michel

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

Uncertainty regarding the position of the search target is a fundamental component of visual search. However, due to perceptual limitations of the human visual system, this uncertainty can arise from intrinsic, as well as extrinsic, sources. The current study sought to characterize the role of intrinsic position uncertainty (IPU) in overt visual search and to determine whether it significantly limits human search performance. After completing a preliminary detection experiment to characterize sensitivity as a function of visual field position, observers completed a search task that required localizing a Gabor target within a field of synthetic luminance noise. The search experiment included two clutter conditions designed to modulate the effect of IPU across search displays of varying set size. In the Cluttered condition, the display was tiled uniformly with feature clutter to maximize the effects of IPU. In the Uncluttered condition, the clutter at irrelevant locations was removed to attenuate the effects of IPU. Finally, we derived an IPU-constrained ideal searcher model, limited by the IPU measured in human observers. Ideal searchers were simulated based on the detection sensitivity and fixation sequences measured for individual human observers. The IPU-constrained ideal searcher predicted performance trends similar to those exhibited by the human observers. In the Uncluttered condition, performance decreased steeply as a function of increasing set size. However, in the Cluttered condition, the effect of IPU dominated and performance was approximately constant as a function of set size. Our findings suggest that IPU substantially limits overt search performance, especially in crowded displays.

Introduction

We are constantly engaged in tasks that require detecting, identifying, or localizing objects in our environment. These tasks typically involve some uncertainty about the position of the target object. For example, while typing this document and scanning a previous draft for reference, my gaze moves back and forth, to and from the previous draft. It always takes a moment to locate matching words or phrases in the draft and relocate the cursor in the current document. Visual tasks like these, that require the detection or localization of an object whose position is uncertain, are called visual search tasks.

As the example of searching through text documents illustrates, position uncertainty can have various origins. For example, an observer scanning a new page of text for a target word will necessarily have some uncertainty regarding the position of the word on the page. This type of uncertainty, which results from poor a priori specification of the target location, is called extrinsic position uncertainty (EPU). On the other hand, for an observer that is already familiar with the page of text (e.g., having previously located the target word on it), but whose gaze is currently directed elsewhere, the inaccuracies of visual memory and peripheral vision will tend to limit the observer's ability to localize the target word. This type of uncertainty, inherent in the observer, is called intrinsic position uncertainty (IPU).

Position uncertainty can have dramatic effects on detection and discrimination performance, with increases in position uncertainty leading to lower accuracy (Burgess & Ghandeharian, 1984; Eckstein, Thomas, Palmer, & Shimozaki, 2000), higher apparent detection and discrimination thresholds (Cohn & Lasley, 1974; Cohn & Wardlaw, 1985; Palmer, Verghese, & Pavel, 2000), and longer search times (Egeth, Atkinsons, Gilmore, & Marcus, 1973; Estes & Wessel, 1966; Treisman & Gelade, 1980). In search tasks, this (extrinsic) position uncertainty is typically manipulated by varying the number possible target locations (e.g., by varying the number of cued locations or by varying the number of distracter and target elements presented

in the display). Research examining the effects of position uncertainty in visual search and detection tasks has tended to focus on the effects of extrinsic, rather than intrinsic, sources of position uncertainty (e.g., Bochud, Abbey, & Eckstein, 2004; Burgess & Ghandeharian, 1984; Swensson & Judy, 1981; but see also Manjeshwar & Wilson, 2001; Pelli, 1985; Tanner, 1961). However, evidence from position discrimination (e.g., Klein & Levi, 1987; White, Levi, & Aitsebaomo, 1992) and crowding (e.g., Levi, 2008; Pelli et al., 2007) suggests that the precision with which we can localize features decreases—or, equivalently, that our uncertainty about the location of features increases—in the visual periphery. If intrinsic position uncertainty increases with retinal eccentricity, then we should expect to observe eccentricity-dependent effects of position uncertainty in visual tasks.

In a recent study, Michel and Geisler (2011) measured the intrinsic position uncertainty of human observers in a single-fixation visual search task. They found that intrinsic position uncertainty increases approximately linearly as a function of retinal eccentricity. Furthermore, they showed that an ideal observer constrained to have the same peripheral sensitivity and intrinsic position uncertainty as a human observer predicted the human detection and localization performance across changes in external position uncertainty. In contrast, an alternative ideal observer that was not constrained by intrinsic position uncertainty failed to predict human performance, systematically overestimating the impact of extrinsic position uncertainty. These results demonstrate that, in the peripheral visual field at least, intrinsic position uncertainty plays an important role in limiting visual search.

The current study extends this approach to overt, or multiple-fixation, searches. For a single-fixation search task in which the target may appear in some location other than the fovea, any imprecision of representations in the peripheral visual field (including IPU) are obviously important. Is this imprecision similarly important in more natural conditions of overt visual search, when observers are allowed to make eye movements?

It is not obvious that factors that depress peripheral search performance in a single fixation will retain their impact across multiple fixations. For example, observers performing overt searches might reduce position uncertainty at scene locations of interest by sequentially directing foveating eye movements to get more precise position information from those locations. One might hypothesize that this information, when integrated across eye movements, should minimize any effects of peripheral IPU. Similar reasoning applied other forms of peripheral degradation, such as color or luminance contrast sensitivity, would suggest that the effect of

these deficits should likewise be minimized when observers are permitted to make multiple fixations. However, existing research examining effects of peripheral deficits in overt search seem to argue against this hypothesis. For example, even subtle and virtually undetectable changes in peripheral contrast sensitivity have been shown to significantly impact search performance (Geisler, Perry, & Najemnik, 2006), and a recent metastudy of overt search tasks showed that differences in search-time slopes across various classic search stimuli can be explained by a model in which visual feature representations in the peripheral visual field are represented by statistical summaries (Rosenholtz, Huang, Raj, Balas, & Ilie, 2012).

Both of these findings suggest that peripheral deficits that limit performance in single-fixation tasks also degrade performance in multiple-fixation searches. The purpose of the current study is to examine this claim with respect to IPU in particular, and to determine how the effect of IPU on search performance is modulated by EPU and by the density of feature clutter in the search environment. In particular, we examined the effect of IPU on overt search performance using a two-pronged approach:

First, we derived normative models to describe the performance of ideal observers whose peripheral sensitivity and fixation strategies were matched to those of a human observer, but whose use of visual information was otherwise optimal. To assess directly how IPU *should* impact search performance, we derived two versions of the ideal observer model—an *IPU-constrained* searcher and an *unconstrained* searcher—and simulated their performance under varying levels of EPU. Both the IPU-constrained and the unconstrained searchers were limited by degrading their peripheral sensitivity to match that of human observers, but the IPU-constrained searcher was additionally limited with peripheral IPU matching that of human observers. Recall that EPU is a property of the search task (i.e., EPU is manipulating by varying the number of potential target locations in the task) so the performance of both model searchers should be influenced by EPU. However, the effect of EPU on performance should vary across observers.

Figure 1 schematically represents the pattern of performance that we expect for the two model searchers, assuming an equally limited search time across conditions. For an "unconstrained" observer search performance should worsen (errors should increase) as EPU increases. This performance drop results primarily from the well-known stimulus uncertainty effect (Cohn & Lasley, 1974; Pelli, 1985; Swensson & Judy, 1981; Tanner, 1961). As the number of possible target locations increases, the observer must monitor an increasing number of "noise" locations, any one of which may generate a false alarm. The result is
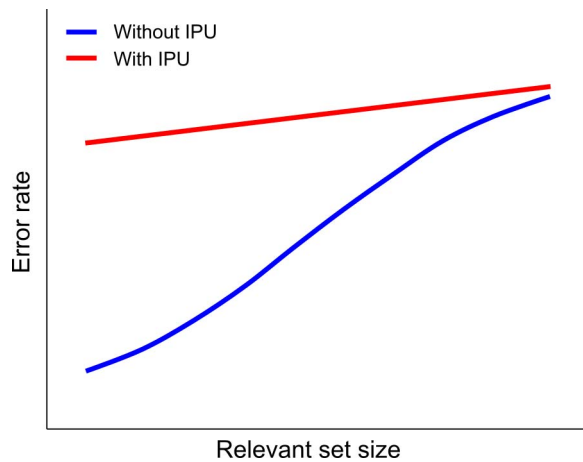
Figure 1. A schematic representation of expected search performance as a function of intrinsic position uncertainty (IPU) and extrinsic position uncertainty (EPU). For observers without IPU (blue curve), error rates should rise dramatically as EPU, indexed by the relevant set size, increases. For observers with significant IPU (red curve), average error rates should be larger than without IPU, but performance differences across changes in set size should diminish, reflecting a reduced sensitivity to variations in EPU.

that the probability of falsely identifying a noise patch as a target increases with EPU. For an IPU-constrained searcher with substantial IPU, overall performance should be worse, but the effect of EPU should be attenuated. This is because the IPU-constrained searcher has some position uncertainty "built in" in the form of IPU, and a proportion of the location uncertainty added by EPU will be subsumed by the effect of this intrinsic uncertainty (see Michel & Geisler, 2011).

Second, we measured human performance in a search task with different background noise conditions designed to probe the effect of intrinsic position uncertainty. For human observers, we cannot remove IPU (which is an intrinsic property of the visual system), but what we can do is to construct environments so as to either enhance or minimize the effects of IPU. In the current study we achieved this by varying the distribution of relevant feature clutter in the background of the search display. We constructed two different background noise conditions: a *Cluttered* condition that exacerbates the effects of IPU on search performance by distributing feature clutter uniformly across the display, and an *Uncluttered* condition that attenuates its effects by removing feature clutter from all scene locations but those cued as potential target locations (see Figure 2). Functionally, these two clutter conditions represent a means of manipulating the "display set size" independently of the "relevant set size" which is defined in terms of the number of cued potential target locations (Palmer, 1994, 1995). The difference in performance between these two clutter

Cluttered
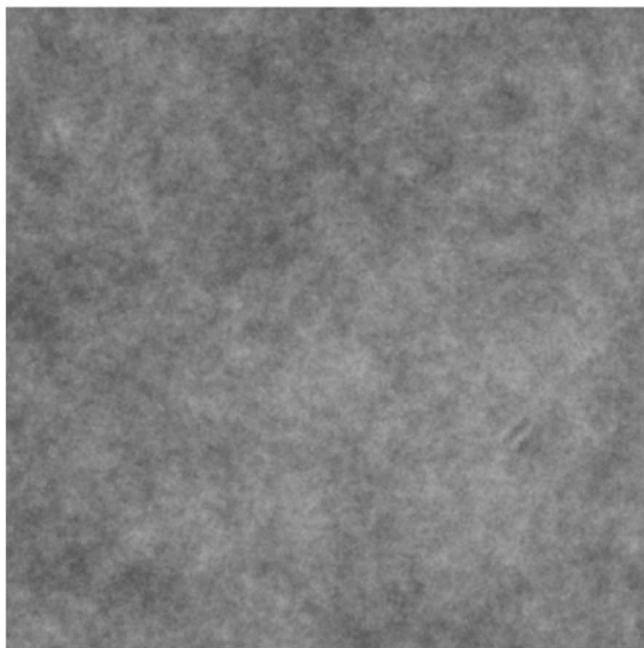
Uncluttered
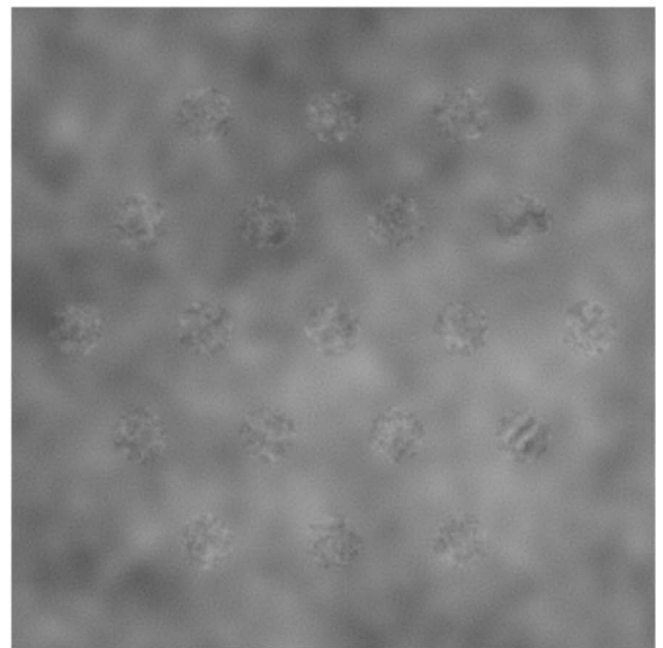

Figure 2. Cluttered and Uncluttered displays. Each display consists of a field of noise with a Gabor target located at one of 19 cued target locations. In the Cluttered condition (left panel), the display is tiled uniformly with relevant feature clutter (in the form of $1/f$ noise). In the Uncluttered condition (right panel), the relevant feature clutter at nontarget locations is removed.

conditions as we increase the relevant set size (i.e., increasing EPU) reveals how IPU limits search performance in cluttered displays.

Using this approach, we show that the overt search performance of human observers is significantly limited by IPU. Specifically, we show that the effects of clutter and set size on human search performance are predicted by an ideal searcher model that includes measured human peripheral intrinsic position uncertainty as a constraint.

Methods

Observers

A total of five human observers participated in the study. Four observers participated in the main search experiment and three (including two of the observers from the main search experiment) participated in the "simultaneous-cue search" control experiment. One of the observers was an author; the other four were naïve to the purpose of the experiment. All observers had normal or corrected-to-normal vision and received compensation for their participation.

Apparatus

Stimuli were presented on a 22-in Philips 202P4 CRT monitor with a resolution of 1280×1024 pixels at 100 Hz. The viewing distance was set to 70 cm from the observer so that the display subtended $15.8° \times 21.1°$ of visual angle. The stimuli were generated and presented using MATLAB software (Mathworks) and the Psychophysics Toolbox extensions (Brainard, 1997). Observers' eye movements were monitored and recorded using an Eyelink 1000 infrared eye tracker (SR Research, Kanata, Ontario, Canada). Head position was maintained using a forehead and chin rest, and eye position signals were sampled from the eye tracker at 1000 Hz.

Stimuli

The target was a 4 cycle/° sine-wave grating, oriented 45° clockwise from vertical and windowed by a raised cosine function with a diameter of 0.875° of visual angle (i.e., a raised-cosine Gabor function; see Figure 2). Contrasts and thresholds for the target are reported in terms of the Michelson contrast of the sinusoidal component. The background was a circular region 24° in diameter filled with 10% contrast (root-mean-square, RMS) luminance noise. Two different types of noise

were used to fill the background. In the Cluttered condition, the background was filled uniformly with $1/f$ noise at a mean luminance of 40 cd/m^2 (Figure 2, left panel). The $1/f$ noise was created by filtering Gaussian white noise, truncating the noise waveform at $\pm 2SD$, and scaling to obtain desired RMS amplitude. The $1/f$ noise contains significant energy within the spatial frequency band of the target, so that filling the stimulus background with this noise adds relevant feature clutter uniformly across the display. In the Uncluttered condition, we created "notched" background noise using a bandstop spatial frequency filter centered on the spatial frequency of the target (Figure 2, right panel). This type of background allows us to limit the spatial locations that might give rise to "distractors" that are similar in spatial frequency and orientation to the target, while maintaining much of the local contrast structure that might influence sensitivity to the target itself via contrast gain control mechanisms (Bex, Mareschal, & Dakin, 2007; Carandini, Heeger, & Movshon, 1997; Geisler & Albrecht, 1992; Sperling, 1989; Wilson, 1993). The bandstop notch was defined as a log-Gaussian function in the frequency domain with a bandwidth of two octaves. Importantly, we only removed the feature clutter at the irrelevant locations. To maintain equivalent local noise contrast masking across the Cluttered and Uncluttered conditions, we always presented $1/f$ noise at the relevant display locations (i.e., the potential target locations). Finally, the area around the circular region was set to a uniform gray with luminance equal to the mean display luminance.

Procedure

Procedures are described below for the Detection and Search tasks. Participants completed a 13-point calibration routine covering the central 22° of gaze angle. The calibration was repeated until the average test-retest calibration error across gaze points fell below 0.25°. If an eye movement (in the detection task) or a blink (in either task) was detected during a trial, the trial was aborted, the observer was notified, and data for that trial were discarded.

Observers completed the study over a total of 10 one-hr sessions, with each hour-long session occurring on a separate day. Observers started by completing three sessions of the detection task (pre-test), followed by five sessions of the search task, followed by 2 more sessions of the detection task (post-test). We measured detection performance in both pre- and post-test sessions in order to account for any changes in the observers' contrast sensitivity over the course of the experiment.
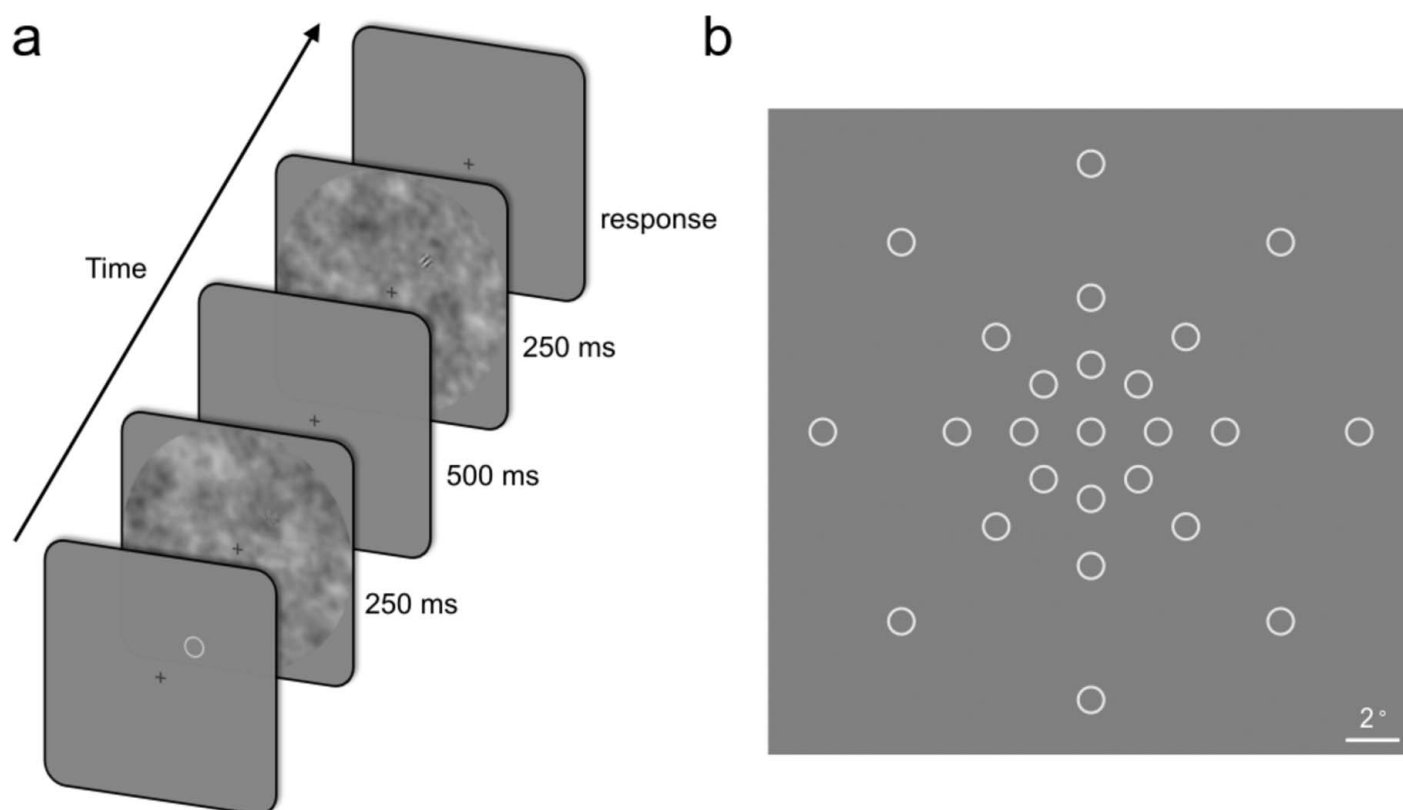
Figure 3. The Detection task. (a) Stimulus sequence for a trial of the detection task. Target size and contrast have both been increased for visibility. In this example, the target is present in the second interval. (b) Visual field locations that were tested to construct visibility maps.

Detection task

At the start of each trial, observers fixated a small cross at the center of the display while an open circle, located at one of eight possible target locations, indicated the location of the target in the current trial. After the observer pressed a start button, the cue disappeared and the stimulus sequence was presented. The stimulus sequence consisted of two stimulus displays each presented for 250 ms, separated by a blank display lasting 500 ms. One of the two stimulus displays contained the target embedded on a $1/f$ noise patch at the cued location while the other contained only the noise patch. The remainder of the background region was always filled with a frequency-notched noise pattern like that used in the Uncluttered search display (see task sequence in Figure 3a). We used frequency-notched noise to minimize any potential effects of IPU on our sensitivity measurements. The observer's task was to report which of the two stimulus displays, the first or the second, contained the target signal. Observers received auditory feedback after each trial indicating whether or not they had selected the correct interval. Gaze position was monitored throughout each trial and trials were discarded if the observer's gaze deviated by more than 1° from the fixation marker.

Observers started each block by completing five practice trials, for which the data were not recorded, followed by the experimental trials. An adaptive procedure (Kontsevich & Tyler, 1999) was used to determine the target contrast for each trial. Trials were blocked by retinal eccentricity for four different eccentricities ($\varepsilon = 0.0°, 2.5°, 5.0°,$ and $10.0°$). Experimental sessions were composed of seven 100-trials blocks, for a total of 700 trials per session. Individual observers completed the Detection Experiment in a total of five sessions, including three pre-test sessions (completed before the search task) and two post-test sessions. The first session was a practice session and its data were excluded from the analysis. The last four sessions (2,800 trials) were used to construct each observer's visibility map (see the Visibility maps section).

Search task

As in the detection task, observers started each trial by fixating a small cross at the center of the display and pressing a start key. Prior to the start of the trial, all potential target locations were marked by circular cues. Once the start key was pressed, the search display appeared and observers were free to make eye
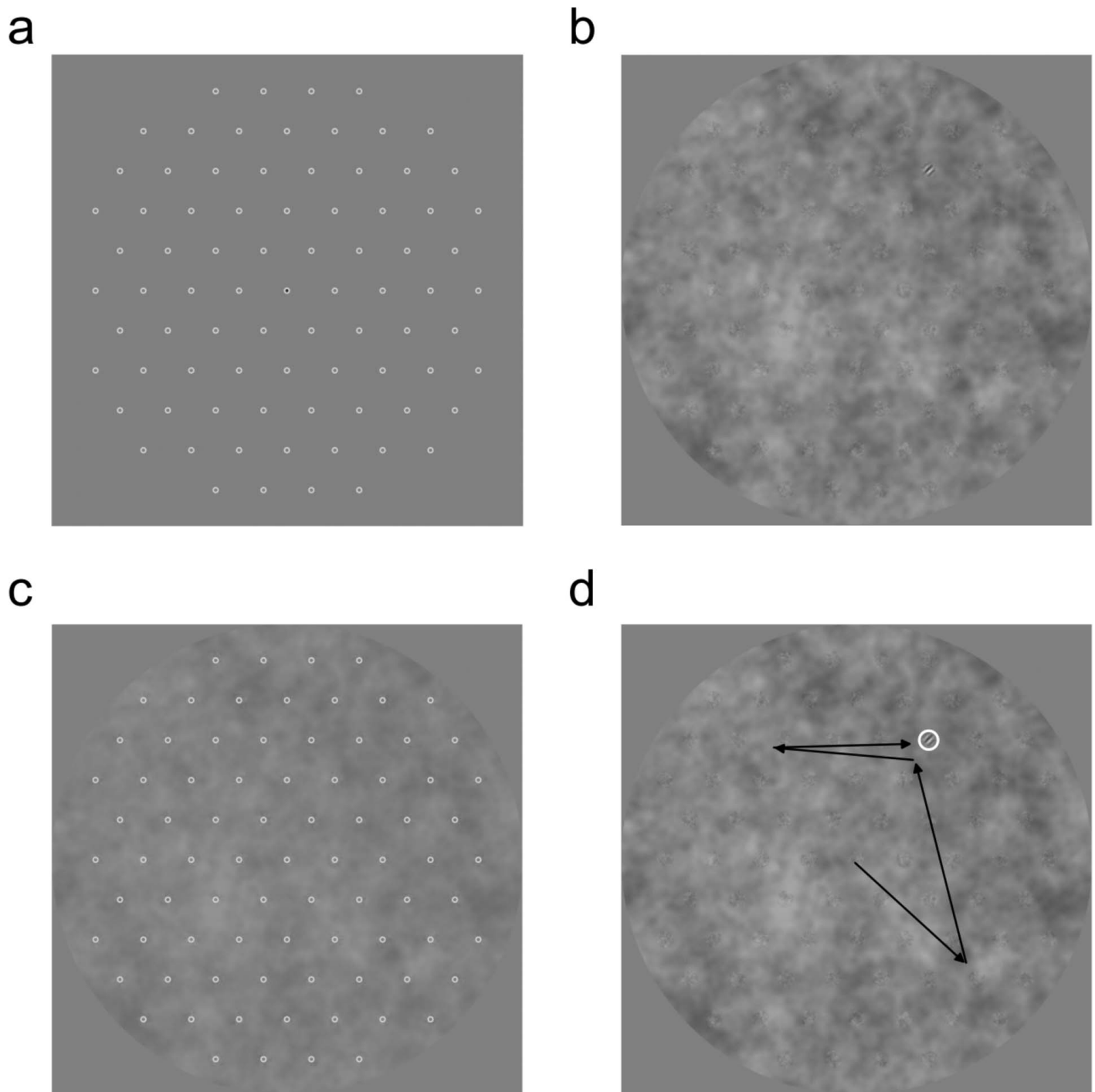
Figure 4. The search task. Subpanels show sample displays for a trial in the Uncluttered condition with a relevant set size of 85. (a) The initial fixation screen, showing cues for 85 potential target locations. (b) The search display. The Gabor target appears in the upper-right quadrant of the display. (c) The response display, showing all possible target locations. (d) The feedback display, showing the sequence of detected fixations (black arrows) and the actual target location (white circle).

movements and to search for the target signal. Observers were instructed to locate the target as quickly and accurately as possible, prioritizing accuracy. Importantly, observers were only allowed a maximum of six fixations or 3 s (whichever expired first) to locate the target.[1] After either six fixations or 3 s had elapsed, the search display disappeared and was replaced by a low-contrast version of the noise background, with all potential target location markers superimposed. Observers were instructed to fixate the marker corresponding to the perceived location of the target and to log their responses with a keypress. Trials
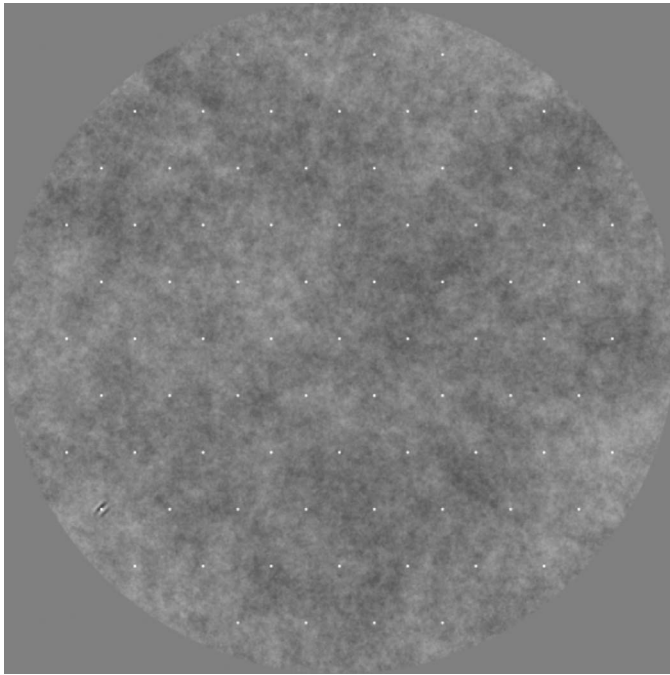
Figure 5. Simultaneous-cue search. A sample display from the Cluttered condition of the simultaneous-cue search with a relevant set size of 85. The target appears in the lower left quadrant of the display.

were considered "correct" if the indicated target location was within 1° of the actual target location. Observers received both auditory feedback (indicating the correctness of the response) and visual feedback (indicating the actual target location and displaying the observer's gaze scanpath).

As in the Detection Experiment observers started each block by completing five practice trials. Trials were blocked both by the clutter condition (i.e., Cluttered or Uncluttered) and by the relevant set size (37, 85, 163, 421, or 817 potential target locations). Each session contained 10 blocks of 50 trials (one for each unique combination of set size and clutter condition), and the block order was randomized across sessions and observers. Each observer required five sessions to complete the search experiment. The first session was a practice session and its data were excluded from the analysis. The last four sessions (2000 trials) were included in the analysis.

Control (simultaneous-cue) search task

We designed the Uncluttered and Cluttered conditions of the search task to attenuate and enhance, respectively, any effect of compulsory spatial integration (due to IPU) on human search performance. However, performance between the two conditions could also differ due to a spatial cueing effect in the Uncluttered condition that is not available in the

Cluttered condition. In particular, the difference between the notched noise and $1/f$ noise backgrounds in the Uncluttered condition could be used to segregate potential target locations from the background, thereby serving as a visual memory aid or an attentional cue to the possible target locations. To control for this possibility, we designed a control version of the search task (the *simultaneous-cue* search task) that was identical to the main search task, except that location cues were superimposed on the search display in both the Uncluttered and Cluttered conditions. The cues were small circular markers centered on the each of the possible target locations (see Figure 5). The size (0.14°) and luminance (80 cd/m^2) of the cue markers were selected on the basis of pilot experiments such that: (a) the cue was equally detectable in the periphery as an isolated patch of $1/f$ noise (in a field of the frequency-notched noise used for the background of the Uncluttered search condition), and (b) the cue did not significantly impact the detectability of the search target in the periphery. Satisfaction of both criteria were confirmed using standard 2IFC detection experiments conducted at 5° and 10° of visual eccentricity.

Two experienced psychophysical observers who had already completed the main search task participated in the simultaneous-cue search task, along one naïve human observer. Aside from the addition of the location cue markers, the stimuli and procedure for the simultaneous-cue search task were identical to those used in the main search task.

Searcher models

To determine how intrinsic position uncertainty *should* influence search performance, we derived and simulated the performance of two different ideal observer models, an *IPU-constrained searcher* that is limited by IPU and an *unconstrained searcher* that is not limited by IPU. Both models were foveated (Geisler & Chou, 1995; Geisler et al., 2006; Legge, Klitz, & Tjan, 1997); that is, they were constructed to model variation in the fidelity of visual representation as a function of position in the visual field. In deriving these models, we considered two classes of intrinsic uncertainty that limit human performance in detection and localization tasks: intrinsic *response* uncertainty, which represents internal factors that contribute to uncertainty in the magnitude of a perceived signal; and intrinsic *position* uncertainty, which represents internal factors that contribute to uncertainty in the spatial source of a perceived signal. Both types of uncertainty vary as a function of visual field position and both were modeled in terms of equivalent internal noise (Ahumada & Watson, 1985; Lu & Dosher, 1999). Both ideal

observer models were limited by intrinsic response uncertainty, which was characterized in terms of the contrast sensitivity for individual observers measured across the visual field (see the Visibility maps section). In addition, the IPU-constrained searcher was also limited by intrinsic position uncertainty as measured in a previous study (Michel & Geisler, 2011). By comparing the performance of the two searcher models, we can determine how the introduction of intrinsic position uncertainty to an otherwise ideal observer limits visual search performance.

Importantly, neither of our searcher models included an explicit fixation selection strategy, instead relying on human fixations. While the choice of fixation strategy can dramatically influence search performance, there is an ongoing debate regarding the optimality of human fixation strategies, with evidence for near-optimal fixation selection in some search tasks (e.g., Michel & Geisler, 2009; Najemnik & Geisler, 2005, 2008) and evidence for markedly suboptimal fixation selection in others (e.g., Acker-mann & Landy, 2013; Clarke, Green, Chantler, & Hunt, 2016; Morvan & Maloney, 2012; Nowakowska, Clarke, & Hunt, 2017; Paulun, Schütz, Michel, Geisler, & Gegenfurtner, 2015; Verghese, 2012). Because the purpose of the current study is to determine the role that IPU plays in search, we do not attempt to engage this debate. Instead, we elide the issue of fixation selection altogether by forcing our model searchers to use the same fixations selected by human observers. This ensures that any performance differences between the human and model searchers cannot be attributed to differences in fixation strategies.

We simulated each trial of the human search task for the *IPU-constrained* and *unconstrained* model searchers. As in the human search task, the display was a field of spatial noise containing an embedded target signal at one of n_C cued target locations, and model observer's task in each trial was to determine which of these cued locations contained the target signal.

Below, we outline our formalization of the search task for the model observers. First, we describe the abstract representation of the display "viewed" by the model observers in a trial, then we derive the optimal decision rules for computing the most probable target location given this representation, and, finally, we describe the details of the simulation procedure itself.

Representing the display

We model the display as a set of n_D discrete image patches, each representing one of the nonoverlapping, 0.875°-diameter spatial regions of the display that contain $1/f$ noise. A subset of these patches, corre-sponding to the n_C regions cued in advance of each trial, represent potential target locations (i.e., in the nomenclature introduced by Palmer, 1994, 1995, n_C represents the relevant set size). For convenience, we only represent the $1/f$ noise-containing regions. This reflects our assumption that only patches containing $1/f$ noise (and/or the target signal itself) have sufficient contrast energy within the frequency band of the target to elicit visual responses that might be confused with the target. As a result, the representation of the display differs slightly between clutter conditions. In the Uncluttered condition, only the cued regions contain $1/f$ noise, so $n_D = n_C$, while in the Cluttered condition, the n_D image patches tile the display and $n_D \geq n_C$.

During each fixation, the observer receives noisy matched-template responses from each of the n_D image patches. Let R_i represent the response obtained from display location $\mu_i = (x_i, y_i)$, where i indexes the display locations $1, \ldots, n_D$, $R_i = r_i + N_r(i)$, and $N_r(i)$ is a sample of Gaussian noise with mean 0 and a standard deviation $\sigma_r(i)$. In addition, let J represent the target location selected randomly from among the n_C possible target locations. For mathematical convenience and without loss of generality (i.e., as in Michel & Geisler, 2011; Najemnik & Geisler, 2005) we assume that the mean response r_i of patch i is 0.5 at the target location (where $i = J$) and -0.5 elsewhere. The standard deviation $\sigma_r(i)$ of each response is a function both of extrinsic factors such as the contrast energy and spectral content of the signal and noise in the corresponding patch, and of various forms of intrinsic sensory uncertainty that we characterize in terms of an additive *equivalent internal noise* (Lu & Dosher, 1999). Importantly, in addition to extrinsic stimulus proper-ties, the intrinsic response uncertainty also varies as a function of the position of the eliciting stimulus in the visual field. We characterize the overall response uncertainty in terms of measured sensitivity d' across the visual field. Because the physical characteristics of our stimulus patches were statistically identical across search conditions, the standard deviation of the equivalent response noise for patch i can be expressed simply as a function of the position of the patch in the visual field

$$\sigma_r(i) = \frac{1}{d'(\varepsilon_i, \theta_i)}, \quad (1)$$

where ε_i represents the retinal eccentricity of the patch, θ_i represents its angular direction, and $d'(\varepsilon_i, \theta_i)$ is computed using the psychophysical measurements and sensitivity model described in the Visibility maps section.

Let $L_i = (X_i, Y_i)$, represent the encoded location of response R_i. Due to the effects of intrinsic position uncertainty, this encoded location is variable. We represent it as a two-dimensional Gaussian variable whose distribution is centered on the true patch

location $\boldsymbol{\mu}_i$ and whose covariance varies as a function of retinal eccentricity ε_i

$$L_i = (X_i, Y_i) \sim \mathcal{N}\left[\mu_i, \sigma_p^2(i)\mathbf{I}\right], \quad (2)$$

where σ_p represents the standard deviation of the IPU,

$$\sigma_p(i) = m_p \varepsilon_i, \quad (3)$$

and $m_p = 0.09$ is a linear coefficient measured in an earlier study (Michel & Geisler, 2011) that characterized human IPU as a linear function of retinal eccentricity. When modeling the unconstrained searcher, we set the standard deviation of the IPU to 0 across all eccentricities.

Computing the target location

The observer's task is to determine the location J of the target as accurately as possible based on the magnitudes $\mathbf{R}_{\phi(t)}$ and locations $\mathbf{L}_{\phi(t)}$ of responses encoded across a sequence of T visual fixations $\phi(1), \ldots, \phi(T)$.

We start by computing the likelihood of the observations. That is, we want to determine the probability of obtaining a set of perceived responses $\mathbf{R}_{\phi(1)}, \ldots, \mathbf{R}_{\phi(T)}$ and perceived response locations $\mathbf{L}_{\phi(1)}, \ldots, \mathbf{L}_{\phi(T)}$ for a display that contains a target at location J and noise elsewhere. First, consider the likelihood for a single fixation. Assuming that the intrinsic response noise and position noise are both independent across display locations,

$$p(\mathbf{R}, \mathbf{L}|J = j) = \sum_{k=1}^{n} p(R_1, \ldots, R_n|j, k)p(L_k|J = j),$$
$$= \sum_{k=1}^{n} p(L_k|j) \prod_{i=1}^{n_D} p(R_i|j, k),$$

$$(4)$$

where

$$p(R_i|j, k) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma_r(i)} \exp\left[-\frac{(R_i + 0.5)^2}{2\sigma_r^2(i)}\right] & \text{if } i = k, \\ \frac{1}{\sqrt{2\pi}\sigma_r(j)} \exp\left[-\frac{(R_i - 0.5)^2}{2\sigma_r^2(j)}\right] & \text{if } i \neq k. \end{cases}$$

$$(5)$$

Here, the subscript j indexes over possible target locations, while the subscript k indexes over the encoded responses and their perceived locations, and $p(L_k|j)$ represents the probability that response R_k with encoded response location L_k was elicited by image patch j in the scene. The likelihood in Equation 4 is effectively a weighted average of the response likelihoods in which the likelihood of each perceived response is weighted by the probability that a target at location j gave rise to that response. This weighted

average illustrates why spatial "pooling" is a sensible strategy for observers with IPU. In the limiting case for which an observer has no IPU at all (e.g., as assumed in Najemnik & Geisler, 2005), the probability weights for all perceived locations but one (the location for which $L_k = \mu_j$) become zero, so that each target index j is associated with exactly one possible response index k. Thus, for an observer with no IPU, the sum in Equation 4 is obviated and the likelihood reduces to

$$p(\mathbf{R}|J = j) = \prod_{i=1}^{n_C} p(R_i|j).$$

In general, however, the probability $p(L_k|j)$ will be small for encoded locations L_k that are distant from μ_j, the position of the jth possible target location, so that actual targets will rarely give rise to sizable perceptual displacements. In detail, the distribution of encoded response locations depends on the retinal eccentricity as described in Equations 2 and 3.

For convenience, we represent the set of possible encoded locations in terms of a finite number of n_D discrete patches. This means that we treat any encoded location L_k falling within patch k as originating from discrete location $\mu_k, k \in [1, n_D]$. The probability $p(L_k|j)$ of associating a response elicited by patch j with location L_k is therefore computed as the integral of the Gaussian distribution described in Equation 2 over the circular patch centered on location μ_k. This distribution becomes broader as the eliciting patch moves farther into the visual periphery, increasing the probability that responses elicited from neighboring patches might be confused.

We can further simplify Equation 4 by observing that the rightmost term $p(R_i|j,k)$ is equal for all $k \neq j$,

$$p(\mathbf{R}, \mathbf{L}|J = j) = K \sum_{k=1}^{n_D} p(L_k|j) \frac{p(R_i|j, k = i)}{p(R_i|j, k \neq i)}, \quad (6)$$

where $K = \prod_{l=1}^{n_D} p(R_l|j, k \neq l)$. Substituting in the Gaussian likelihoods (Equation 5) and simplifying yields

$$p(\mathbf{R}, \mathbf{L}|J = j) = K \sum_{k=1}^{n_D} p(L_k|j) \frac{\sigma_r(k) \exp\left[-\frac{(R_k - 0.5)^2}{2\sigma_r^2(j)}\right]}{\sigma_r(j) \exp\left[-\frac{(R_k + 0.5)^2}{2\sigma_r^2(k)}\right]},$$
$$= K \sum_{k=1}^{n_D} p(L_k|j) \frac{\sigma_r(k)}{\sigma_r(j)}$$
$$\times \exp\left[\frac{(R_k + 0.5)^2}{2\sigma_r^2(k)} - \frac{(R_k - 0.5)^2}{2\sigma_r^2(j)}\right].$$

$$(7)$$

Finally, if we assume temporally independent response and position noise,[2] the likelihood for the

entire sequence of fixations $\phi(1:T) = \phi(1), \cdots, \phi(T)$ can be computed as a product of the likelihoods for the individual fixations. That is,

$$p(\mathbf{R}_{\phi(1:T)}, \mathbf{L}_{\phi(1:T)} | J = j)$$
$$= \prod_t^T p(\mathbf{R}_{\phi(t)}, \mathbf{L}_{\phi(t)} | J = j),$$
$$\propto \prod_t^T \sum_{k=1}^{n_D} p(X_{\phi(t)k} | j) \frac{\sigma_r(k)}{\sigma_r(j)}$$
$$\times \exp\left[\frac{(R_{\phi(t)k} + 0.5)^2}{2\sigma_r^2(k)} - \frac{(R_{\phi(t)k} - 0.5)^2}{2\sigma_r^2(j)} \right]. \quad (8)$$

Recall that the observer's goal is to estimate the location of the target J given the observations $(R_{\phi(1:T)}, L_{\phi(1:T)})$. We use Bayes' rule to obtain the posterior $p(J = j | \mathbf{R}_{\phi(1:T)}, \mathbf{L}_{\phi(1:T)})$ from the likelihood

$$p(J = j | \mathbf{R}_{\phi(1:T)}, \mathbf{L}_{\phi(1:T)})$$
$$\propto p(J = j) p(\mathbf{R}_{\phi(1:T)}, \mathbf{L}_{\phi(1:T)} | j). \quad (9)$$

Because the set of possible target locations is discrete, we choose the most probable location—the *maximum a posteriori* (MAP) estimate—to maximize the response accuracy. Furthermore, because the prior probability over $p(J = j)$ potential target locations in the search experiment is uniform, the MAP estimate of the target location j^* is equal to the maximum likelihood estimate. That is,

$$j^* = \underset{j}{\operatorname{argmax}} \, p(\mathbf{R}_{\phi(1:T)}, \mathbf{L}_{\phi(1:T)} | J = j). \quad (10)$$

Simulation procedure

We used Monte Carlo simulation to estimate the localization performance of the ideal searcher models. Unconstrained and IPU-constrained versions of the ideal searcher were simulated for each human observer using that observer's measured visibility map. Additionally, because we were interested primarily in the effects of intrinsic position uncertainty on the integration of visual information across fixations (rather than on the effects of fixation selection per se), each simulated ideal searcher also used the fixation sequences from the corresponding human observer. Each trial was simulated based on the corresponding human trial as follows:

1. The target location J was selected to match that of the corresponding human trial. The mean response magnitudes were set to $r_J = 0.5$ and $r_i = -0.5$ for $i \neq J$.
2. For each fixation, Gaussian noise samples R_i were generated at each of the n_D locations in the display as described in the Representing the display section.

3. For each fixation, intrinsic position noise at the target location was simulated by randomly selecting one of the discrete n_D locations in the grid L_J as the encoded target location according to $p(L_J | \boldsymbol{\mu}_j, \sigma_p(j))$. For the unconstrained searcher, this encoded target location was always veridical (i.e., $L_j = \boldsymbol{\mu}_j$). Because randomly changing the position of a noise patch has no effect on performance, the position noise was not explicitly simulated for noise patches. Instead, as in Michel and Geisler (2011), the perceived locations of the noise patches L_i, $i \neq J$ were set equal to their mean locations $L_i = (x_i, y_i)$.
4. The observer determined the MAP target location by integrating response and location information optimally across fixations using Equation 10.

To compute the final estimated performance curves for each searcher model (Figure 7), we simulated each human trial 10 times.

A Python language implementation of these ideal searcher simulations, along with some sample human data, can be found at https://github.com/mmmlab/ipu_searcher.

Results

Visibility maps

Visual sensitivity for each human observer was characterized in terms of a "visibility map" that describes the effective signal-to-noise ratio d' as a function of retinal eccentricity ε and angular direction θ. The visibility map was obtained by taking the inverse standard normal integral of a function describing accuracy in the combined pre- and post-test detection data. That is,

$$d'(c, \varepsilon, \theta) = \sqrt{2} \Phi^{-1}[PC(c, \varepsilon, \theta)], \quad (11)$$

where Φ^{-1} represents the standard normal integral and $PC(c, \varepsilon, \theta)$ is a psychometric function representing the expected proportion of correct answers in a 2IFC detection experiment. As in Najemnik and Geisler (2005), the factor $\sqrt{2}$ takes into account the fact that there were two intervals in the forced-choice detection task, but only a single interval per fixation of the search task (Green & Swets, 1966). We modeled the detection accuracy at a particular retinal location (ε, θ) as a cumulative Weibull function of target contrast c

$$PC(c; \varepsilon, \theta) = 1 - 0.5 \exp\left[-\left(\frac{c}{\alpha(\varepsilon, \theta)} \right)^\beta \right], \quad (12)$$

where β is a parameter controlling the steepness of the psychometric function and α is a contrast threshold parameter that varies with the retinal position of the
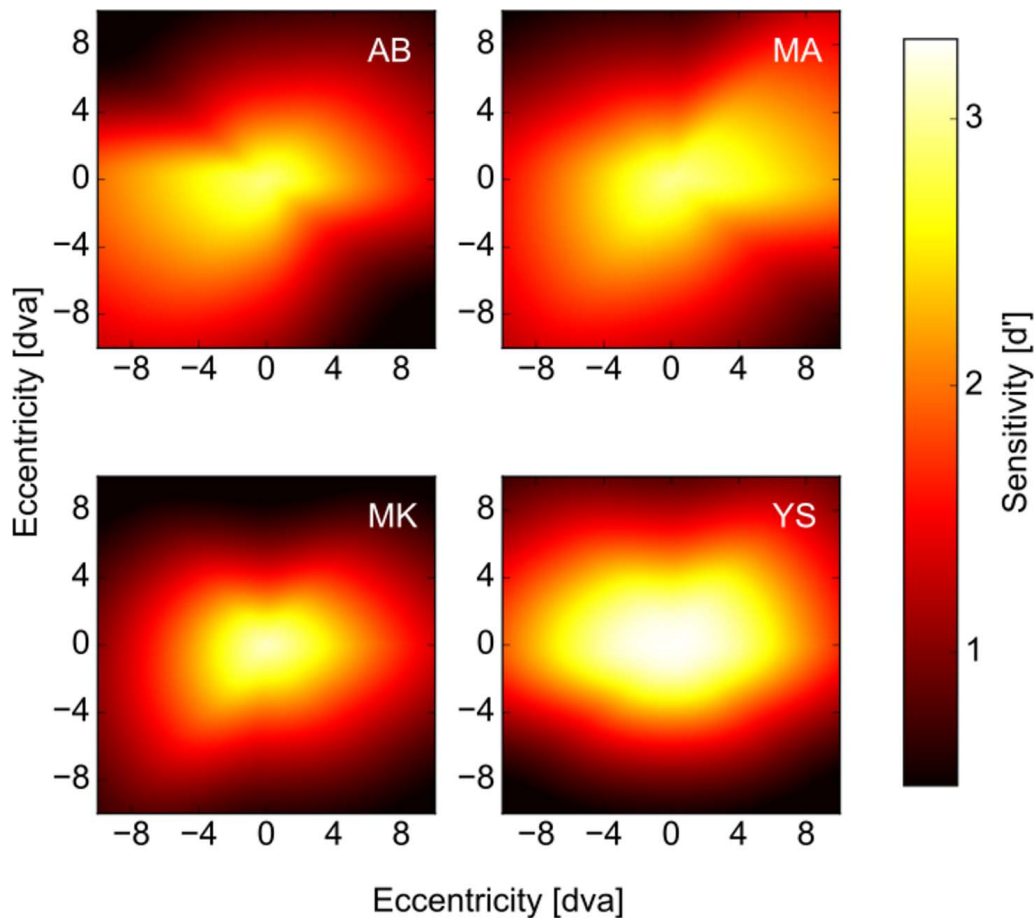
Figure 6. Visibility maps. Each panel shows the sensitivity (d') of an individual human observer to the 20% contrast search target, measured as a function of the target's position in the visual field.

target. Individual estimates of the steepness parameter computed for different target locations did not vary significantly from each other, so in fitting the visibility maps, we assumed a common steepness parameter β. This is consistent with recent measurements reported for low spatial frequency targets (Ackermann & Landy, 2013) and for a detection model that controls for the effects of intrinsic position uncertainty (Michel & Geisler, 2011).

We modeled contrast thresholds $\alpha(\varepsilon, \theta)$ using a log-linear model

$$\alpha(\varepsilon, \theta) = \alpha(0) \exp(\tau_\theta \varepsilon), \quad (13)$$

where $\alpha(0)$ represents the foveal threshold and τ_θ is a log slope parameter controlling the rise in contrast thresholds as a function of eccentricity. This function has been shown to accurately describe the rise in contrast thresholds with increasing eccentricity across a variety of visual tasks (Peli, Yang, & Goldstein, 1991). The resulting psychometric model has 10 parameters: $\alpha(0), \beta$, and a separate log slope parameter τ_{θ_i} for each of the eight directions along which we measured detection performance $\theta_i \in \{0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°\}$, where $i = \theta/45°$ represents the

direction index. We used maximum likelihood to fit this model to each observer's detection data. For intermediate values of θ not measured in the detection experiment (i.e., for nonintegral i), we computed τ_θ by linearly interpolating between the nearest measured values $\tau_{\theta_{\lfloor i \rfloor}}$ and $\tau_{\theta_{\lceil i \rceil}}$.

Figure 6 shows the visibility maps measured for each of our observers. These maps represent the sensitivity (d') to the 4 cycles/° search target across the visual field when presented at 20% contrast. The visibility maps show quite a bit of individual variability across observers, but they all demonstrate the features previously reported in similar (photopic) visibility measurements of normal human observers (Ackermann & Landy, 2013; Michel & Geisler, 2009; Najemnik & Geisler, 2005; Paulun et al., 2015): (a) sensitivity is highest at the fovea and decreases as a function of retinal eccentricity, (b) sensitivity decreases more slowly along the horizontal axis than along the vertical axis, and (c) sensitivity is greater in the lower visual field than in the upper visual field.
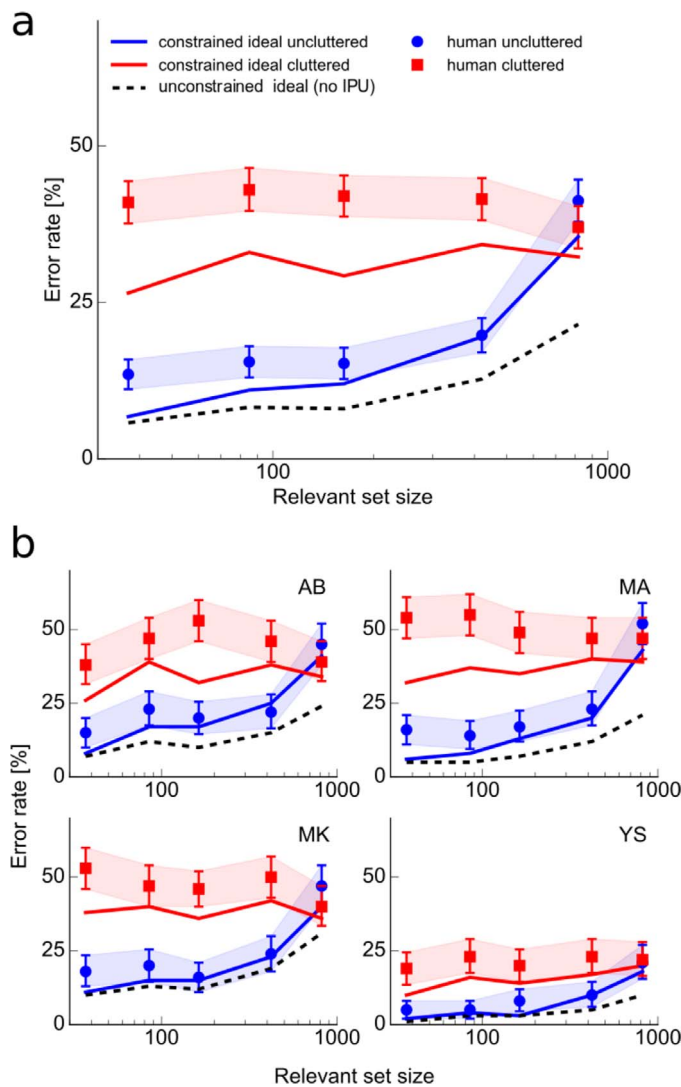
Figure 7. Search performance for human and simulated observers. (a) Aggregate performance averaged across human observers. (b) Performance computed for individual observers. Each panel plots error rate (%) as a function of the number of possible target locations in Cluttered (red) and Uncluttered (blue) backgrounds. Markers and error bars indicate mean performance and 95% confidence intervals for human performance. Solid curves represent expected performance for IPU-constrained ideal observers, whereas dashed curves represent expected performance for unconstrained ideal observers.

Search performance

Main search task

Performance in search tasks generally involves a tradeoff between search speed and localization accuracy. To simplify the characterization of performance in the current search task, we constrained the speed of the search by limiting the number of fixations (to a maximum of six) and the search time (to a maximum of 3 s). We then characterized human search performance in terms of the localization error rate, where localiza-

tion errors were defined as trials in which the target location reported by observers differed from the actual target location by more than 1°.

Figure 7a (markers) shows the aggregate localization error rates for human observers. A two-way, within-subjects ANOVA revealed main effects of noise condition, $F(1, 3) = 43.85$, $p = 0.007$, and of the relevant set size, $F(4, 12) = 8.87$, $p = 0.001$, as well as a significant noise condition $\times$ set size interaction, $F(4, 12) = 13.88$, $p < 0.001$. In particular, the average error rate was nearly twice as large in the Cluttered condition ($M = 40.2\%$, $SE = 1.6\%$) than in the Uncluttered condition ($M = 20.6\%$, $SE = 1.5\%$). The error rate also increased as a function of the relevant set size, but a simple effects analysis showed that this effect was only significant for the Uncluttered condition, $F(4, 12) = 21.18$, $p < 0.001$. In the Cluttered condition, human performance did not vary significantly as a function of set size, $F(4, 12) = 0.94$, $p = 0.474$. Importantly, the lack of a set size effect in the Cluttered condition cannot be explained as a ceiling effect on the localization error rate. Due to the large effective set size (37–817 potential target locations) and the large area of the search region ($\sim 450^{o2}$), chance error rates exceeded 97% in all conditions, which is more than twice the average error rate recorded in the Cluttered condition. Results for individual observers (Figure 7b) show the same overall pattern as the aggregate data, including (a) an approximate doubling of the average error rate between the Uncluttered and Cluttered conditions, (b) a relatively flat error rate across set sizes for the Cluttered condition, and (c) an increase in error rate as a function of set size in the Uncluttered condition that converges onto the error rate for the Cluttered condition.

For each human observer, we simulated the unconstrained and IPU-constrained ideal searchers as described in the Searcher models section. Both ideal searchers modeled sensitivity using the visibility maps obtained from the detection task (Figure 6) and used human fixations recorded in the search task.

The results of these simulations are shown in Figure 7. The dashed curves indicate the performance of the unconstrained observer, and the solid curves indicate the performance of the IPU-constrained observer. The individual curves (Figure 7b) represent the localization error rates computed for 20,000 simulated trials (100 simulated repetitions of each human trial) at each of the five relevant set sizes, while the aggregated curves (Figure 7a) represent the average of the individual localization error rates.

The unconstrained searcher (Figure 7, dashed curves) has low error rates ($M = 11.25\%$) that increase moderately as a function of the relevant set size. Importantly, the performance of the unconstrained searcher does not vary as a function of the background
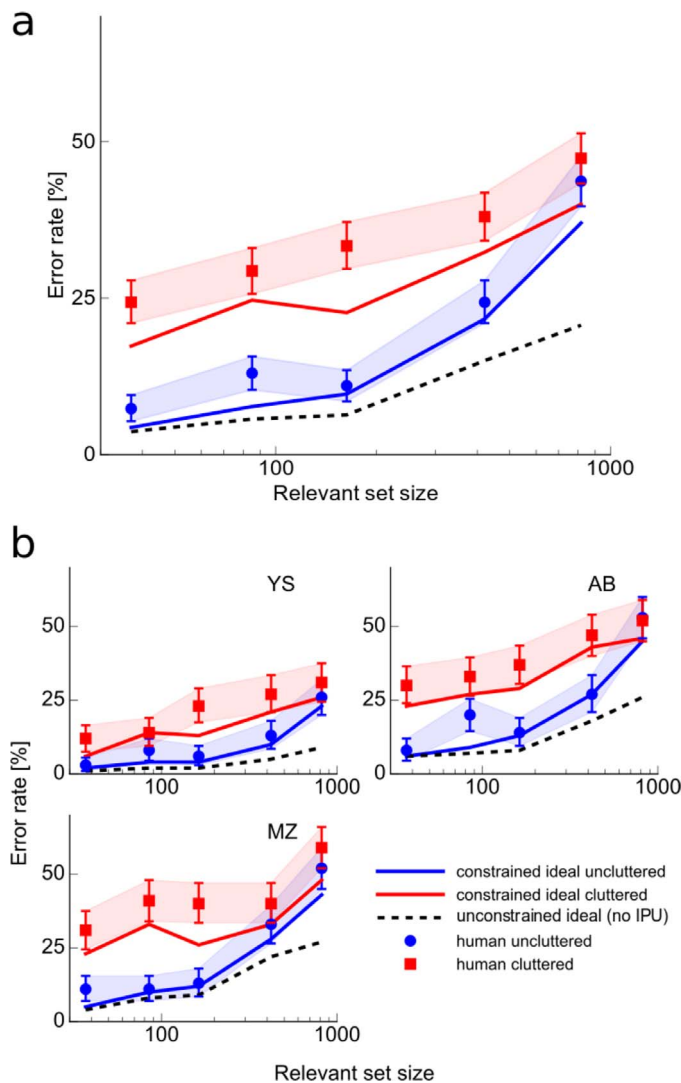
Figure 8. Performance in the simultaneous-cue search task. (a) Aggregate performance averaged across human observers. Each panel plots localization error rate (%) as a function of the number of possible target locations in Cluttered (red) and Uncluttered (blue) backgrounds. Markers and error bars indicate mean performance and 95% confidence intervals for human performance. Solid curves represent expected performance for IPU-constrained ideal observers, while dashed curves represent expected performance for unconstrained ideal observers. (b) Performance for individual observers.

type. For an ideal searcher without IPU, signals at nontarget locations are irrelevant and are ignored, so there is no distinction between the Cluttered and Uncluttered conditions.

In contrast, the performance of the IPU-constrained searcher varies dramatically as a function of the background type, exhibiting a pattern that mirrors that of the human observers: (a) the average error rate approximately doubles between the Uncluttered ($M = 16.95\%$) and Cluttered ($M = 31.05\%$) conditions; (b) the error rate for the Cluttered condition is elevated but

relatively insensitive to set size; and (b) the localization error rate in the Uncluttered condition rises dramatically as a function of the relevant set size (from $M = 6.75\%$ for a set size of 37 to $M = 35.50\%$ for a set size of 817).

Simultaneous-cue search task

Figure 8a (markers) shows the aggregate localization error rates for human observers in the simultaneous-cue experiment. As in the main search experiment, a two-way, within-subjects ANOVA revealed main effects of noise condition, $F(1, 2) = 41.84$, $p = 0.02$, and of the relevant set size, $F(4, 8) = 37.38$, $p < 0.001$. However, the noise condition × set size interaction did not reach statistical significance, $F(4, 8) = 3.02$, $p = 0.086$. The average error rate was larger in the Cluttered condition ($M = 33.40\%$, $SE = 1.7\%$) than in the Uncluttered condition ($M = 19.18.6\%$, $SE = 2.0\%$), though not by quite as large a margin as in the Main experiment, and the error rate increased as a function of the relevant set size.

Results for individual observers (Figure 7b) show the same overall pattern as the aggregate data, though the trends are less apparent for subject YS, who as an author and a participant in the main search experiment, had extensive previous experience with the search task and exhibited substantially lowered error rates.

As in the main experiment, we simulated the unconstrained and IPU-constrained ideal searchers for each human observer, using the visibility maps obtained from the detection task and the human fixations recorded in the search task. The simulation results were similar to those obtained for the main experiment, although overall performance was somewhat better (i.e., localization error rates were smaller).

Once again, the unconstrained searcher (Figure 8, dashed curves) exhibits low error rates ($M = 10.26\%$) that increase moderately as a function of the relevant set size, while the IPU-constrained searcher exhibits a pattern similar to that of the human observers: (a) The average error rate approximately doubles between the Uncluttered ($M = 16.07\%$) and Cluttered ($M = 27.40\%$) conditions; (b) the localization error rate in the Uncluttered condition rises dramatically as a function of the relevant set size; and (c) the error rate for the Cluttered condition is elevated but less sensitive to set size, so that the error rates for the Cluttered and Uncluttered conditions converge as the relevant set size increases.

Discussion

The purpose of the current study was to determine the effect of intrinsic position uncertainty (IPU) in

overt search. Specifically, we sought to determine whether IPU substantially limits performance in search tasks with high extrinsic uncertainty, or tasks for which the relevant set size is large. Our results suggest that IPU significantly limits overt search performance, especially in search displays that include a large amount of clutter. Evidence for this conclusion comes from both the simulated performance of the ideal searchers and from the performance of the human observers.

First, the expected error rates for the IPU-constrained searcher increased substantially over those of the unconstrained searcher, showing that IPU substantially impairs performance and that this impairment is exacerbated by high extrinsic uncertainty (i.e., large relevant set sizes). Second, the expected error rates rose sharply, approximately doubling for the IPU-constrained searcher in the Cluttered condition versus the Uncluttered condition. This result shows that the clutter manipulation was effective. Removing the clutter at irrelevant locations did in fact attenuate the effect of IPU in our search task. Third, the parameter-free predictions of the IPU-constrained searcher model accurately accounted for most of the trends of human performance across clutter and relevant set size conditions. This suggests that human observers, like the IPU-constrained searcher, are limited by IPU.

Though the clutter manipulation worked well as a method of attenuating and exacerbating the effects of IPU for the IPU-constrained searcher, and appeared to be successful in modulating the effects of IPU for the human observers as well, it introduced a potential confound in that the difference in frequency content between the notched noise and $1/f$ noise backgrounds in the Uncluttered condition was visible and might have served as a kind of spatial cue marking the potential target locations. Human observers might have used these "cues" as visual memory aids for the set of potential target locations or treated them as fixation targets. Moreover, because these cues are only available in the Uncluttered condition, we were concerned that they might account for some of the improvement in performance over the Cluttered location.

We ran the simultaneous-cue experiment to control for this possibility, by including small target location markers (detectable in the periphery) that were visible throughout the search trial in both the Cluttered and Uncluttered conditions. The results of the simultaneous-cue search task were very similar to those of the main search experiment, suggesting that location-cueing cannot account for the performance differences between the Cluttered and Uncluttered conditions. One minor difference between the main search experiment and the simultaneous-cue experiment was that the error rates for both human and simulated observers were consistently smaller across conditions in the simultaneous-cue experiment. The fact that this improved

performance occurred both in human and simulated searchers suggests that including the location cues may have helped the human observers more optimally select fixation locations. Recall that the simulated searchers were forced to use the same fixations the human observers used on each trial. Thus we would expect any improvement in fixation selection to boost performance for the simulated searchers as well as for their corresponding human observers.

Another source of uncertainty that can degrade search performance is target-size uncertainty (Judy, Kijewski, Fu, & Swensson, 1995). Though the size of the target was specified and fixed across all conditions of our search experiments, explicit target-size cues were not provided during the search displays. Thus it is possible that some *intrinsic* uncertainty regarding target size (i.e., due to poor memory for the target) contributed to the degraded performance of the human observers. If so, then the Uncluttered condition might inadvertently (i.e., via the visibility of the relevant-noise regions) include a simultaneous target-size cue that the Cluttered condition does not. However, although we cannot absolutely rule out effects of such intrinsic target-size uncertainty, we find it unlikely that this uncertainty was large enough to play a significant role in our results. First, the target size was fixed across experimental conditions and never changed throughout the duration of the experiment. Additionally, before any of the data included in the search performance plots were collected, each human observer had completed at least four 1-hr sessions (including three detection sessions with size cued explicitly) and a practice search session. Moreover, two of the three observers in the simultaneous-cue condition (YS and AB) had completed participation in the main experiment (which may have contributed to their improved fixation selection). By the time these observers began participation in the simultaneous-cue experiment, they already had a minimum of 10 hrs of experience searching for the target. Thus, we doubt that the performance of the human observers in the Cluttered conditions was degraded significantly by target-size uncertainty.

The predictions of the IPU-constrained searcher were not perfect. Though the performance trends were similar for the IPU-constrained searchers and the human observers, the ideal searchers consistently outperformed the human observers, particularly in the Cluttered condition of the search experiment. This gap between human performance and the model prediction is not surprising given that our model is normative model of search behavior that considers only peripheral sensitivity and IPU as limiting factors. There are undoubtedly other inefficiencies (e.g., limitations of visual memory, suboptimal integration across saccades) that limit human performance, but are not included in

the IPU-constrained searcher model. Another limitation of the IPU-constrained searcher model is the IPU of our human observers might be larger than the IPU that we built into the model. In the interest of expediency, we used the based our estimates of IPU on measurements from a previous study (Michel & Geisler, 2011) that used a different set of human observers.

Overall, our results fit in with a growing literature demonstrating the importance of including the effects of intrinsic position uncertainty in models of visual detection and search. We previously found that IPU can strongly impair the detection and localization of signal within a noisy environment in single-fixation search tasks (Michel & Geisler, 2011). In the current study, we showed that, despite the potential of eye movements to resolve IPU through sequential foveation of visual targets, the impact of IPU persists even under naturalistic search tasks involving sequences of voluntary saccades.

Keywords: clutter, crowding, ideal observer analysis, intrinsic position uncertainty

Acknowledgments

Commercial relationships: none.
Corresponding author: Melchi M. Michel.
Email: melchi.michel@rutgers.edu.
Address: Department of Psychology, Rutgers University, New Brunswick, NJ, USA.

Footnotes

[1] In a pilot version of this task, we allowed observers to make as many eye movements as they need to find the target within 12 sec. In this version of the task and observers required an average of 6 fixations to locate the target. Note that performance in visual search comprises a tradeoff between speed or duration (e.g., the number of fixations) and accuracy (i.e., the proportion of correct responses). Limiting the duration of the search effectively eliminates a degree of freedom from the task so that we can characterize performance in terms of accuracy.

[2] Strictly speaking, the assumption of temporally independent response noise above is incorrect. Within a trial, the (external) noise mask is fixed across fixations, so that the (internal + external) response noise should be correlated across fixations (Najemnik & Geisler,

2005), placing a lower limit on the amount of noise reduction that can be achieved by integrating across fixations. However, because the response noise in this task is dominated overwhelmingly by internal factors (i.e., measured human detection efficiency for the target signal was well below 5% over most of the visual field) and because the search was limited to allow only a small number of fixations, any performance differences between the independent model described above and a more accurate model that takes into account the effect of the static external noise should be negligible.

References

Ackermann, J. F., & Landy, M. S. (2013). Choice of saccade endpoint under risk. *Journal of Vision*, *13*(3):27, 1–20, doi:10.1167/13.3.27. [PubMed] [Article]

Ahumada, A. J., & Watson, A. B. (1985). Equivalent-noise model for contrast detection and discrimination. *Journal of the Optical Society of America A, Optics and Image Science*, *2*(7), 1133–1139.

Bex, P. J., Mareschal, I., & Dakin, S. C. (2007). Contrast gain control in natural scenes. *Journal of Vision*, *7*(11):12, 1–12, doi:10.1167/7.11.12. [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, doi:10.1118/1.1630493.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Burgess, A. E., & Ghandeharian, H. (1984). Visual signal detection. II. Signal-location identification. *Journal of the Optical Society of America A, Optics and Image Science*, *1*(8), 906–910, doi:10.1364/JOSAA.1.000906.

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience*, *17*(21), 8621–8644.

Clarke, A. D. F., Green, P., Chantler, M. J., & Hunt, A. R. (2016). Human search for a target on a textured background is consistent with a stochastic model. *Journal of Vision*, *16*(7):4, 1–16, doi:10.1167/16.7.4. [PubMed] [Article]

Cohn, T. E., & Lasley, D. J. (1974). Detectability of a luminance increment: Effect of spatial uncertainty. *Journal of the Optical Society of America*, *64*, 1715–1719, doi:10.1364/JOSA.66.001426.

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, doi:10.1364/JOSAA.2.000820.

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*(3), 425–451.

Egeth, H., Atkinsons, J., Gilmore, G., & Marcus, N. (1973). Factors affecting processing mode in visual search. *Perception & Psychophysics, 13*(3), 394–402, doi:10.3758/BF03205792.

Estes, W. K., & Wessel, D. L. (1966). Reaction time in relation to display size and correctness of response in forced-choice visual signal detection. *Perception & Psychophysics, 1*(5), 369–373, doi:10.3758/BF03207411.

Geisler, W. S., & Albrecht, D. G. (1992). Cortical neurons: Isolation of contrast gain control. *Vision Research, 32*(8), 1409–1410.

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.

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, doi:10.1167/6.9.1.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics.* New York, NY: Wiley.

Judy, P. F., Kijewski, M. F., Fu, X., & Swensson, R. G. (1995). Observer detection efficiency with target size uncertainty. *Proceedings of the SPIE, 2436,* 10–17.

Klein, S. A., & Levi, D. M. (1987). Position sense of the peripheral retina. *Journal of the Optical Society of America A, 4*(8), 1543–1553.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research, 39*(16), 2729–2737, doi:10.1016/S0042-6989(98)00285-5.

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997, July). Mr. Chips: An ideal-observer model of reading. *Psychological Review, 104*(3), 524–553.

Levi, D. M. (2008). Crowding-an essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654, doi:10.1016/j.visres.2007.12.009.

Lu, Z. L., & Dosher, B. A. (1999). Characterizing human perceptual inefficiencies with equivalent internal noise. *Journal of the Optical Society of America A, 16*(3), 764–778, doi:10.1364/JOSAA.16.000764.

Manjeshwar, R., & Wilson, D. (2001). Effect of inherent location uncertainty on detection of stationary targets in noisy image sequences. *Journal of the Optical Society of America A, 18*(1), 78–85.

Michel, M. M., & Geisler, W. S. (2009). Gaze contingent displays: Analysis of saccadic plasticity in visual search. *Society for Information Display Technical Digest, 40*(1), 911–914.

Michel, M. M., & Geisler, W. S. (2011). Intrinsic position uncertainty explains detection and localization performance in peripheral vision. *Journal of Vision, 11*(1):18, 1–18, doi:10.1167/11.1.18. [PubMed] [Article]

Morvan, C., & Maloney, L. T. (2012). Human visual search does not maximize the post-saccadic probability of identifying targets. *PLoS Computational Biology, 8*(2), e1002342, doi:10.1371/journal.pcbi.1002342.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature, 434,* 387–391, doi:10.1167/5.8.778.

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, doi:10.1167/8.3.4. [PubMed] [Article]

Nowakowska, A., Clarke, A. D. F., & Hunt, A. R. (2017). *Human visual search behaviour is far from ideal. Proceedings of the Royal Society of London B: Biological Sciences, 284*(1849), 20162767-6.

Palmer, J. (1994). Set size effects in visual search: The effects of attention is independent of the stimulus for simple tasks. *Vision Research, 34*(13), 1703–1721.

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-Size effect. *Current Directions in Psychological Science, 4*(4), 118–123, doi:10.1111/1467-8721.ep10772534.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*(1012), 1227–1268, doi:10.1016/S0042-6989(99)00244-8.

Paulun, V. C., Schütz, A. C., Michel, M. M., Geisler, W. S., & Gegenfurtner, K. R. (2015). Visual search under scotopic lighting conditions. *Vision Research, 113*(B Part), 155–168, doi:10.1016/j.visres.2015.05.004.

Peli, E., Yang, J., & Goldstein, R. B. (1991). Image invariance with changes in size: The role of peripheral contrast thresholds. *Journal of the*

Optical Society of America A, *8*(11), 1762–1774, doi:10.1364/ JOSAA.8.001762.

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, doi:10.1364/JOSAA.2.001508.

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, doi:10.1167/7.2.20. [PubMed] [Article]

Rosenholtz, R., Huang, J., Raj, A., Balas, B. J., & Ilie, L. (2012). A summary statistic representation in peripheral vision explains visual search. *Journal of Vision*, *12*(4):14, 1–17, doi:10.1167/12.4.14. [PubMed] [Article]

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision*, *4*(2–3), 183–207.

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, doi:10.3758/ BF03207369.

Tanner, W. P. (1961). Physiological implications of psychophysical data. *Annals of the New York Academy of Sciences*, *89*, 752–765, doi:10.1111/j. 1749-6632.1961.tb20176.x.

Treisman, A. M., & Gelade, G. (1980, January). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136, doi:10.1016/0010-0285(80)90005-5.

Verghese, P. (2012). Active search for multiple targets is inefficient. *Vision Research*, *74*, 61–71, doi:10.1016/ j.visres.2012.08.008.

White, J., Levi, D., & Aitsebaomo, A. (1992). Spatial localization without visual references. *Vision Research*, *32*(3), 513–526.

Wilson, H. (1993). Nonlinear processes in visual pattern discrimination. *Proceedings of the National Academy of Sciences, USA*, *90*(21), 9785–9790, doi: 10.1073/pnas.90.21.9785.