THE ROLE OF PERIPHERAL POSITION UNCERTAINTY IN OVERT VISUAL SEARCH

BY YELDA SEMIZER

A thesis submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Master of Science

Graduate Program in Psychology

Written under the direction of

Melchi M. Michel

and approved by

_____

_____

_____

New Brunswick, New Jersey

October, 2015

ABSTRACT OF THE THESIS

The Role of Peripheral Position Uncertainty in Overt Visual Search

by Yelda Semizer

Thesis Director: Melchi M. Michel

Uncertainty regarding the position of the search target is a fundamental component of visual search. This position uncertainty can be either extrinsic (EPU)–uncertainty regarding where a stimulus might appear, or intrinsic (IPU)–uncertainty regarding the distal source of the perceived stimulus. Previous measurements indicate that IPU increases approximately linearly with retinal eccentricity and that it accounts for impaired detection and localization performance in the periphery (Michel & Geisler, 2011). Our aim in the current project was to characterize the role of IPU in overt visual search and to determine whether it is a limiting factor in search performance. Human observers completed two tasks. First, the observers completed a detection task to measure sensitivity to the target as a function of visual field position. Then, they completed a search task, which required localization of the target signal within a noisy environment. Observers were allowed to make a maximum of six fixations. To examine the effect of IPU, two different experimental conditions were created. In the "cluttered" condition,

the display was tiled uniformly with feature clutter (in the form of $1/f$ noise) to maximize the effect of IPU. In the "uncluttered" condition, the clutter at irrelevant locations was removed to decrease the effect of IPU. The amount of EPU was also manipulated across conditions. We developed a constrained ideal searcher model, in which the searcher is limited by IPU measured for human observers. Individual ideal searchers were simulated for each human observer using the sensitivity measured in the detection task and the fixations sequences measured in the search task. Introducing IPU to the ideal searcher impaired overall overt search performance, but not uniformly. In the "uncluttered" condition, performance decreased steeply as a function of increasing EPU. However, in the "cluttered" condition, the effect of IPU dominated and performance flattened as a function of EPU. Measured performance for human searchers showed similar trends. Our findings suggest IPU as a limiting factor in overt search performance.

Acknowledgements

I would like to thank my advisor Dr. Melchi Michel, my committee members, Dr. Eileen Kowler and Dr. Manish Singh, and the members of the Computational Vision Lab for their insightful comments on this project and for their continuous support.

Table of Contents

List of Figures

Chapter 1
Introduction

Everyday tasks, which require either identification or detection of an object, such as searching for your car in a parking lot, for your keys in your backpack or for your friend within a crowd, typically involve some uncertainty about the position of the target object. Similarly, other more crucial tasks, such as a radiologist's search for a tumor in an X-ray image, a travel security agent's scan of a luggage for hazardous items, or an air-traffic controller's monitoring of the air-traffic require decision-making under great degree of uncertainty within highly noisy environments. Failure in these latter tasks can have detrimental consequences. For example, missing a tumor might result in the death of a patient while a false alarm might result in unnecessary treatment, which itself might lead to other severe health problems. An understanding of the sources and effects of uncertainty in such environments can aid the design of training and visual displays to improve the performance of these decision-makers. The role of position uncertainty in visual tasks and its detrimental effect on detection performance has been documented (Green & Swets, 1966; Peterson, Birdsall, & Fox, 1954; Tanner, 1961). However, there is still need to investigate how this uncertainty affects the visual search performance.

Visual search can be defined as an active scan of a visual environment for a specified target object (Eckstein, 2011). Visual search has been heavily studied in perceptual science (e.g., Geisler, Perry, & Najemnik, 2006; Najemnik & Geisler, 2008, 2009; Wolfe, Alvarez, Rosenholtz, & Sherman, 2011). The visual

environment is usually cluttered with distracters which may have similar features as the target object, making it hard to localize the object. One of the defining characteristics of visual search is an uncertainty about the position of the target. Therefore, when we talk about visual search, it is necessary to talk about position uncertainty. There are two types of position uncertainty: *extrinsic* and *intrinsic* (Michel & Geisler, 2011). Before providing the definitions, let's give an example to illustrate these two different types of position uncertainty. Assume that a friend asks you to meet him either "on campus" or "in the Psychology building". In this case, you would have extrinsic uncertainty, that is higher in the former compared to the latter because you have more possible locations in which you need to look for your friend. If you are not familiar with the campus and don't know where the department is, then this extra information would not help you to find your friend. This is similar to having intrinsic position uncertainty because you need to search most of the locations anyway. More generally, extrinsic position uncertainty (EPU) refers to a prior uncertainty regarding where a stimulus might appear in a visual environment. In a typical search task, even if the observer knows all of the possible target locations beforehand, he will nonetheless have some degree of uncertainty about which of the locations contains the target object. Therefore, EPU can be seen as a property of the task.

Intrinsic position uncertainty (IPU), refers to the uncertainty regarding the source of a perceived stimulus or feature within the visual environment. In other words, when an observer detects a feature the visual periphery, he might still have some degree of uncertainty regarding the exact spatial location of its distal source. Unlike EPU, IPU is a characteristic of the human visual system and can be seen as the property of the observer rather than of the task. This later characteristic inevitably makes IPU harder to study. Nonetheless, there is research showing that IPU increases linearly with increasing retinal eccentricity and it impairs detection

performance in the periphery (Michel & Geisler, 2011). However, the effect of IPU on performance within more natural environments still needs to be investigated. Hence, the primary goal of this study is to examine the effect of IPU on overt visual search performance.

Pelli (1985) reported that typically when researchers investigate uncertainty, they merely focus on extrinsic uncertainty and ignore intrinsic uncertainty. Studies showed that EPU increases linearly as the number of possible target locations increases. The detrimental effects of EPU on performance was found in a number of detection and localization experiments, both in the fovea in the context of luminance change detection (Cohn & Lasley, 1974) and in the periphery marked by various eccentricities in the context of signal-noise localization (Burgess & Ghandeharian, 1984; Swensson & Judy, 1981). The detection performance under EPU was much poorer than the performance under certainty where observers knew where the target will appear (Cohn & Wardlaw, 1985). This performance decline because of greater number of possible target locations is called "set-size effect", which refers to the performance drop as a result of an increase in the number of stimuli (Baldassi & Burr, 2000; Cameron, Eckstein, Tai, & Carrasco, 2004; Palmer, Ames, & Lindsey, 1993). Set-size effects have been reported in detection and identification tasks as well as in localization tasks (Cameron et al., 2004). When combined with the findings on EPU, the findings on set-size effects demonstrate that they should be taken into account to model an accurate representation of the visual system. However, a model which focuses only on EPU would be inadequate; therefore, we also need to represent IPU.

The aforementioned studies that merely focus on EPU fail to take IPU into account and assume that the observer is 100% sure about the exact stimulus. However, the ideal observer models which are limited only by EPU showed that their performance does not seem to align perfectly with actual human performance,

suggesting that there has to be another source of uncertainty (Shiffrin, McKay, & Shaffer, 1976). Tanner (1961) reported that even if the observers exactly knew the stimulus (signal), detection performance still suggested uncertainty. He developed an ideal observer model that has knowledge about all potential stimuli but is uncertain about the exact stimulus at a given time. Comparison of ideal and human observers' performances revealed similar psychometric functions, suggesting that the assumption that humans have exact stimulus information needs to be reconsidered. Later studies that modeled IPU reported that it increases approximately linearly with retinal eccentricity (e.g., Levi & Tripathy, 1996; Michel & Geisler, 2011; Pelli, Palomares, & Majaj, 2004).

The harmful effect of IPU on peripheral vision is also implied by the findings in visual crowding studies (Bouma, 1970; Levi, 2008; Van den Berg, Johnson, Anton, Schepers, & Cornelissen, 2012). Visual crowding refers to the phenomenon where identification of a peripheral object is impaired by its neighboring objects (for a comprehensive review, see Levi, 2008). For example, performance for letter identification is worse when the letter is presented with other adjacent letters than when it is presented alone, and this effect becomes larger as the retinal eccentricity increases (Bouma, 1970). Although investigations on crowding typically focus on letter identification tasks (e.g., Pelli et al., 2004), crowding is also evident in other perceptual processes, such as perception of natural scenes (Wallis & Bex, 2012). Research on crowding (Bouma, 1970; Pelli et al., 2004) and position uncertainty (Levi & Tripathy, 1996) suggest that performance drop in the periphery (depending on the retinal eccentricity) is much higher than it is expected when only the decline in the spatial resolution is considered. This implies that there might be another source of uncertainty that leads to the observed performance decrements and IPU is a strong candidate for this source (e.g., Michel & Geisler, 2011). Therefore, models for human visual system needs to take IPU into

consideration as well as EPU.

To investigate IPU, we can simulate normative or "ideal" observer with and without IPU and compare their performances under different visual environments. Figure 1.1 schematically represents the pattern of performance that we expect for these ideal observers under various degrees of EPU. For an "unconstrained" ideal observer, that is not limited by IPU, search performance should get worse as EPU increases. This performance drop results primarily from the increase in the possible number of target locations. As the number of possible target locations increases, the observer has to monitor more and more locations. Hence, the effective stimulus uncertainty increases, leading to reduced detection and discrimination performance (Pelli, 1985). When we introduce IPU to the ideal observer ("constrained"), the performance should similarly drop, since the observer is effectively monitoring extra target locations. More importantly, however, the effect of EPU should diminish when IPU is high. Because the observer is already monitoring extra target locations, the location uncertainty added by EPU is largely subsumed by the effect of IPU.

For the human observer, since we cannot remove IPU, what we can do is to create environments that either enhance or minimize the effect of IPU. This can be achieved by varying the amount or distribution of feature clutter in the scene. For example, if the clutter has similar features as the target and it is uniformly distributed across the scene, even if the observer perceives a signal at the visual periphery, the localization of the perceived signal will be harder. The neighboring areas around the perceived signal might be misinterpreted as the actual signal. Hence, the effect of IPU on localization performance is enhanced in such an environment. When the clutter which has similar features as the target is removed from the neighboring or irrelevant areas, the misinterpretation will be less likely. Once the observer perceives a signal, localization will be easier compared to the

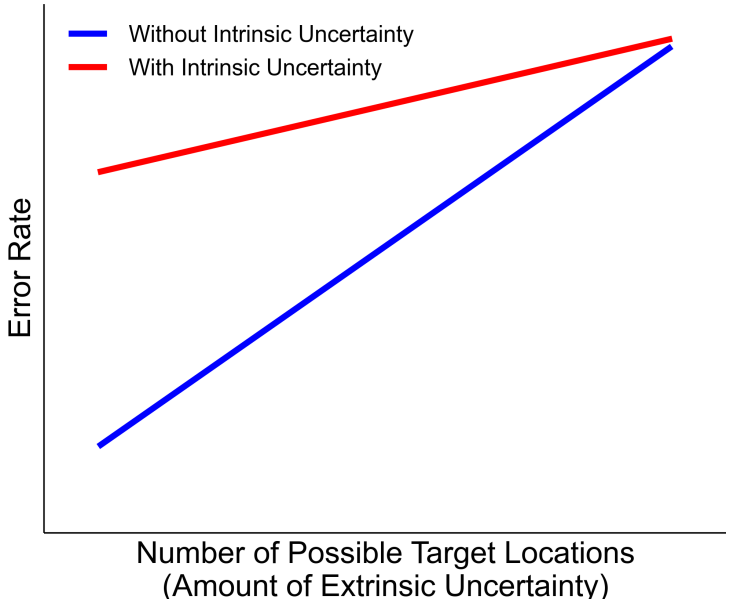

Figure 1.1: A schematic representation of error rate as a function of the number of possible target locations for observers with and without intrinsic position uncertainty, red and blue lines respectively. Note that performance of an observer without intrinsic uncertainty decreases as the amount of extrinsic uncertainty increases. Performance of an observer with intrinsic uncertainty is worse than the observer without intrinsic uncertainty. Also, it is less influenced by the amount of extrinsic uncertainty.

other environment because there is no features around the perceived signal that would lead to mislocalization. Hence, the effect of IPU on localization performance will be minimized in such an environment. After measuring performance of human and ideal observers across these conditions, we can compare them to determine whether IPU is a limiting factor in overt search performance, as we did in this study.

Chapter 2
Method

2.1 Observers

Four human observers participated in the data collection. One of the observers was an author; the other three were naïve to the purpose of the experiment. All observers had normal or corrected-to-normal vision and received compensation for their participation.

2.2 Apparatus

Stimuli were presented on a 22in Philips 202P4 CRT monitor with resolution 1280×1024 pixels at 100 Hz. The viewing distance was set to 70 cm from the observer so that the display subtended $15.8^\circ \times 21.1^\circ$ 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 500 and 1000 Hz, for detection and search experiments, respectively. A custom 13-point calibration routine where the target position was sampled within 11° of visual angle was used. If an eye movement (in the detection task) or a blink (both in the detection and search tasks) were detected during a trial, the observer was

notified and the trial was discarded.

2.3 Stimuli

The target was a 4 cycle per degree sine-wave grating, tilted 45° clockwise from the vertical orientation, presented at 20% contrast (see Figure 2.1). The background was a circular region 24° in diameter filled with 10% contrast (rootmean-square, RMS) luminance noise. Two different types of noise were used to fill the background. In the "cluttered condition", the background was filled with $1/f$ noise at a mean luminance of 40 cd/m^2, which adds feature clutter uniformly across the display. $1/f$ noise was created by filtering Gaussian white noise and truncating the noise at $\pm 2SD$ and then scaling to obtain desired RMS amplitude. In the "uncluttered condition", we used a notched spatial frequency filter which was centered on the spatial frequency of the target. Specifically, the filter was defined as a log-gaussian function with an octave bandwidth of 2. We applied this filter to $1/f$ noise to remove the part of the spectrum that overlaps with the target frequency band. We only removed the feature clutter at the irrelevant locations. In other words, we kept $1/f$ noise at the relevant locations (i.e., possible target locations) to make the two conditions comparable. Finally, the area around the circular region was set to mean luminance.

2.4 Procedure

2.4.1 2IFC Detection Task

To start each trial, observers had to fixate at the cross presented at the center of the screen. They had to retain their fixation during the trial. One of the 8 possible target locations was randomly cued with an open circle in each trial.

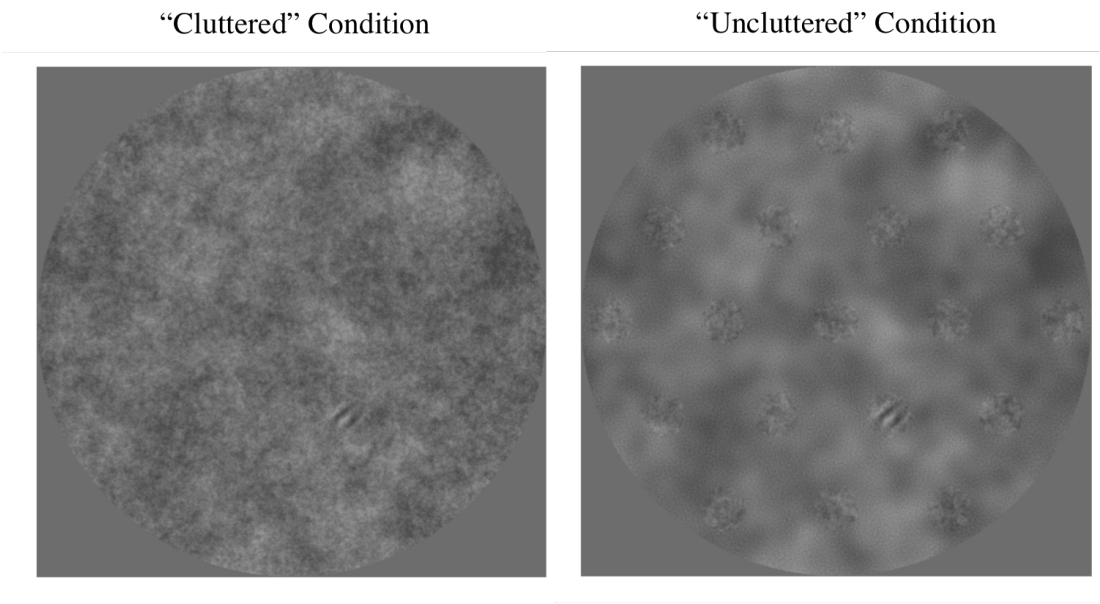

Figure 2.1: Example displays for the "cluttered" and "uncluttered" condition, with an embedded sine-wave grating as the target. In the "cluttered" condition, the display was tiled uniformly with feature clutter (in the form of $1/f$ noise) to enhance the effect of intrinsic position uncertainty. In the "uncluttered" condition, the clutter at irrelevant locations was removed to minimize the effect of intrinsic position uncertainty.

Observers initiated each trial with a key press. Once the key press was detected and observers' fixation was checked, two displays were flashed one after another, each for 250 ms with an inter stimulus interval of 500 ms, in which a blank screen was flashed. Only one of the displays contained the target embedded on a $1/f$ noise patch at the cued location while the other contained only the noise patch. The circular background was always filled with the filtered $1/f$ noise which was "uncluttered". We used an "uncluttered" background to minimize the effect of IPU on performance and thus get local estimates of sensitivity that are minimally influenced by IPU. Observers were instructed to decide which of the two displays contained the target as accurately and as quickly as possible by key press. They received auditory feedback. At the start of each block, observers completed 5 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 at each trial. The retinal eccentricity of the cue (the distance from the fixation point) varied across blocks. We used 4 different eccentricities ($\varepsilon = 0.0°, 2.5°, 5.0°, \text{and } 10.0°$). Finally, each session contained 7 blocks, each with 100 trials, for a total of 700 trials per session. The task was completed in 5 sessions. Stimulus sequence for this task can be seen in Appendix A1.

2.4.2 Search Task

To start each trial, observers had to fixate at the center cross but they were allowed to make eye-movements during a trial. All possible target locations were cued with circular markers at the beginning of each trial. Observers started each trial with a key press. Once the key press was detected and observers' fixation was checked, the stimulus display was presented for either a maximum number of 6 fixations were detected or 3 sec passed.[1] Observers were instructed to find the target as quickly and as accurately as possible. Next, cues for possible target locations were presented again, but this time with the dimmed stimulus display. Observers were required to fixate at the location of the target. As they moved their eye across the display, a green circle showed their eye position. Once their eye landed on the desired location, they logged their responses by a key press. They received both auditory and visual feedback. Observers' saccades and the actual target location were drawn on the stimulus display to serve as visual feedback. Observers started each block with completing 5 practice trials, for which the data were not recorded. The type of the background noise and the number of possible target locations varied across blocks. The background noise was either $1/f$ noise

[1]In another version of this task, we allowed observers to make as many eye movements as they need to find the target within 12 sec. One observer, the first author, completed this task and the average number of fixations was 6. The reason for fixing the number of fixations is that performance in the unrestricted search task comprises a tradeoff between speed (i.e., the number of fixations) and accuracy (i.e., the proportion of correct responses). Fixing the speed of the search effectively eliminates a degree of freedom from the task so that we can measure performance entirely in terms of accuracy.

which adds clutter uniformly across the display ("cluttered" condition) or filtered $1/f$ noise which does not contain clutter at the irrelevant locations ("uncluttered" condition). There were 5 different number of possible target locations (37, 85, 163, 421, and 817). Each session contained 10 blocks, each with 50 trials, for a total of 500 trials per session. The task was completed in 5 sessions. A schematic illustrating the stimulus sequence appears in Appendix A2.

The experiment took approximately 10 hour long sessions, each carried out on a separate day. Each observer started with 3 sessions of the 2IFC detection task (pre-test), then they completed 5 sessions of the search task, followed by 2 extra sessions of the 2IFC detection task (post-test). The first session of each task served as a practice day; hence, these data were excluded from the analysis. We measured detection performance in pre- and post-test sessions in order to account for any changes in the observers' sensitivity levels over the course of the experiment.

2.5 Intrinsic Uncertainty Searcher

To represent human visual system in our model, we developed two different ideal observer models, "constrained" and "unconstrained", which were limited by the factors that are known to have an effect on human observers' overt search performance. First, we know that human observers' sensitivity decreases in the periphery. Hence, both of these ideal observers were limited by the human observers' sensitivity measured as a function of retinal eccentricity. Only the "constrained" ideal observer was also limited by IPU measured previously for human observers (Michel & Geisler, 2011) while the "unconstrained" ideal observer had no IPU. This differentiation between two models allows us to investigate the effect of IPU on overt search performance, which cannot be achieved with human observers. Below we define the "constained" ideal observer. Since our task is a

search task, we will refer to this ideal observer as "constrained" ideal searcher.

2.5.1 The Generative Model

The search task for the "constrained" ideal searcher was the same as that for the human observer. We created a display filled with randomly distributed spatial noise. Then, we discretized the visual display into N locations, representing non-overlapping circular patches of the display. The task was to determine which of these locations contains the target in each trial. In the "uncluttered condition", these locations were the same as the possible target locations. In the "cluttered condition", we used more locations than the actual number of target locations to represent the greater possibility of mislocalization as a result of feature clutter.

We assumed that the perceived responses are influenced by the "equivalent internal noise" (Lu & Dosher, 1999, p. 764). In this context, equivalent internal noise refers to the uncertainty about the magnitude of the perceived response and it depends on the retinal position of the response. Here, it is represented as the visual sensitivity as a function of the retinal eccentricity. Hence, for each location i and fixation $k(t)$, the perceived response $R_{i,k(t)}$ is generated from a $1D$ Normal distribution, given by

$$R_{i,k(t)} \sim \begin{cases} \mathcal{N}(+0.5, \sigma) & \text{if } i = \text{target location}, \\ \mathcal{N}(-0.5, \sigma) & \text{if } i \neq \text{target location}, \end{cases}$$

where for a given target contrast level c and retinal eccentricity ε, $\sigma(c, \varepsilon)$ represents combined extrinsic and intrinsic response uncertainty, given by

$$\sigma(c, \varepsilon) = \frac{1}{\sqrt{2}\phi^{-1}\Big(1 - 0.5 \exp\Big[-\Big(\frac{c}{\alpha}\Big)^{\beta}\Big]\Big)}, \tag{2.1}$$

where ϕ^{-1} is the standard normal integral, and c and β represents the target contrast level and the slope of the estimated psychometric function, respectively. Note that for the perceived response $R_{i,k(t)}$, we set mean difference between signal and signal+noise to 1 and vary that by the visual sensitivity as a function of retinal eccentricity. We can use σ to represent uncertainty regarding the response because it can account for the fall-off of the accuracy of responses as retinal eccentricity increases.

Because of the spatial uncertainty represented by IPU, these responses were perceived as coming from locations which were shifted perceptually. For each location i and fixation $k(t)$, the perceived location $L_{i,k(t)} = [X_i, Y_i]$ is generated from a $2D$ Normal distribution which centered on the actual target location with a standard deviation of IPU, given by

$$L_{i,k(t)} = \big[X_{i,k(t)}, Y_{i,k(t)}\big] \sim \mathcal{N}\big[(x_i, y_i), \sigma_p(i, k(t))\big], \tag{2.2}$$

where σ_p represents IPU, given by

$$\sigma_p(\varepsilon) = m_p \varepsilon, \tag{2.3}$$

where ε is the retinal eccentricity, and m_p is a scalar value ($m_p = 0.09$, taken from Michel and Geisler (2011)). As can be seen from the equation, IPU is assumed to increase approximately linearly with increasing retinal eccentricity. We can use IPU to represent the distribution of perceived locations because it accounts for the displacement of locations, which becomes greater as the retinal eccentricity increases. To represent the ideal observer which is not limited by IPU, we set the standard deviation of the perceived locations to 0. Finally, we assumed a flat prior over all possible target locations: $i \in \text{Uniform}\big[0, N\big]$.

2.5.2 Computation of Signal Location

For a series of fixations $\phi_1, \ldots, \phi_T$, 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. We assume that the response noise and the location noise both depend on the retinal position of the target. The likelihood ratio is given by:

$$l(\mathbf{R}, \mathbf{L}) = \frac{p(\mathbf{R}, \mathbf{L}|k = i)}{p(\mathbf{R}, \mathbf{L}|k \neq i)}. \tag{2.4}$$

First, let's consider the likelihood for the first fixation.

$$\begin{aligned} p(\mathbf{R}, \mathbf{L}|J = j) &= \sum_{k=1}^{n} p(R_1, \ldots, R_n|j, k)\ p(L_k|J = j), \\ &= \sum_{k=1}^{n} p(L_k|j) \prod_{i=1}^{n} p(R_i|j, k), \end{aligned} \tag{2.5}$$

where

$$p(R_i|j, k) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma_r(i)} \exp\left[- \dfrac{(R_i + 0.5)^2}{2\sigma_r^2(i)} \right] & \text{if } i = k, \\ \\ \dfrac{1}{\sqrt{2\pi}\sigma_r(j)} \exp\left[- \dfrac{(R_i - 0.5)^2}{2\sigma_r^2(j)} \right] & \text{if } i \neq k. \end{cases} \tag{2.6}$$

Here, j and k index over possible target locations and encoded responses, respectively. $p(L_k|j)$ represents the probability of encoded location for R_k being generated from the location j and it depends on the retinal eccentricity. Since we represented the locations as discrete components, $p(L_k|j)$ is given by the integral over the region that is centered on the perceived location L_k for a Gaussian centered on the actual target location L_j with covariance $\sum_p(j)$. This distribution will become broader as the location L_j moves to the greater eccentricities because

the observer integrates over more locations. Moreover, since $p(R_i|j,k)$ is equal for all $i \neq k$,

$$\begin{aligned} p(\mathbf{R},\mathbf{L}|J=j) &= K\sum_{k=1}^{n} p(L_k|j)\ \frac{p(R_i|j,k=i)}{p(R_i|j,k\neq i)}, \\ &= K\sum_{k=1}^{n} p(L_k|j)\ \frac{\sigma_r(k)\exp\left[-\dfrac{(R_k-0.5)^2}{2\sigma_r^2(j)}\right]}{\sigma_r(j)\exp\left[-\dfrac{(R_k+0.5)^2}{2\sigma_r^2(k)}\right]}, \\ &= K\sum_{k=1}^{n} p(L_k|j)\ \frac{\sigma_r(k)}{\sigma_r(j)}\exp\left[\frac{(R_k+0.5)^2}{2\sigma_r^2(k)}-\frac{(R_k-0.5)^2}{2\sigma_r^2(j)}\right], \end{aligned} \tag{2.7}$$

where $K=\prod_{l=1}^{n} p(R_l|j,k\neq l)$.

Now, let's consider the likelihood for the sequence of fixations in temporally independent response and position noise:

$$\begin{aligned} p(\mathbf{R}_{\phi(1:T)},\mathbf{L}_{\phi(1:T)}|J=j) &= \prod_{t}^{T} p(\mathbf{R}_{\phi(t)},\mathbf{L}_{\phi(t)}|J=j), \\ &\propto \prod_{t}^{T}\sum_{k=1}^{n} p(X_{\phi(t)k}|j)\ \frac{\sigma_r(k)}{\sigma_r(j)}\exp\left[\frac{(R_{\phi(t)k}+0.5)^2}{2\sigma_r^2(k)}-\frac{(R_{\phi(t)k}-0.5)^2}{2\sigma_r^2(j)}\right]. \end{aligned} \tag{2.8}$$

What we really want to compute is the most probable target location. Using Bayes' rule we find that

$$p(J=j|\mathbf{R}_{\phi(1:T)},\mathbf{L}_{\phi(1:T)}) \propto p(J=j)\ p(\mathbf{R}_{\phi(1:T)},\mathbf{L}_{\phi(1:T)}|j). \tag{2.9}$$

Thus, if $p(J=j)$ is uniform, the MAP estimate of the target location is simply

$$\underset{j}{\operatorname{argmax}}\ p(\mathbf{R}_{\phi(1:T)},\mathbf{L}_{\phi(1:T)}|J=j). \tag{2.10}$$

2.5.3 Simulating the Uncertainty Searcher

Based on each human observer's visibility map and the fixations from the search task, each ideal searcher was simulated using Monte Carlo simulation. We started by setting $r_j = 0.5$ for the target location and $r_i = -0.5$ for all other locations, where $i \neq j$ to represent responses. We then generated random response variable vector as described in the Generative model section. For all possible target locations, we computed the likelihood of the response vector given the signal is present or absent to get the target present likelihood ratios. Next, by computing the joint likelihood for each possible target location and normalizing this likelihood, we obtained the posteriors.

The searcher decided whether the target was found by comparing Equation 2.7 to a threshold criterion (Green & Swets, 1966). If the target was not found, we updated the posterior and computed the optimal position for the next fixation by calculating the probability of being correct if a specific fixation location was chosen as the next fixation. Search was terminated when maximum posterior is greater than the posterior threshold (0.99) or when maximum number of fixation counts is performed (given by the number of human observer's fixations).

After simulating all trials for a given condition, we computed the threshold for the ideal searcher required to reach the human observer's accuracy. We simulated each trial within the human data 10 times.

Chapter 3

Results

3.1 Visibility Maps

To represent each observer's sensitivity level as a function of retinal eccentricity and angular direction, we constructed visibility maps using maximum likelihood estimation. We started by combining the data from pre- and post-test in the 2IFC detection task. We then fit a single psychometric function steepness parameter across all potential target locations using maximum likelihood estimation. We assumed that the slope would be similar across locations, based on previous work (Ackermann & Landy, 2013). Next, we used maximum likelihood to estimate each observer's foveal threshold. To estimate the threshold at all other locations, we used the following formula:

$$\alpha(\varepsilon) = \alpha(0)\exp(\tau_\theta \varepsilon), \tag{3.1}$$

where α is the threshold, ε is the retinal eccentricity 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 (Michel & Geisler, 2011; Peli, Yang, & Goldstein, 1991). We estimated the best-fitting τ_θ 's using maximum likelihood for each angular direction for which we collected data. Then, we used linear interpolation between angular and radial coordinates to compute

τ_θ 's across the visibility map. Next, we obtained the proportion correct responses given by the psychometric function. The psychometric function is modeled as a cumulative Weibull function given by

$$p = 1 - 0.5 \exp\Big[- \Big(\frac{c}{\alpha}\Big)^{\beta} \Big], \tag{3.2}$$

where β represents a steepness parameter and c represents target contrast. Also, we fixed c to 0.2 to represent the target contrast of 20% used in the search experiment. Finally, we obtained d' measurements by converting the proportion correct responses using the following formula:

$$d' = \sqrt{2}\phi^{-1}(p), \tag{3.3}$$

where ϕ^{-1} is the standard normal integral and p is the proportion correct responses.

Figure 3.1 represents the visibility map combined across all individual observers (for individual visibility maps, see Appendix A3). The visibility map shows that sensitivity is highest at the fovea and decreases as a function of retinal eccentricity. Also, observers are more sensitive along the horizontal axis than along the vertical axis.

3.2 Search Performance

We calculated each human observer's proportion correct responses in each of the 10 conditions (2×5, the type of background noise and the number of possible target locations, respectively). For each human observer, we simulated the "constrained" and "unconstrained" ideal searchers, both of which used individual visibility maps from the 2IFC detection task (pre- and post-test) and human

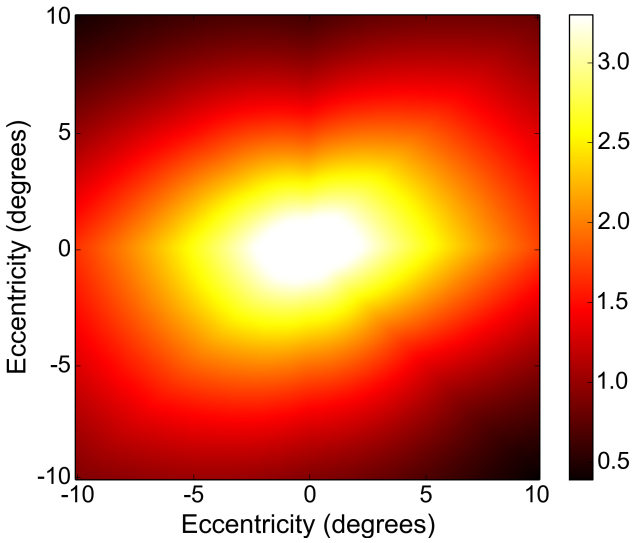

Figure 3.1: Average visibility map showing sensitivity (d') as a function of retinal eccentricity for the target contrast level of 20% across all participants. Note that sensitivity decreases as a function of retinal eccentricity as indicated by darker colors.

fixations from the search task.

Figure 3.2 represents the average search performance given by the error rate as a function of the number of the possible target locations in each of the background types (for individual data, see Appendix A4). As expected, the ideal searcher which is not limited by any intrinsic uncertainty performs better than all others. Performance of the constrained ideal searcher in the "uncluttered" condition gets worse as the number of possible target locations increases. Its performance somewhat flattens in the "cluttered" condition. Measured performance for human observers was also consistent with our expectations. They perform similarly to the ideal searchers although their error rate is higher. Human observers' poorer performance implies that there might be other limiting factors that our model does not account for.

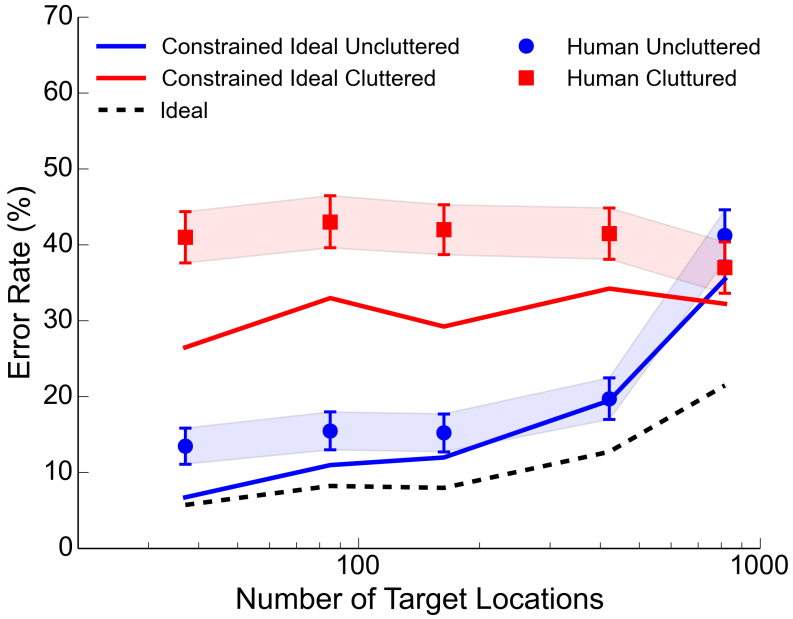

Figure 3.2: Average search performance across all four participants showing the error rate (%) as a function of the number of possible target locations in each background type. Error bars indicate SEMs. Note that ideal searcher without IPU (shown by the dashed line) outperforms all the others. Introducing IPU to the constrained ideal searcher impairs overall overt search performance, but not uniformly. Human observers perform worse than the ideal searchers but with the same pattern.

Chapter 4
Discussion

The purpose of the present study was to examine the effect of position uncertainty in overt search. Specifically, we investigated whether IPU is a limiting factor in overt search performance. Results suggest that IPU limits performance in our task. The evidence comes from both the ideal and human observers. For the human observers, when the environment does not include any irrelevant clutter ("uncluttered" condition), the performance seems to decrease as the observer needs to monitor more locations. When the environment is cluttered uniformly ("cluttered" condition), additional locations does not impair the performance; as a result, the performance somewhat flattens. This pattern of results was consistent across individual observers (see Appendix A4 for individual results). For the ideal searchers with IPU, the performance decline shows a similar pattern. More importantly, the ideal searcher without any IPU reveals the best performance, which is a clear indication that IPU limits overt search performance in our task.

Our results also show that the ideal searchers always outperform the human observers. Although this performance difference is not unexpected given that our model is an "ideal" model of human behavior, it might result from two main factors. First, there might still be some factors that limit overt search performance, which are not accounted by our model. Our model considers only the peripheral sensitivity and IPU as limiting factors. Another possibility is that the amount of IPU given to the ideal searchers might be less than IPU for our human observers. We used the value of IPU measured for human observers by Michel and Geisler

(2011).

Also, our findings can explain the performance decline in the periphery in crowding tasks. This implication might not be straightforward given the methodological differences between our task and the usual crowding tasks. The usual crowding tasks require a steady fixation at a central point with a goal to identify features in the periphery. In our overt search task, although the observer is allowed to make multiple fixations, he constantly gathers information from the periphery, similar to a classical crowding task. Additionally, there is already evidence showing that similar to the regions of crowding (Bouma, 1970; Levi, 2008), IPU increases approximately linearly in the periphery (Michel & Geisler, 2011).

Although our expectation that observers would perform better in the "uncluttered" condition than the "cluttered" condition was confirmed, there might be an alternative explanation for these findings. Namely, it is possible that the difference between the notched noise and $1/f$ noise backgrounds in the "uncluttered" condition can be used to segregate potential target locations from the background, thereby serving as a type of location cue. This cuing effect might itself improve search performance. To control for this possibility, we designed another version of the search task. In this new task, we used tiny location markers (0.14° degrees in diameter) as cues which were centered on the possible target locations in both conditions. These markers remained visible during the search displays. Two experienced and one naïve human observers completed this task. Then, we simulated corresponding ideal searchers for each observer through the same process defined above. Our findings show that although the observers seem to benefit from the cues in their fixation selection, the pattern of results remained similar to the earlier search task (see Appendix A5). This was also true for the ideal searchers.

These findings are in line with the earlier studies which show that while modeling human visual system, IPU should be taken into account as well as EPU (Pelli, 1985). It has been found that IPU influences detection and localization of signal within a noisy environment (Michel & Geisler, 2011). However, these tasks require steady fixations which are not sufficient to represent natural human behavior. Here, we were able to show that the effect of IPU persists even under a more natural task which involves eye-movements. The next step would be to investigate these effects with more natural images than ours or under more natural environments (e.g., outside the lab). Future studies are need to complete this step. As another future direction, measuring IPU directly and investigating individual differences would give additional evidence for this process.

In conclusion, this study measured overt search performance as a function of peripheral position uncertainty, for both extrinsic uncertainty regarding where a stimulus might appear (EPU), and intrinsic uncertainty regarding the distal source of the perceived stimulus (IPU). We developed a constrained ideal searcher model, in which searcher is limited by IPU measured for human observers. Our results suggest that IPU is a limiting factor in overt search performance and should be taken into account in the models of human visual system to prevent the possibility of overestimating the performance. These findings are also significant in terms of real life tasks. Position uncertainty is one of several factors which constraint real search performance within noisy environments. An understanding of the sources and effects of uncertainty in such environments can aid the design of training and visual displays to improve the performance of these decision-makers.

Chapter 5

Appendicies

A1. 2IFC Detection Task

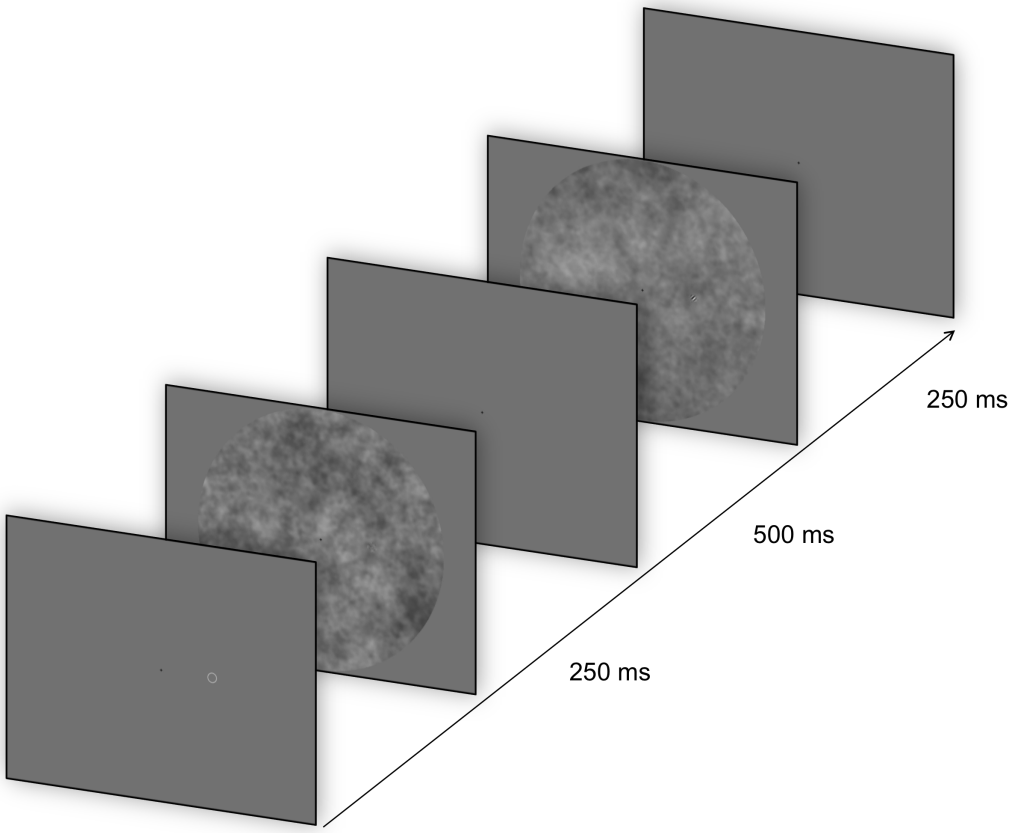

Figure A1: Trial sequence for 2IFC detection task. Target appears at the cued location in the second interval.

A2. Search Task

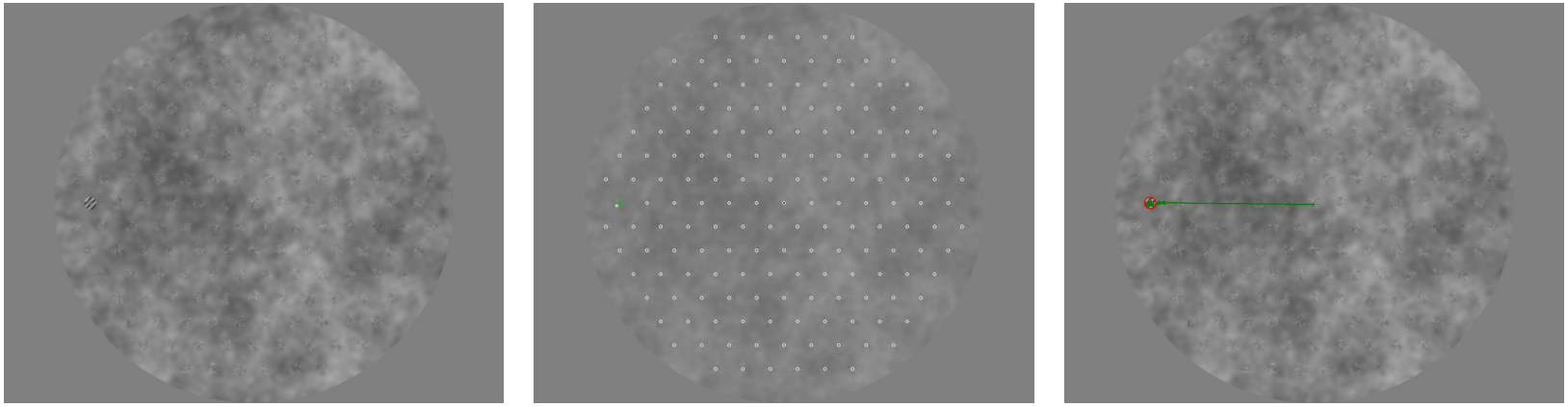

Figure A2: Trial sequence for search task with 163 locations in the "uncluttered" condition. Observer searches for the target in the first display. Then, in the second display, the observer reports the signal location by fixating at that location while pressing a key. Green dot represents the location of the observer's gaze. In the last display, feedback is given by the green arrows showing saccades and a red circle showing the actual target location, along with a sound.

A3. Visibility Maps

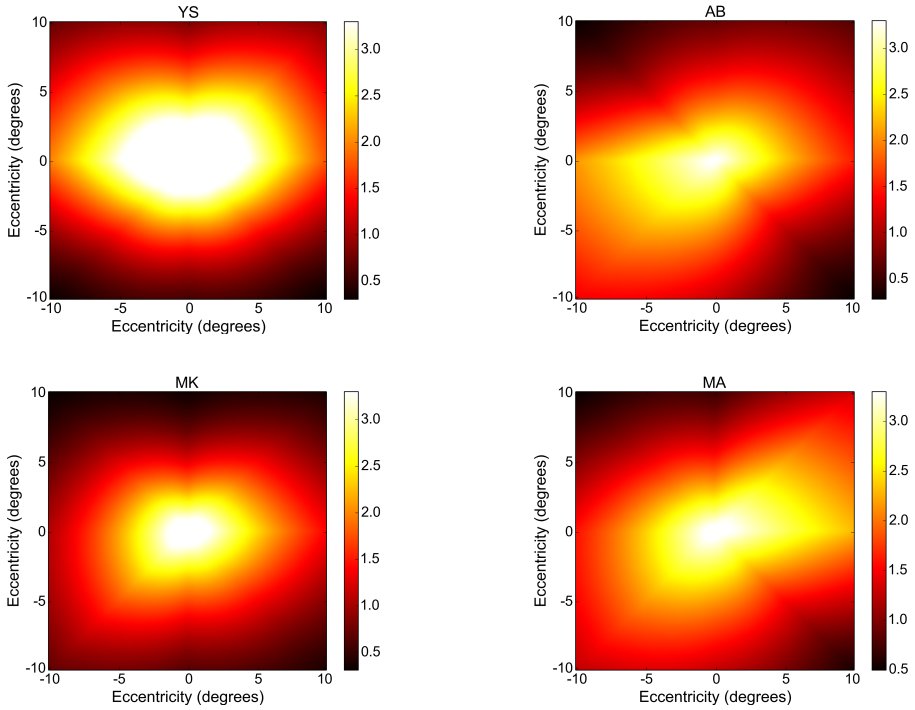

Figure A3: Individual visibility maps showing sensitivity (d') as a function of retinal eccentricity for the target contrast level of 20% across all participants. Note that sensitivity decreases as a function of retinal eccentricity, indicated by darker colors.

A4. Search Performance

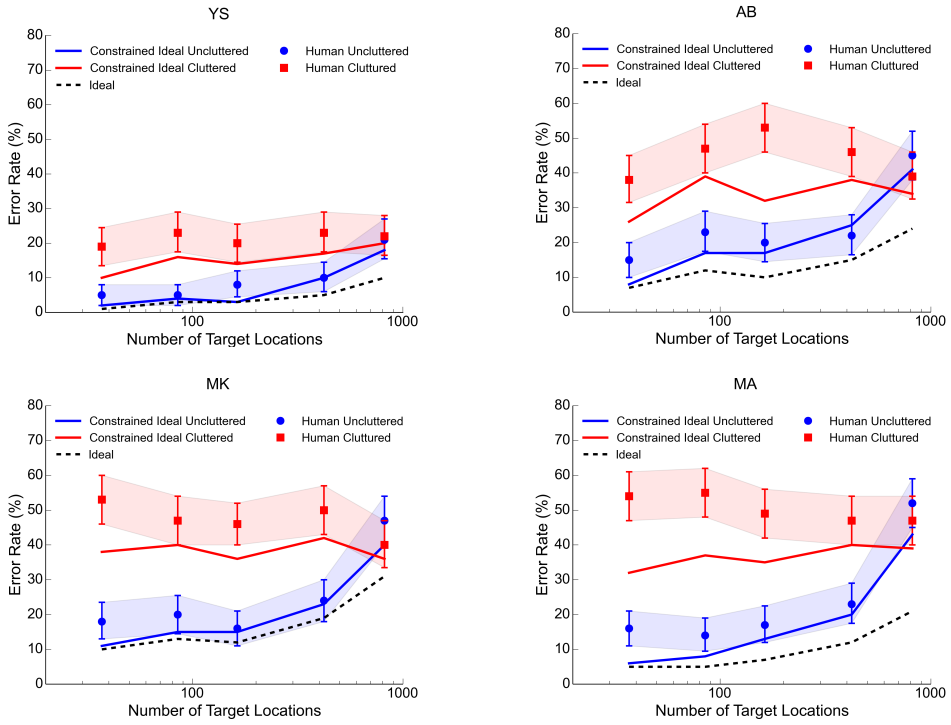

Figure A4: Individual search performance for four participants showing the error rate (%) as a function of the number of possible target locations in each background type. Error bars indicate SEMs. Note that ideal searcher without IPU (shown by the dashed line) outperforms all the others. Introducing IPU to the constrained ideal searcher impairs overall overt search performance, but not uniformly. Human observers perform worse than the ideal searchers but with the same pattern.

A5. Cued Search Performance

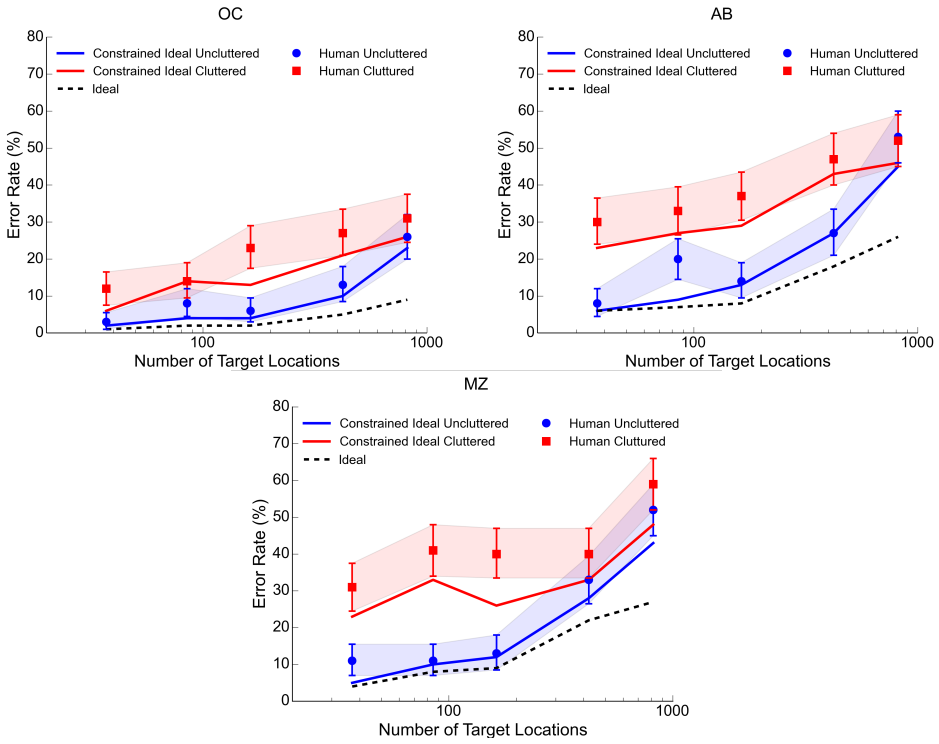

Figure A5: Cued Search Performance for three observers. Note that results show a similar pattern to that of the original search task.

References

Ackermann, J. F., & Landy, M. S. (2013). Choice of saccade endpoint under risk. *Journal of Vision*, *13*(3), 1–20. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/24023277` doi: 10.1167/13.3.27

Baldassi, S., & Burr, D. C. (2000). Feature-based integration of orientation signals in visual search. *Vision Research*, *40*, 1293–1300. doi: 10.1016/S0042-6989(00)00029-8

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, *226*, 177–178. doi: 10.1038/226177a0

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

Cameron, E. L., Eckstein, M., Tai, J., & Carrasco, M. (2004). Signal detection theory applied to three visual search tasks identification, yes/no detection and localization. *Spatial Vision*, *17*, 295–325. doi: 10.1163/1568568041920212

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. (2011). Visual search : A retrospective. *Journal of Vision*, *11*, 1–36. doi: 10.1167/11.5.14.

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), 858–873. doi: 10.1167/6.9.1

Green, D. M., & Swets, J. A. (1966). *Signal Detection Theory and Psychophysics*. New York: Wiley.

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

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

Levi, D. M., & Tripathy, S. P. (1996). Localization of a peripheral patch: The role of blur and spatial frequency. *Vision Research*, *36*, 3785–3803. doi: 10.1016/0042-6989(95)00343-6

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

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

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

Najemnik, J., & Geisler, W. S. (2009). Simple summation rule for optimal

fixation selection in visual search. *Vision Research*, *49*(10), 1286–1294. Retrieved from `http://dx.doi.org/10.1016/j.visres.2008.12.005` doi: 10.1016/j.visres.2008.12.005

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 108–130. doi: 10.1037/0096-1523.19.1.108

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., Palomares, M., & Majaj, N. J. (2004). Crowding is unlike ordinary masking: Distinguishing feature integration from detection. *Journal of Vision*, *4*(12), 1136–1169. doi: 10.1167/4.12.12

Peterson, W., Birdsall, T., & Fox, W. (1954). The theory of signal detectability. *IRE Professional Group on Information Theory*, *4*(4), 171–212. doi: 10.1109/TIT.1954.1057460

Shiffrin, R. M., McKay, D. P., & Shaffer, W. O. (1976). Attending to forty-nine spatial positions at once. *Journal of Experimental Psychology: Human Perception and Performance*, *2*, 14–22. doi: 10.1037/0096-1523.2.1.14

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

Van den Berg, R., Johnson, A., Anton, A. M., Schepers, A. L., & Cornelissen, F. W. (2012). Comparing crowding in human and ideal observers. *Journal of Vision*, *12*(8), 1–15. doi: 10.1167/12.6.13

Wallis, T. S. A., & Bex, P. J. (2012). Image correlates of crowding in natural scenes. *Journal of Vision*, *12*(7), 1–19. doi: 10.1167/12.7.6

Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention Perception Psychophysics*, *73*(6), 1650–1671. doi: 10.3758/s13414-011-0153-3