
Quantifying the performance limits of human saccadic targeting during visual search

Miguel P Eckstein[¶], Brent R Beutter, Leland S Stone

Human Information Processing Research Branch, Human Factors Research and Technology Division, NASA Ames Research Center, Moffett Field, CA 94035-1000, USA; [¶] also at Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA 93106, USA;

e-mail: eckstein@psych.ucsb.edu

Received 6 September 2000, in revised form 26 April 2001

Abstract. In previous studies of saccadic targeting, the issue how visually guided saccades to unambiguous targets are programmed and executed has been examined. These studies have found different degrees of guidance for saccades depending on the task and task difficulty. In this study, we use ideal-observer analysis to estimate the visual information used for the first saccade during a search for a target disk in noise. We quantitatively compare the performance of the first saccadic decision to that of the ideal observer (ie absolute efficiency of the first saccade) and to that of the associated final perceptual decision at the end of the search (ie relative efficiency of the first saccade). Our results show, first, that at all levels of salience tested, the first saccade is based on visual information from the stimulus display, and its highest absolute efficiency is $\sim 20\%$. Second, the efficiency of the first saccade is lower than that of the final perceptual decision after active search (with eye movements) and has a minimum relative efficiency of 19% at the lowest level of saliency investigated. Third, we found that requiring observers to maintain central fixation (no saccades allowed) decreased the absolute efficiency of their perceptual decision by up to a factor of two, but that the magnitude of this effect depended on target salience. Our results demonstrate that ideal-observer analysis can be extended to measure the visual information mediating saccadic target-selection decisions during visual search, which enables direct comparison of saccadic and perceptual efficiencies.

1 Introduction

Humans make frequent rapid eye movements—saccades—to point the high-resolution fovea at the location of current interest. A fundamental question is how the brain decides when to make a saccade and where to direct its endpoint—a process known as saccadic targeting. One visual task for which saccades are likely to play a particularly important role is search. Many studies of search have only measured perceptual performance (ie the reaction times and the percentage of correct perceptual decisions when looking for a ‘target’ object hidden among ‘distractor’ objects). Far less is known about saccades during active visual search. In previous studies, the accuracy of the first saccade during visual search has been measured in order to determine whether the first saccade is guided by displayed visual information about the target or by cognitive expectations, previous experience, or other a priori strategies (He and Kowler 1989; Hooge 1996; Zelinsky 1996; Findlay 1997). There have been somewhat disparate results. Some early studies found that color information guided eye movements, but shape information did not (Williams 1966, 1967). More recently, Zelinsky (1996) found little evidence for visual guidance in searches using symbolic displays. On the other hand, Viviani and Swenson (1982), as well as other more recent studies (Findlay 1997; Motter and Belky 1998; Hooge and Erklens 1999), found strong evidence for visual guidance in search. The difference in results might at least in part be due to the variations in element density in the displays (spaced versus cluttered), target salience, and task design.

The present work extends previous studies by shifting away from the earlier focus on the binary question whether search saccades are visually guided or not. The present objective is to quantify “how well” humans direct their first search saccade using

information about target location and to compare this performance with “how well” humans perform other visual tasks. A fundamental problem in determining “how well” humans target their first search saccades is that performance generally depends on the stimulus used and the task difficulty.⁽¹⁾ For example, in the case of extensively studied perceptual decisions such as discrimination of two signals (disks, Gaussian-windowed sine waves, etc) with differing contrast, performance is an increasing function of the contrast difference between the target and the nontarget. So, in one sense, the answer to the question “how good” are search saccades at using stimulus information in their targeting is the uninteresting statement “as good as the salience of the target”. One way to overcome this problem is to use ideal-observer analysis. Ideal-observer analysis allows the investigator to compute the best-possible performance for a given visual task. This ‘ideal’ performance defines an absolute reference standard with respect to which human performance can be evaluated and compared across levels of difficulty within a task and across visual tasks. Human observers’ performance can then be measured with respect to that of the ideal observer with a measure known as ‘efficiency’. Measuring human efficiency rather than absolute performance enables direct comparisons between the accuracy of first search saccades and that of visually based decisions in a wide variety of tasks previously reported (eg detection, contrast discrimination, letter identification).

In this study, we used ideal-observer analysis to make quantitative comparisons between a number of different performance decisions in a target localization task. Most previous oculomotor search studies have used high-salience ‘popout’ targets for which perceptual decisions are nearly always correct; however, because salience is known to affect perceptual decisions during search (Burgess and Ghandeharian 1984b; Eckstein and Whiting 1996), it is reasonable to assume that salience will have profound effects on search saccades as well. We therefore examined saccadic and perceptual performance across a wide range of target saliences. Our first goal was to quantify “how good” the first saccade is by comparing its accuracy with that of the ideal observer. Our second goal was to quantify the information acquired during an active search above and beyond that available for the first saccade by comparing the accuracy of the first saccade with that of the final perceptual decision after multiple fixations. Given that saccades are a critical component of normal search behavior, our third goal was to quantify the increase in efficiency of the perceptual decision achieved by active search with eye movements over passive search during central fixation. Although previous studies have shown that precluding observers from executing saccades to explore the visual scene can decrease search performance (eg Scialfa and Joffè 1998), we quantified the effect of allowing active search with eye movements on perceptual performance for a wide range of target saliences.

2 Methods

Stimuli were viewed binocularly on a Philips Brilliance 21A monitor, luminance linearized with a lookup table (mean: 35 cd m⁻²). On each trial, the target, a 21 min of arc diameter Gaussian-blurred ($\sigma = 3.5$ min of arc) disk, appeared with equal probability at the center of one of ten boxes (2.4 deg \times 2.4 deg) equidistant along a circle of radius 5.9 deg (figure 1a). Gaussian-distributed, spatially uncorrelated (white), static luminance noise was added to each pixel (RMS contrast = 26%). Target (signal) contrast was adjusted to achieve three different signal-to-noise ratios (SNRs): 2.99, 5.19, and 7.26. In the case of white noise, the SNR is computed by taking the square root of the sum

⁽¹⁾For example, human performance in a ten-alternative forced-choice task might be 60%, while performance for the same target in a two-alternative forced-choice task might be 89%. This difference, however, would not reflect any difference in how well observers perform the underlying visual search, but rather on the complexity of the task.

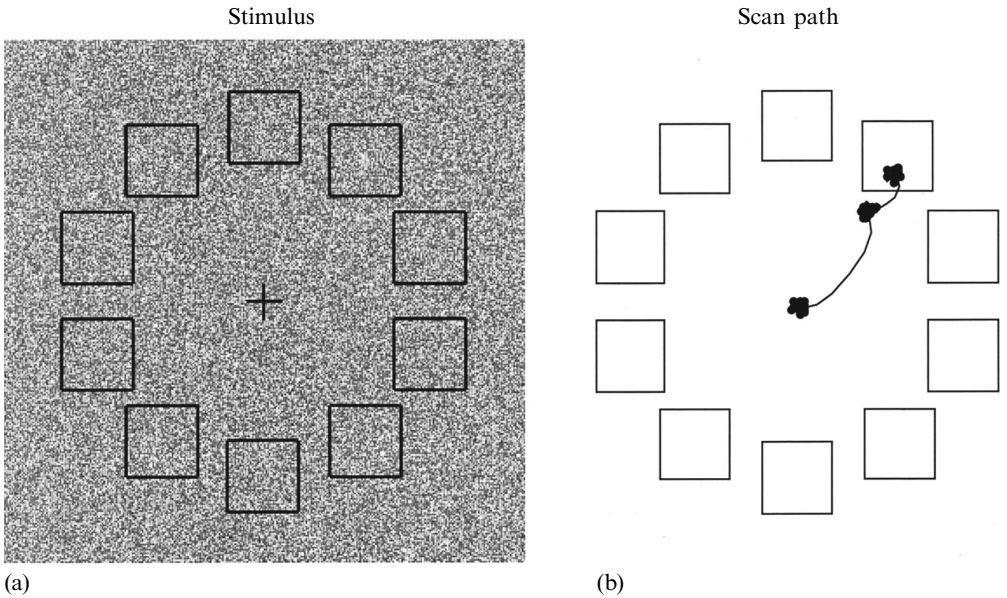


Figure 1. (a) Example high-SNR stimulus. (b) Scan path of a single high-SNR trial for observer ME, illustrating the first two search saccades (thin lines) and fixations (clusters of eye-position samples).

of the squared target-pixel contrasts and dividing by the standard deviation of the pixel-noise contrast (see Appendix).

Observers started each trial by fixating a small cross. With a mouse press, they then triggered the presentation of a 'test image' (figure 1a) and with another mouse press they indicated when they found the target. After responding or after 6 s (whichever came first), the test image was replaced by a 'response image' that contained the ten empty boxes and a rotatable central arrow. Observers used mouse buttons to point the arrow towards their ten-alternative forced choice (10AFC) and to record their decision. In a second, control, experiment run at a single SNR of 2.99, we imposed fixed stimulus display times of 4 and 6 s.

The position of the left eye was measured by an infrared video-based eye tracker sampling at 240 Hz (ISCAN Inc, NASA prototype), synchronized with the 60-Hz monitor. Eye-tracker calibration was performed by using nine crosses arranged in a 12 deg by 12 deg grid (precision was within 0.2 deg). Head movements were minimized by means of a bite bar. For each trial, the eye position from the initial trial fixation was used to correct for any small residual head movements. Saccades larger than 0.7 deg were detected with a digital template.

In the 'eye-movement' condition, observers were allowed to make eye movements, but no specific eye-movement strategy was encouraged. In the 'fixation' condition, observers were instructed to fixate the central cross at all times during the trial. Observers participated in three sessions, each consisting of six runs (two conditions \times three SNRs) of a hundred trials performed in a randomly permuted order. Each run consisted of a single block of a particular SNR and condition. Eye movements were monitored to record saccade endpoints (measured as the average location of the eye-position samples of the postsaccadic fixation), and to discard trials with anticipatory saccades (< 90 ms latency) in the eye-movement condition and trials in which observers broke central fixation (1.7 deg threshold) in the fixation condition. Because of the parametric nature of the experiments and the general consistency across observers, data were only collected from the three authors, each with normal vision.

2.1 Eye-position analysis

To facilitate the comparison of the first saccadic decision with the perceptual 10AFC decision, each saccadic endpoint was assigned to one of the ten possible target locations. We initially explored two criteria. The ‘direction’ criterion considered the first saccade correct if its endpoint was closer to the target than to any of the distractors. The advantage of this criterion is that it uses the true first saccade without penalizing for saccadic hypometria, eg the actual first saccade in figure 1b is deemed correct by this criterion. The disadvantage is that a saccade deemed correct could correspond to a deliberate saccade to a location corresponding to neither target nor distractor. The ‘box’ criterion considered the first saccade that landed inside a box. It was deemed correct if the box contained the target. The advantage of this second criterion is that it corresponds to an unambiguous target selection. The disadvantage is that the measured performance does not necessarily correspond to the true first saccade, eg the actual second saccade in figure 1b is deemed the correct ‘first’ saccade for this trial by this criterion. However, our pilot analysis revealed no significant difference between the performances measured with these two criteria ($p < 0.05$, paired t tests, Bonferroni corrected), reflecting the fact that the number of trials for which these two criteria yielded different decisions was small compared to the performance variation across sessions. Indeed, even when the two criteria selected different ‘first’ saccades (the situation in figure 1b), the saccadic decisions were nonetheless usually the same. Hereafter, we report saccadic decision accuracy using only the direction criterion.

2.2 Comparing human performance to the ideal-observer performance

Because a crucial aspect of the present work involves making a number of performance comparisons (human versus ideal; first saccades versus final perceptual decisions; search with eye movements versus search without eye movements), the question arises as to what measure of accuracy is appropriate for these comparisons. Because the difficulty of a given visual task depends not only on the nature of the task, but also on the target and distractor salience, this choice is critical. One possibility would simply be to take ratios of the proportion of correct trials for the two conditions [ie $P_{c(\text{human})}/P_{c(\text{ideal})}$]. Given that we used a constant level of external noise for all SNRs, an observer with suboptimal performance produced by a constant level of inefficiency resulting from limitations in spatial resolution (eg the contrast-sensitivity function) and/or internal noise (eg the temporal variance in neural responses) would generate P_c ratios that vary dramatically as a function of signal salience (see Appendix and figure A1). In other words, a ratio of $P_{c(\text{human})}$ to $P_{c(\text{ideal})}$ that changes with target salience (or number of locations) does not indicate that the nature of the underlying information processing is actually varying, because any inefficiency in the process could nonetheless be constant. In the present work, we applied statistical decision theory and ideal-observer analysis (Green and Swets 1966) to provide a quantitative metric of saccadic performance, which takes into consideration task difficulty (eg varying salience and set size).

3 Results

3.1 Percent-correct accuracy analysis

Figure 2a shows the accuracy of the first saccadic decision for all three observers. Figure 2b shows the accuracy of the final perceptual decision in the eye-movement condition. Final perceptual accuracy was always significantly higher (t test, $p < 0.05$) than initial saccadic accuracy (except for LS at the highest SNR), and both increased with increasing SNR. This SNR effect cannot be accounted for by a speed–accuracy trade-off because both perceptual mean reaction times and saccadic latencies increased with decreasing SNR (from 0.9 to 3.6 s and from 217 to 264 ms, respectively, averaged

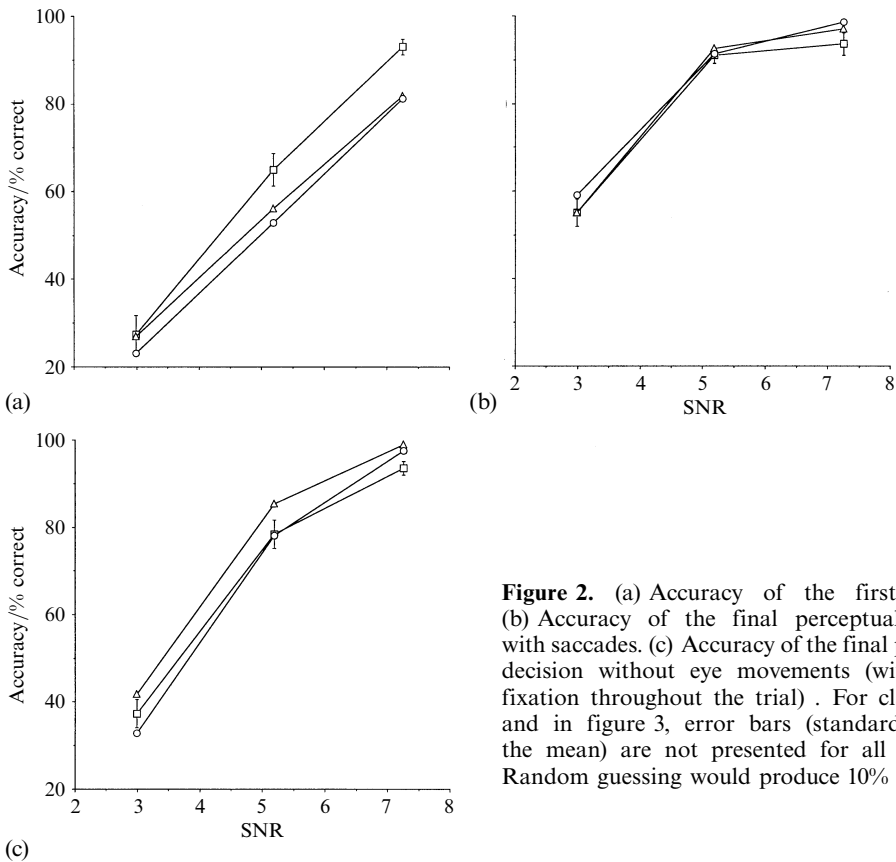


Figure 2. (a) Accuracy of the first saccade. (b) Accuracy of the final perceptual decision with saccades. (c) Accuracy of the final perceptual decision without eye movements (with central fixation throughout the trial). For clarity, here and in figure 3, error bars (standard error of the mean) are not presented for all observers. Random guessing would produce 10% accuracy.

across observers). Figure 2c shows the final perceptual accuracy without the benefit of eye movements under otherwise identical conditions. Final perceptual accuracy was always significantly higher ($p < 0.05$) in the eye-movement condition (figure 2b) than in the fixation condition (figure 2c) at the lower two SNRs, but not at the highest SNR. On average, reaction times were faster in the fixation (~ 2.2 s) than in the eye-movement (~ 3.6 s) condition at the lowest SNR. However, the performance enhancement from multiple fixations cannot be attributed to speed–accuracy trade-off. In a control experiment that imposed a fixed viewing time of 4 or 6 s in both conditions at a single SNR of 2.99, eye movements still dramatically increased performance (on average, from 44.1% to 58.3% at 4 s and from 46.2% to 60.7% at 6 s).

3.2 Efficiency analysis

To allow meaningful quantitative comparisons between perceptual, saccadic, and ideal decisions, we transformed percent correct into the d' index of detectability, defined as the distance in standard-deviation units between the target and distractor response distributions [we used a lookup table for a 10AFC task, see Green and Swets (1966)]. Unlike direct comparisons of percent correct, the squared d' ratios (efficiencies) provide quantitative measures of the relative performance, independent of the number of possible target locations or distractors (Burgess and Ghandeharian 1984b; see Appendix). Absolute efficiency specifies performance relative to the ideal observer, with d'_{ideal} being simply the SNR. The absolute efficiency of the first saccade, $(d'_{\text{saccade}}/d'_{\text{ideal}})^2$, ranged from 4% to 20% (across observers and SNRs) while that of the final perceptual decision, $(d'_{\text{perceptual}}/d'_{\text{ideal}})^2$, ranged from 20% to 38% (eye-movement condition) and from 10% to 33% (fixation condition). The relative efficiency of the first saccade

versus the final perceptual decision in the eye-movement condition, $(d'_{\text{saccade}}/d'_{\text{perceptual}})^2$, increased with SNR and reached a mean of 60% (across observers) at the highest SNR (figure 3a). The relative efficiency of the final perceptual decision in the fixation condition versus that in the eye-movement condition also increased with SNR, reaching a mean of 104% (although not significantly different than 100%) at the highest SNR (figure 3b). At the lowest SNR, when tested with fixed presentation times, the average (\pm SE across observers) relative efficiency was $0.57 \pm 3\%$ and $0.58 \pm 5\%$ for 4 and 6 s presentations, respectively, indicating that this less-than-unity efficiency does not result from a speed–accuracy trade-off. Lastly, the relative efficiency of the first saccadic decision with respect to the final perceptual decision in the fixation condition (figure 3c) is less than unity and appears constant with SNR for two of the three observers.

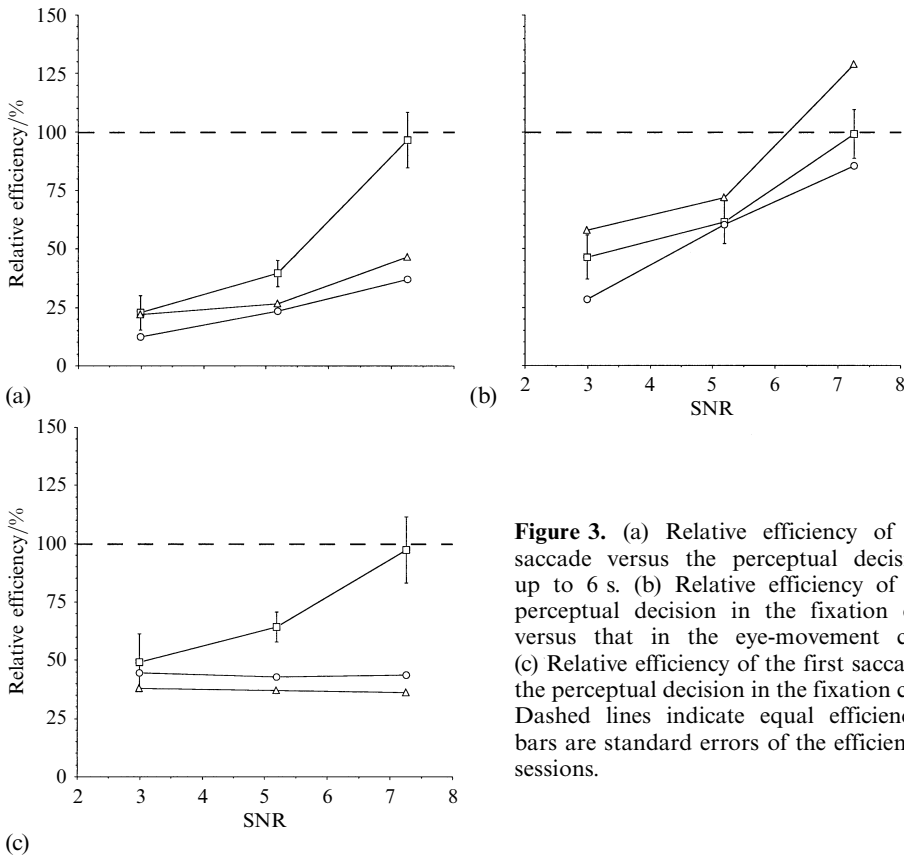


Figure 3. (a) Relative efficiency of the first saccade versus the perceptual decision after up to 6 s. (b) Relative efficiency of the final perceptual decision in the fixation condition versus that in the eye-movement condition. (c) Relative efficiency of the first saccade versus the perceptual decision in the fixation condition. Dashed lines indicate equal efficiency. Error bars are standard errors of the efficiency across sessions.

4 Discussion

4.1 Absolute efficiency

Our data show that the decision accuracy of the initial search saccade was higher than chance for all observers at all levels of salience tested. These results agree with those of previous authors who examined search for a disk embedded among dimmer disks (Findlay 1997) and orientation-, color-, and conjunction-defined targets (Motter and Belky 1998). We have extended these previous studies by using statistical decision theory to measure the information in the display available to an ideal observer and to compare perceptual and saccadic performance with this theoretical optimum. The absolute efficiency of the initial saccadic decision in our task was 4%–20%, which can be directly compared with other behavioral performance—eg 10%–56% for disk

detection in noise (Burgess and Ghandeharian 1984b; this study), 70% for contrast discrimination (Burgess et al 1981), 12%–20% for letter identification (Solomon and Pelli 1994), and 3%–8% for object recognition (Tjan et al 1995). The efficiency of the first saccade most likely will depend on display parameters such as element eccentricity and density, on the spatial characteristics of the target and distractors, and on the visual task (detection, contrast discrimination, identification, etc). It may also be a function of saccade latency in the same way that perceptual decisions are subject to a speed–accuracy trade-off. Further studies are required to resolve these issues.

4.2 *First saccade versus final perceptual decision*

Our results show that relative efficiency of the first saccadic decision versus the final perceptual decision for the eye-movement condition increases with SNR. One could argue that the low efficiency of the first saccadic decisions relative to the perceptual decisions at the lower SNRs might reflect a fundamental difference in the type and quality of the visual information processing used by the saccadic and perceptual systems. The concept of parallel visual pathways with different underlying processing abilities is not unprecedented. For example, the parvocellular pathway is not as sensitive to luminance-based motion (high-temporal-frequency, low-spatial-frequency stimuli) as the magnocellular pathway (Merigan and Maunsell 1993). Indeed, some have explicitly proposed separate visual pathways for perception and motor action (Goodale and Milner 1992; for an alternate view see Krauzlis and Stone 1999). However, the difference in the accuracies of the perceptual and the first saccadic decisions might be more simply explained by differences in the visual information available to make the decisions. There simply is more time to view the stimulus prior to the final perceptual decision than prior to the first search saccade, and the multiple fixations available for the perceptual decision allow access to high-resolution foveal processing of the possible target locations, which is not available when making the first saccadic decision.

To evaluate the impact of these two differences (long versus short viewing time and foveal versus peripheral processing), we compared the relative efficiency of the first saccadic decision with that of the perceptual decision in both the fixation and eye-movement conditions. When the perceptual decision is constrained to the use of only peripheral processing (fixation condition), the relative efficiency of the first saccade is higher than when foveal processing is allowed (eye-movement condition). Averaged across observers, it increases from 19.1% to 43.9% at the lowest SNR and from 29.9% to 48.1% at the middle SNR. For the highest SNR, the relative efficiency remains approximately the same (60.1% versus 59.1%). These results suggest that for the two lowest SNRs, the peripheral processing underlying the first saccadic decision accounts for much of its inefficiency with respect to the final perceptual decision after multiple fixations. However, the less-than-unity relative efficiency with respect to the perceptual decision generally observed even during central fixation (figure 3c) demonstrates that peripheral versus foveal processing cannot account for all of the poorer saccadic performance.

The second simple factor possibly contributing to the higher performance of the perceptual decision relative to the first saccadic decision is the fact that the saccadic decision is based at best on a few hundred milliseconds of visual processing, while the perceptual decision is based on many seconds of visual processing. To control for this difference, one must restrict perceptual processing to the brief time available for the first saccade. We are currently exploring this issue. Preliminary results suggest that the relative efficiency of the first saccadic decision with respect to the perceptual decision is close to unity when the stimulus duration is restricted to match the time available for saccadic processing (Stone et al 1999). Together with the current findings, this result suggests that the peripheral nature and the brevity of the presaccadic central viewing can largely account for the low efficiency of the first saccadic decision relative to that of the

perceptual decision after multiple fixations without the need to postulate a difference in the quality of the visual information-processing mechanisms.

4.3 *Perceptual decision with and without eye movements*

A number of studies have shown that visual-detection performance degrades as the stimulus location increases in retinal eccentricity in tasks ranging from the detection of sinusoidal patterns (Robson and Graham 1981) to the detection of letters (Scialfa et al 1987). In this context, one reason that perceptual decisions in visual tasks improve when eye movements are allowed is because saccades allow foveation of the possible target locations (Scialfa and Joffe 1998). However, it is unknown how much or under what circumstances eye movements will improve the absolute efficiency of the perceptual decision during search. Our finding that the absolute efficiency of the final perceptual decision was lower in the fixation condition (no eye movements) than in the eye-movement condition confirms that foveation or near foveation can play an important role in increasing efficiency. Our results show that the benefit of the multiple foveations provided by saccades is most prominent at low saliences (SNRs), reaching approximately a factor of two increase in absolute efficiency at the lowest salience. On the other hand, multiple foveations provided by saccades have less of an effect on the final perceptual decisions at high SNRs.

4.4 *Implication of reduced efficiency with decreasing target salience*

The inefficiencies causing human performance to be lower than that of the ideal observer are typically modeled as resulting from several sources: (i) the observer's use of a sub-optimal filter (template) to process the image data [ie a suboptimal sampling efficiency; Burgess et al (1981)], (ii) loss of low-spatial-frequency and high-spatial-frequency information due to the contrast-sensitivity function (Burgess 1994), and (iii) an additive or proportional internal noise source (Burgess et al 1981; Pelli 1981). A common property of all of these sources of inefficiency is that they degrade efficiency equally at all levels of target salience and therefore predict constant absolute efficiencies at all levels of target salience.⁽²⁾ Approximately constant efficiency as a function of signal contrast is found for detection of aperiodic signals in noise and contrast-discrimination tasks (Burgess and Ghandeharian 1984b). If the lower relative efficiency of saccades versus perceptual decisions is attributed to differences in one or more of these three sources of inefficiency, we would expect relative efficiency to be independent of target salience. However, as stated above, at low salience, the relative efficiency of the first saccade versus the final perceptual decision in the eye-movement condition is low, while at high salience, it is closer to unity. Therefore these results cannot be explained by any of the three common sources of inefficiency described above. The same argument can be made for the relative efficiency of the perceptual decision in the fixation condition versus that in the eye-movement condition.

So how might the variable efficiency be explained? Similar reductions in efficiency with signal salience have been found for detection of periodic signals (Burgess and Ghandeharian 1984a). One common interpretation for decreases in efficiency with low signal salience is intrinsic uncertainty about target location or phase (Tanner 1961; Pelli 1985). Such 'spatial uncertainty' will degrade performance for lower signal contrasts more than for higher signal contrasts (see Pelli 1985; Eckstein et al 1997). In this uncertainty framework, the interpretation of our results is that saccadic and perceptual decisions based on peripheral processing are subject to more spatial uncertainty than perceptual decisions based on foveal processing. This greater uncertainty would then lead to a decrease in relative efficiency with decreasing salience. The idea of higher

⁽²⁾This assumes constant external noise, consistent with our methods. If salience is changed by changing external noise rather than signal strength, then constant internal noise will not yield constant efficiency across SNR.

spatial uncertainty in the periphery than in the fovea is consistent with previous studies (Hess and Hayes 1994). Furthermore, this view is consistent with our finding that the relative efficiency of saccades versus perception in the fixation condition does not vary with target saliency (for two of the three observers), because, in this case, both the first saccadic and final perceptual decisions are based on peripheral processing and so presumably have similarly high levels of spatial uncertainty (figure 3c). Another possible explanation of the decreases in efficiency at low salience is that non-linear transducers may mediate contrast coding (Lu and Doshier 1999). However, this interpretation requires that the nonlinear transducer mediating peripheral perceptual and saccadic decisions be different from that mediating foveal decisions.

4.5 Physiological implications

The performance limits of the first saccade during search must reflect neural information about target location. A number of cortical and subcortical areas have recently been implicated in saccadic targeting within search and search-like paradigms, most notably, the frontal eye fields or FEF (eg Schall and Hanes 1993), the superior colliculus (eg Basso and Wurtz 1997), and the parietal (eg Gottlieb et al 1998) and inferior temporal (eg Chelazzi et al 1993) cortex. Within this context, we propose a tool to compare saccadic targeting with neural and perceptual responses using the same metric. The ability of a neuron (or neuronal population) to locate the target could be estimated by computing the mean and standard deviations of the neuron's responses when the target is in its receptive field and when a distractor is in its receptive field. This neural d' could then be directly compared with those of saccadic and perceptual decisions. This method differs from the receiver operating characteristic (ROC) analysis (Green and Swets 1966) used by others (Thompson et al 1996) to measure target–distractor discriminability by FEF neurons.⁽³⁾

References

- Barlow H B, 1957 "Incremental thresholds at low intensities considered signals/noise discrimination" *Journal of Physiology* **136** 469–488
- Basso M A, Wurtz R H, 1997 "Modulation of neuronal activity by target uncertainty" *Nature* **389** 66–69
- Butter B R, Eckstein M P, Stone L S, 2000 "Performance differences in contrast discrimination and detection visual search are similar for saccades and perception" *Investigative Ophthalmology & Visual Science* **41**(4) S424
- Burgess A E, 1994 "Statistically defined backgrounds: performance of a modified nonprewhitening observer model" *Journal of the Optical Society of America A* **11** 1237–1242
- Burgess A E, Ghandeharian H, 1984a "Visual signal detection. I. Ability to use phase information" *Journal of the Optical Society of America A* **1** 900–905
- Burgess A E, Ghandeharian H, 1984b "Visual signal detection. II. Effect of signal location identification" *Journal of the Optical Society of America A* **1** 906–910
- Burgess A E, Wagner R F, Jennings R J, Barlow H B, 1981 "Efficiency of human visual signal discrimination" *Science* **214** 93–94
- Chelazzi L, Miller E K, Duncan J, Desimone R, 1993 "A neural basis for visual search in inferior temporal cortex" *Nature* **363** 345–347
- Eckstein M P, 1998 "The lower efficiency for conjunctions is due to noise and not serial visual attention" *Psychological Science* **9** 111–118
- Eckstein M P, Ahumada A J, Watson A B, 1997 "Visual signal detection in structured backgrounds. II: Effect of contrast gain control, background variations and white noise" *Journal of the Optical Society of America A* **14** 2406–2419

⁽³⁾ Their method measures the neuron's ability to discriminate between a stimulus with the target in the receptive field from one with the target in one of the locations on the side opposite the receptive field. While analysis of the area under the ROC curve is appropriate for a 2AFC task, quantitative comparison between their ROC analysis and the monkey's 8AFC performance is inappropriate. The method we propose, however, would allow a direct quantitative comparison of neural, oculomotor, and perceptual performance in an 8AFC search or even across different M-AFC tasks.

- Eckstein M P, Beutter B R, Stone L, 1998 "The effect of set-size on the relation between saccadic and perceptual decisions during search" *Investigative Ophthalmology & Visual Science* **39**(4) S1031
- Eckstein M P, Whiting J S, 1996 "Visual signal detection in structured backgrounds. I: Effect of number of possible signal locations and signal contrast" *Journal of the Optical Society of America A* **13** 1777–1787
- Findlay J M, 1997 "Saccade target selection during visual search" *Vision Research* **37** 617–631
- Goodale M A, Milner A D, 1992 "Separate visual pathways for perception and action" *Trends in Neurosciences* **15** 20–25
- Gottlieb J P, Kusunoki M, Goldberg M E, 1998 "The representation of visual salience in monkey parietal cortex" *Nature* **391** 481–484
- Green D M, Swets J A, 1966 *Signal Detection Theory and Psychophysics* (New York: John Wiley)
- He P, Kowler E, 1989 "The role of location probability in the programming of saccades: implications for center-of-gravity tendencies" *Vision Research* **29** 1165–1181
- Hess R F, Hayes A, 1994 "The coding of spatial position by the human visual system: effects of spatial scale and retinal eccentricity" *Vision Research* **34** 625–643
- Hooge I, 1996 *Control of Eye Movement in Visual Search* PhD thesis, Utrecht University, Utrecht, The Netherlands
- Hooge I, Erkelens C J, 1999 "Peripheral vision and oculomotor control during visual search" *Vision Research* **39** 1567–1575
- Krauzlis R J, Stone L S, 1999 "Tracking with the mind's eye" *Trends in Neurosciences* **22** 544–550
- Lu Z L, Doshier B A, 1999 "Characterizing human perceptual inefficiencies with equivalent internal noise" *Journal of the Optical Society of America A* **16** 764–778
- Merigan W H, Maunsell J H, 1993 "How parallel are the primate visual pathways?" *Annual Review of Neuroscience* **16** 369–402
- Motter B C, Belky E J, 1998 "The guidance of eye movements during active visual search" *Vision Research* **38** 1805–1815
- Palmer J, 1994 "Set-size effects in visual search: the effect of attention is independent of the stimulus for simple tasks" *Vision Research* **34** 1703–1721
- Pelli D G, 1981 *Effects of Visual Noise* PhD thesis, Cambridge University, Cambridge, England
- 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
- Robson J G, Graham N, 1981 "Probability summation and regional variation in contrast sensitivity across the visual field" *Vision Research* **21** 409–418
- Schall J D, Hanes D P, 1993 "Neural basis of saccade target selection in frontal eye field during search" *Nature* **366** 467–469
- Sialfa C T, Joffe K M, 1998 "Response times and eye movements in feature and conjunction search as a function of target eccentricity" *Perception & Psychophysics* **60** 1067–1082
- Sialfa C T, Kline D W, Lyman B J, 1987 "Age differences in target identification as a function of retinal location and noise level: examination of the useful field of view" *Psychological Aging* **2** 14–19
- Solomon J A, Pelli D G, 1994 "The visual filter mediating letter identification" *Nature* **369** 395–397
- Stone L S, Beutter, B R, Eckstein M P, 1999 "Salience effects on perceptual and saccadic target localization during visual search" *Society for Neuroscience Abstracts* **25** 548
- Tanner W P, 1961 "Physiological implications of psychophysical data" *Annals of the New York Academy of Sciences* **89** 752–765
- Thompson K G, Hanes D P, Bichot N P, Schall J D, 1996 "Perceptual and motor processing stages identified in the activity of macaque frontal eye field neurons during visual search" *Journal of Neurophysiology* **76** 4040–4055
- Tjan B S, Braje W, Legge G E, Kersten D J, 1995 "Human efficiency for recognizing 3-D objects in luminance noise" *Vision Research* **25** 1661–1674
- Tolhurst D J, Movshon J S, Dean A F, 1982 "The statistical reliability of signals in single neurons in cat and monkey visual cortex" *Vision Research* **23** 775–785
- Viviani P, Swensson R G, 1982 "Saccadic eye movements to peripherally discriminated visual targets" *Journal of Experimental Psychology: Human Perception Performance* **16** 459–478
- Wickelgren W A, 1968 "Unidimensional strength theory and component analysis of noise in absolute and comparative judgements" *Journal of Mathematical Psychology* **5** 102–122
- Williams L G, 1966 "The effects of target specification on objects fixated during visual search" *Acta Psychologica* **27** 355–360
- Williams L G, 1967 "Target conspicuity and visual search" *Human Factors* **8** 80–92
- Zelinsky G J, 1996 "Using eye saccades to assess the selectivity of search movements" *Vision Research* **36** 2177–2187

APPENDIX
Signal detection theory and the definition of d'

Models based on signal detection theory (Green and Swets 1966) have been successfully applied to predict human accuracy in perceptual decisions during visual search as a function of signal salience and number of possible signal locations for target localization tasks in white noise (Burgess et al 1981; Burgess and Ghandeharian 1984) and more realistic backgrounds (Eckstein and Whiting 1996). Signal detection theory has also been used for detection of a target among a set of distractors differing from the target along a single (Palmer 1994) or many physical attributes (eg Eckstein 1998). More recently, the model has been applied to predict the accuracy of the first saccade in a search for a bright disk among dimmer disks (Eckstein et al 1998; Beutter et al 2000).

In signal detection theory, the target and nontarget elements in the image are assumed to elicit a noisy response within the observer. Noise will arise owing to variability in the firing of the neurons (eg Tolhurst et al 1982), fluctuations in decision criteria (Wickelgren 1968), and the presence of external noise in the image. A common assumption is that the internal responses of the observer to the target and nontarget elements can be described by independent Gaussian probability distributions with the same standard deviations, but different means. A measure of target/nontarget discriminability is d' , the index of detectability, which is the distance between the target and nontarget probability distributions in standard-deviation units (Green and Swets 1966):

$$d' = \frac{\lambda_t - \lambda_n}{\sigma_i}, \quad (\text{A1})$$

where λ_t is the mean response to the target, λ_n is the mean response to the nontarget, and σ_i is the standard deviation of the internal responses.

Percent correct (P_c) in a task where the target might appear in one of M locations (M -alternative forced choice) is the probability of the target location taking a larger value than the $M - 1$ nontarget locations. With the above assumptions for the internal responses, P_c is given by (Green and Swets 1966):

$$P_c(d', M) = \int_{-\infty}^{+\infty} g(x - d')G(x)^{M-1}dx, \quad (\text{A2})$$

where $g(x)$ is the probability of the target location producing a response x :

$$g(x) = \left(\frac{1}{2\pi}\right)^{1/2} \exp\left(-\frac{x^2}{2}\right) \quad (\text{A3})$$

and $G(x)$ is the cumulative probability of one of the $M - 1$ nontarget locations producing a response less than x :

$$G(x) = \int_{-\infty}^x g(y)dy. \quad (\text{A4})$$

Equations (A2) through (A4) can then be used to generate a lookup table converting P_c to d' for any given value of M .

Ideal observer

An ideal observer optimally uses all the image data, and knowledge of the stimulus to achieve the best possible performance. On each trial, the ideal observer chooses the target location with the highest likelihood ratio (Green and Swets 1966). When the target appears with equal probability in one of the M possible locations, the ideal observer computes a measure of the likelihood for each of the possible signal locations, by computing the correlation between a filter with the spatial profile of the signal,

and the image values at that location. On each trial, the ideal observer chooses the location that results in the highest likelihood (correlation). It can be shown that, for an ideal observer detecting a target added to white noise, d' is related to the displayed target and noise as follows (Burgess and Ghandeharian 1984):

$$d'_{\text{ideal}} = \text{SNR} = \left(\frac{E}{N_0} \right)^{-1/2}, \quad (\text{A5})$$

where E is the signal energy defined as

$$E = \iint \left[\frac{s(x, y)}{I_0} \right]^2 dx dy, \quad (\text{A6})$$

where I_0 is the mean luminance of the background, $s(x, y)$ is the luminance profile of the signal added to the background, and N_0 is the noise spectral density expressed in contrast units (in the case of white noise, this is just the luminance standard deviation of the pixel noise divided by I_0).

Quantifying performance

A common method to quantify how well humans perform a given visual task is to compare their performance with that of the ideal observer by using a metric known as the absolute efficiency. The absolute efficiency (Barlow 1957) of human performance for a given task is defined as $(d'_{\text{human}}/d'_{\text{ideal}})^2$. To compare the amount of information available with the first saccade relative to that available to the final perceptual decision, we introduce the relative efficiency of the saccadic decision versus the perceptual decision, $(d'_{\text{saccade}}/d'_{\text{perception}})^2$. This can be done for either the fixation or eye-movement condition. Similarly, to compare the information available in the final perceptual decision under fixation with that available with eye movements, we introduce the relative efficiency of the fixation condition versus the eye-movement condition, $(d'_{\text{fixation}}/d'_{\text{eye movement}})^2$.

Why compute efficiencies by using d' rather than a simpler measure of performance such as the percentage of correct trials? One argument is that it is desirable to use absolute and relative efficiency metrics that measure “how well” humans process the information in the display in a way which is independent of the task difficulty. If there is a constant source of inefficiency within the observer that degrades performance, then measurements of efficiency should reflect this constant inefficiency, and be independent of specific task parameters such as signal salience, display duration, or the number of possible signal locations. If so, the investigator could use this efficiency metric to draw conclusions about the constancy of observers’ inefficiencies.

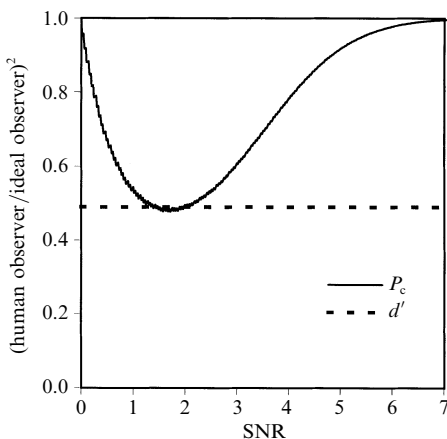


Figure A1. Squared percent correct (P_c) ratio and d' ratio between an observer with a constant source of inefficiency (ie constant additive/proportional noise; sampling inefficiency) and the ideal observer as a function of SNR (manipulated by increasing the signal contrast).

It can be shown that, for an observer who has a constant source of inefficiency, the ratio of $P_{c(\text{human})}$ to $P_{c(\text{ideal})}$ will vary with signal contrast and/or number of locations in the task. Figure A1 shows the effects of a constant source of inefficiency (internal noise) on the P_c ratio. The results show that, if relative efficiency were based on the P_c ratio, it would vary with signal contrast despite the fact that the observers' source of inefficiency is constant. However, the d' ratio (or any function thereof) does remain constant across signal contrast, and thus directly reflects the constant source of inefficiency. Similar arguments rule out the use of P_c to measure relative efficiencies.

