Title
Eye movement optimization in visual search

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 35(35)

ISSN
1069-7977

Authors
Hope, Ryan M.
Gray, Wayne D.

Publication Date
2013

Peer reviewed

Eye movement optimization in visual search

Ryan M. Hope and Wayne D. Gray
Cognitive Science Department
Rensselaer Polytechnic Institute

Abstract

In the present study we investigated whether eye movements in visual search are optimized to reduce time on task. Subjects task was to find a target object in a large field of objects that differed based on shape, color, size and numeric label. The target specification was manipulated, directly influencing the average number of fixations it took subjects to find the target object. Although a microstrategy that allowed for parallel saccade programming and information processing was found to be more efficient in terms of time, a serial microstrategy where saccade programming always follows information processing was found to be the more prevalent microstrategy.

Keywords: visual search, eye movements, microstrategy, optimization, return saccades

Introduction

Visual search may well be our most ubiquitous cognitive task. Many (dozens? hundreds?) times a day we scan our desk for books or memos, our fridge for eggs or beer, the streets for oncoming traffic, cable television for shows we want to watch, and crowded rooms for faces we recognize. Although our natural scan environments are seldom completely novel, the objects in them and their places in these environments are seldom constant.

How do we do such searches? Most of them have the flavor of being a "one off" on at least one of several dimensions. Even the clutter on top of my dresser varies, if only slightly, from day to day. Do we develop optimal search strategies for each environment? – My fridge? Driving on I-87? Looking for friends to sit next to during Cognitive Science talks in Sage 4101? This strikes us as a likely possibility, meaning that there may be no general high level strategy for scanning my dresser top, the fridge, and I-87. But what about lower level strategies? What about the process of moving our eyes to a location, fixating that location, processing the perceptual and semantic information at that location, and deciding whether that location contains the target of our search, or whether we need to saccade to the next location? Can this strategy be optimized? Can it be optimized for all or many search environments? What would such an optimization strategy optimize?

Although in such a small paper it is obvious that we cannot address most of these issues, we believe we have a good start on addressing the later; namely, what would an optimized fixation and saccade strategy look like – this is the subject of our paper.

Background

The key characteristic of visual search for an *active vision* task (Findlay & Gilchrist, 2003), is one of moving the eyes from one location to another until we find our search target. But what does this really entail and can we bring to bear strategies in this task that cannot be applied to the simpler case of being tachistoscopically presented with a single item and asked whether that is our search target or not?

The case of a serial sequence of single items, each of which demands a "yes" or "no" answer before we are shown the next item, seems to define a procedure in which we perceptually process the visual object, semantically process the results of that perception, and decide whether the current object is a member of the target set defined for us by the experimenter. This strategy or procedure seems well suited for the given task environment and maybe defines the optimal strategy in this environment. Indeed, this procedure could be applied when I am searching my dresser top or searching for the large-red-star in Figure 1. Indeed, vision researchers often describe searches that entail the visual scanning of a busy screen in exactly these terms (e.g., Deubel & Schneider, 1996; Henderson, 1992). From this point forward we will refer to this strategy as the *serial microstrategy*.

An alternative strategy differs from the serial microstrategy only in that the programming and subsequent execution of the next eye movement does not wait for the semantic processing and decision processing to complete. Programming of the next eye movement could start as soon as the previous saccade is completed, in parallel with information processing and decision making processes. From this point forward we will refer to this strategy as the *parallel microstrategy*.

Evidence for the parallel strategy comes from a number of visual search studies (Engel, 1977; Gould, 1973; Hooge & Erkelens, 1996) where it was reported that subjects often fixated the target, made an eye movement away from the target, and then, on the next fixation, returned to the target (return saccades). Hooge and Erkelens (1996) also reported a number of missed targets. This evidence implies that the fixation durations can be too short to recognize the target. The occurrence of return saccades and missed targets suggest that saccade preparation may start before foveal processing is complete and that complete foveal processing is not necessarily the trigger for the subsequent saccade. For further details of both serial and parallel models of eye movements in visual search see Hornof and Halverson (2003), Hornof and Kieras (1997).

The parallel and serial microstrategies can also be differentiated based on their temporal costs. Since the parallel strategy allows for the concurrent processing of information and saccade preparation, fixation durations could be shorter when the parallel strategy is used than when the serial strategy is used because target identification can continue during the saccade. However, Becker and Jürgens (1979) showed that it is possible for saccades to be aborted during the preparation phase. This could lead to someone using the parallel strategy with apparent fixation durations closer to those observed in the serial strategy. On the other hand, since saccades in the parallel strategy do not necessarily wait for a decision of target presence, the parallel strategy could involve two extra saccades and fixations as a result of return saccades. Therefore in the extreme cases, the parallel strategy has a fixed cost of 2 saccades and 2 fixations whereas the serial strategy has a cost that grows with each fixation. Therefore there exists some threshold, in terms of number of fixations, where the parallel strategy will eventually become more efficient than the serial strategy.

The task used in the present study is a visual search task first used by Williams (1966). In the Williams search task, subjects have to find a target object in a very large field of objects that differ by size, shape, color and a numbered label (e.g., “11”, “25”, etc). Williams found that when he manipulated which target features were known (e.g., “large blue circle” versus “small yellow” versus “triangle”), a high proportion of fixations were on objects of the specified color and only a moderate proportion of fixations were on objects of the specified size or shape. When two or more characteristics were specified, fixations were generally based on a single characteristic. Additionally, the average number of fixations required to find the target differed based on the target specifications. This aspect of the task makes it perfect for eliciting the use of different microstrategies as the optimal strategy should depend on target specifications.

It seems likely that at least some subjects on at least some trials used a mix of serial and parallel microstrategies. However, the parallel and the serial strategies predict the same pattern of saccades for all but the last two saccades to and from target object. Hence, for this initial report, we make the simplifying assumption that if we find a return saccade, that we can classify the entire trial as having been accomplished using the parallel strategy. Likewise, the absence of return saccades were used to classify trials as having been accomplished using a serial strategy. Based on this classification scheme we had the following hypotheses:

1. Due to the temporal costs of the serial microstrategy with increasing number of fixations, the proportion of trials with return saccades (indicating the use of the parallel microstrategy) should be higher on trials where search is inefficient than it is on trials where search is efficient. Here efficiency corresponds to the number of fixations need to find a target, a function of the target specification or number of cue features.

2. The average fixation duration (all fixations on a given trial) should be shorter on trials that exhibit return saccades (indicating the use of the parallel microstrategy) than on trials that do not exhibit return saccades.

3. For trials that do not contain return saccades, fixation durations should increase as a function of number of cue features due to increasing processing requirements.

4. Subjects will satisfice by using the microstrategy that results in the most time savings across all trial types.

Method

Subjects

Subjects were 15 undergraduate students at Rensselaer Polytechnic Institute (10 men, 5 women) who were given course credit for their participation. Subjects were prescreened for their dependence on eyeglasses or contact lenses; only subjects that reported needing neither were allowed to participate in the experiment.

Apparatus

The experiment was displayed on a 22" Dell widescreen LCD with a resolution of 1680 × 1050 (pixels) and physical dimensions of 473.76 × 296.1 (mm). Eyetracking was performed with a SensoMotoric Instruments RED500 eyetracker running at a sample rate of 500 Hz. On average, subjects were positioned 700 ± 100 mm away from the LCD.

Task

Subjects task on each trial was to find a target in a field of 48 randomly dispersed objects. Each object had a unique combination of shape (4 levels), color (4 levels) and size (3 levels). On each shape was a randomly assigned numeric id which ranged from 01 to 48. Each trial starts with the search objects masked and a cue at the center of the display. The cue was a text description of the target object. The probe always contained the numeric id and up to three other features (shape, size and color). The particular features shown in the cue, in addition to the numeric id, was systematically manipulated throughout the experiment such that each subject experienced one trial of each object and cue combination. The non-id cue features were ordered randomly with the id always showing last (see Figure 1 for an example of the cue). The 48 objects and 8 cue combinations results in 384 unique trials. Because of the random dispersion of objects on each trial, no subjects experienced identical trials. Subjects were instructed to study the cue until they felt they had memorized it, at which point, they were instructed to press spacebar on the keyboard to reveal the search field and begin searching for the target. The cue remained on screen during the search phase. Once subjects find the target, they end the trial by using the mouse to click on the target. Search time was measured as the time between the spacebar press and the first correct click on the target object. Subjects were given no explicit instruction to emphasize speed, their task was simply to find the target.

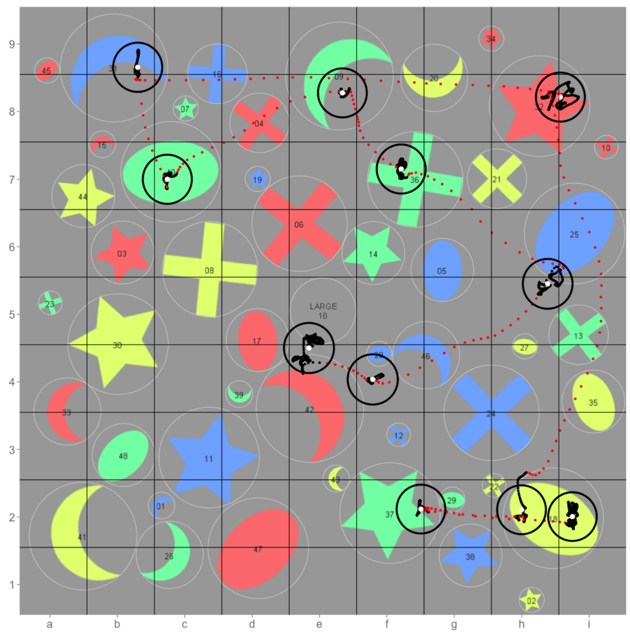

Figure 1. An example of a trial search display that contains an end of trial return saccade. In this trial, the cue is "large 18". The target is the large yellow oval in the bottom right corner. Eye gaze data is overlaid; red dots correspond to samples that belong to saccades, black dots correspond to samples that belong to fixations. Circles with a 1 degree visual angle radius have been drawn around the center of mass for each fixation. As indicated by the grid formed by the x and y axes, the scanpath for this Ss is: e4, f4, h5, h5, f7, e8, c7, b9, i8, h2, f2, i2. Notice how the last 3 fixations include 2 fixations on the target object separated in time by a fixation on a different object. Neither the grid nor the light gray circles around each object were visible during the trial.

Stimuli

The four shapes used in the task were star, oval, crescent and cross. The four colors used in the task were red, yellow, green and blue and had hue values (in HSV space) of 0, 72, 144 and 216 respectively. The saturation and value of all four colors was set to 50 and 100 respectively. The search field was a 1050 ×1050 (pixel) square centered on the screen and had 65% gray background color. See Figure 1 for an example of the shapes and colors used in this task. The three object sizes were small (48 pixels), medium (119 pixels), and large (191 pixels) and correspond to visual angles of 1.1, 2.7 and 4.4 degrees with a potential error up to 15% depending on head position.

Gaze Data Classification

Raw gaze data were classified into events (saccades and fixations) by an algorithm that uses both velocity and acceleration thresholds. The algorithm followed the following general procedure:

1. Convert the x,y screen coordinates from pixels to degrees of visual angle relative to the center of the screen.
2. Compute smoothed first and second order derivatives of the x,y visual angle components.
3. Label gaze samples with corresponding velocity and acceleration components that both exceed their respective thresholds as saccades, the rest are labeled as fixations.
4. Identify saccades that last for less than 20 ms, reclassify them as fixations.

Velocity and acceleration (first and second order derivites) were calculated using a Savitzky-Golay filter (as recommended by Nyström and Holmqvist (2010)) for its ability to preserve local minima and maxima. The Savitzky-Golay filter used was a second order filter with a window length of 11 samples which allowed for accurate detection of saccades down to 20ms in duration (Nyström & Holmqvist, 2010). The velocity and acceleration thresholds used in the present study were 30°/s and $8000°/s^2$ respectively and were based on the "cognitive configuration" of the EyeLink software (SR Research Ltd. 2007).

Results

Number of Fixations

Our first hypothesis stated that the proportion of return saccades should increase with number of fixations and that number of fixations was directly related to the cue specification (trial type) and number of cues. Therefore, in order to evaluate our first hypothesis we first need to show that trial type has an impact on number of fixations. In order to accomplish this we performed one-way repeated measures analysis of variance. The analysis of variance revealed a significant effect of trial type on proportion of return saccades, $F(7,98) = 76.56$, $p < .001$, ${\eta_g}^2 = 0.73$. The means and standard errors can be seen in Figure 2. The general trend revealed from this analysis is that the number of fixations decreased as the number of cue features increased. We performed a second analysis of variance to test the effect of number of cues on number of fixations. As expected by hypothesis 3, the greater the number of cues, the fewer fixations, $F(3,42) = 96.62$, $p < .001$, ${\eta_g}^2 = 0.76$. The effect size of number of cue was stronger than the effect of trial type on number of fixations.

Return Saccades

Return saccades were identified through scanpath analysis. First, the search field was divided into a 9×9 array which resulted in 81 total cells. The width and height of each cell was approximately 2.6 degrees of visual angle. Second, fixations were recoded as belonging to one of the 81 cells based on their center of mass. After recoding, consecutive fixations occurring in the same cell were combined. This effectively removes any microsaccades and other small amplitude corrective saccades. Finally, for scanpaths of length 3 or greater, the last fixation was compared to the fixation 2 back. In order for a trial to be classified as containing a return saccade two

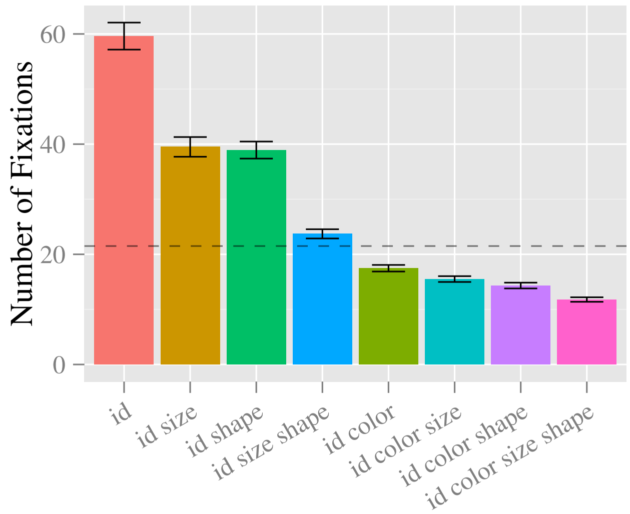

Figure 2. The average number of fixation for each trial type. Error bars represent standard error. The dashed line represents the threshold where the parallel strategy becomes more efficient than the serial strategy, determined empirically using Equation 5.

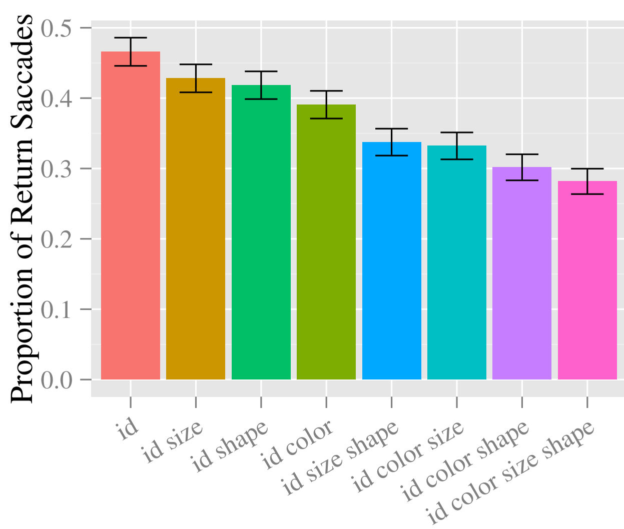

Figure 3. The average proportion of return saccades for each trial type. Error bars represent standard error.

criterion had to be met. First, the fixation 2 back from the final fixation had to be in the same cell, or any of the surrounding cells, as the last fixation. Second, the fixation 1 back from the last fixation could not be on the center cell where the cue was located (see center of Figure 1). Of the 5939 total trials, 2273 trials contained a return saccade.

In order to test the effect of trial type on the proportion of return saccades a one-way repeated measures analysis of variance was performed. The analysis of variance revealed a significant effect of trial type on proportion of return saccades, $F(7,98) = 12.62$, $p < .001$, ${\eta_g}^2 = 0.33$. The means and standard errors can be seen in Figure 3. The general trend revealed from this analysis is that the proportion of return saccades decreased as the number of cue features increased. In order to confirm this trend we performed a second one-way repeated measures analysis of variance to test the effect of number of cue features on the proportion of return saccades. The analysis of variance revealed a significant effect of number of cue features on proportion of return saccades, $F(3,42) = 25.57$, $p < .001$, ${\eta_g}^2 = 0.46$. Similar to the analysis of variance involving number of fixations, the effect size of number of cues was stronger than the effect of trial type for proportion of return saccades. In addition, the correlation between proportion of return saccades and number of fixations with respect to trial type, which can be seen by comparing Figure 2 and Figure 3, is strong; $r = .90$, $n = 8$, $p = .002$.

Fixation Durations

In order to test the effect of return saccades and number of cue features on average fixation duration (all fixations within a trial) we performed a 2×4 (return saccade by number of cue features) repeated measures analysis of variance. The analysis of variance revealed a significant main effect of return saccade, $F(1,14) = 27.44$, $p = .001$, ${\eta_g}^2 = 0.22$; a significant main effect of number of cues, $F(3,42) = 6.68$, $p < .001$, ${\eta_g}^2 = 0.11$; and a marginally significant interaction effect $F(3,42) = 2.62$, $p = .063$, ${\eta_g}^2 = 0.02$. The means and standard errors can be seen in Figure 4.

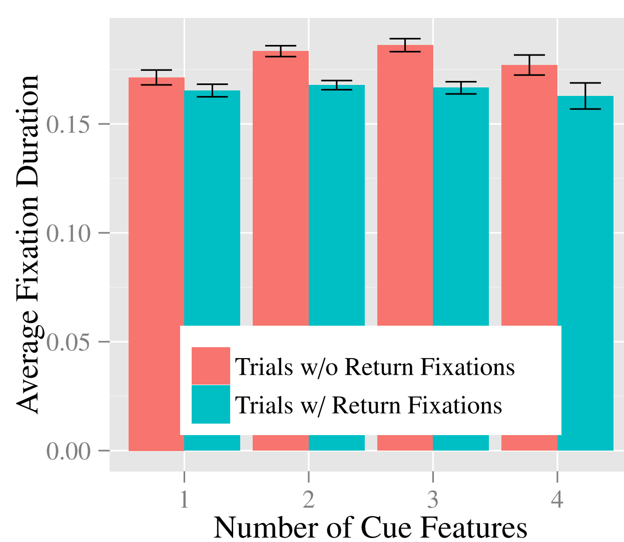

Figure 4. The average fixation duration (all fixations within a trial) for each level of number of cue features. Error bars represent standard error.

Comparison of Search Strategies

The serial strategy search time T_{serial} can be approximated by the following equation:

$$T_{serial} = (N - 2) * (F_{serial} + S) \tag{1}$$

where N is the average number of fixations on a given trial type, and S is the average saccade duration. Similarly, the parallel strategy search time $T_{parallel}$ can be approximated by the following formula:

$$T_{parallel} = N * (F_{parallel} + S) \tag{2}$$

In order to approximate the values for F_{serial} and $F_{parallel}$, we first performed mixed effects regression on mean fixation duration with return saccade and number of cues as fixed factors and with subjects as a random factor. The regression yielded the following equation (rounded to the nearest millisecond):

$$F = 213 - E * 24 + C * 5 \tag{3}$$

where F is fixation duration, E is 1 for return saccade trials (0 for non-return saccade trials) and C is the number of cues. The intercept value of 213 ms is consistent with previous research on average fixation durations in visual search (Rayner, 2009; Salthouse & Ellis, 1980). The 24 ms difference between trials with and with out return saccades (assumed to be associated with the time required to make a decision of target presence) is psychologically plausible (Neisser, 1963; van Diepen, De Graef, & D'Ydewalle, 1995). In addition, the 5 ms per cue feature seems psychologically plausible. Using Equation 3 we can compute the average fixation duration, F_{serial} and $F_{parallel}$, for a given trial type and search strategy by using the average number of fixations, N, as shown in Figure 2. The value of S was set to 45 ms in all computations based on the average of saccade durations in the empirical data.

The time savings (or loss) from using the parallel search strategy can then be computed as follows:

$$T_{diff} = T_{serial} - T_{parallel} \tag{4}$$

The results of Equation 1 and Equation 2 applied to each of the 8 probe combinations applied to Equation 4 is shown in Figure 5. The parallel search strategy saved more time than the serial search strategy for all trials where color was not an available cue feature. The sum of all T_{diff} values was 972.02 ms indicating that overall, the parallel strategy could be more efficient.

This analysis can be taken one step further by setting F_{serial} equal to $F_{parallel}$ and solve for N, resulting in Equation 5, to find the threshold in terms of number of fixations where the parallel strategy becomes more efficient than the serial strategy.

$$N = \frac{2 * F_{serial} + 2 * S}{F_{serial} + F_{parallel}} \tag{5}$$

By using a value of 213 for F_{serial} and 189 for $F_{parallel}$ (computed from Equation 3) and a value of 45 for S, the threshold turns out to be 21.5 fixations. This threshold is depicted as the dashed horizontal line in Figure 2. Interestingly, for all trials in which color is available as a probe feature, the average number of fixations is less than this threshold. In other words, when color is not available, the parallel search strategy will be more efficient in terms of time.

Discussion

The goal of this study was to determine if people optimize their eye movements during visual search. In order to accomplish this goal we had subjects perform a difficult visual search task where they had to find a target object in a field of objects that differed in size, shape color and numeric label. We hypothesized that subjects would optimized their eye movements by using the more efficient of two microstrategies: the serial and parallel microstrategies. We used the presence of return saccades as a marker of the parallel microstrategy. Consistent with our first hypothesis, the proportion of trials that contained return saccades were higher on trials that required lots of fixations to find the target as well as on trials where there were few target cues compared to trials that required only a few fixations to find the target or trials that had many target cues. One interpretation of this result is that subjects are indeed sensitive, consciously or unconsciously, to the temporal costs of serial strategy which ensures a target presence decision before the eye is moved to the next location.

Our second hypothesis predicted that fixation durations would be shorter on trials that contained return saccades that on trials that did not contain return saccades. This indeed was the case. We also predicted in our third hypothesis that on trials that did not include a return saccade that fixation durations should increase with respect to the number of known target features. This prediction was mildly supported by the marginally significant interaction we found in our analysis of variance involving return saccades and number of cue features on fixation duration.

Our conservative cost analysis of the serial and parallel microstrategies does show that the parallel microstrategy can more cost effective in terms of time on task. However, return saccades were only observed on 38% of the trials in our experiment. This could be interpreted as evidence that subjects did not or could not optimize their eye movements. However, we are not positive that is the case. It's possible that return saccades are just not a good enough measure of the parallel search strategy since it is possible for early eye movement programming (that would have cause the eye to move before target analysis) to be aborted in the parallel strategy. This could not only potentially reduce the number of observed return saccades but it could also affect the average fixation durations in trials with and with out return saccades. It is possible that someone could be using the parallel strategy throughout most of a trial but then abort an eye movement that would

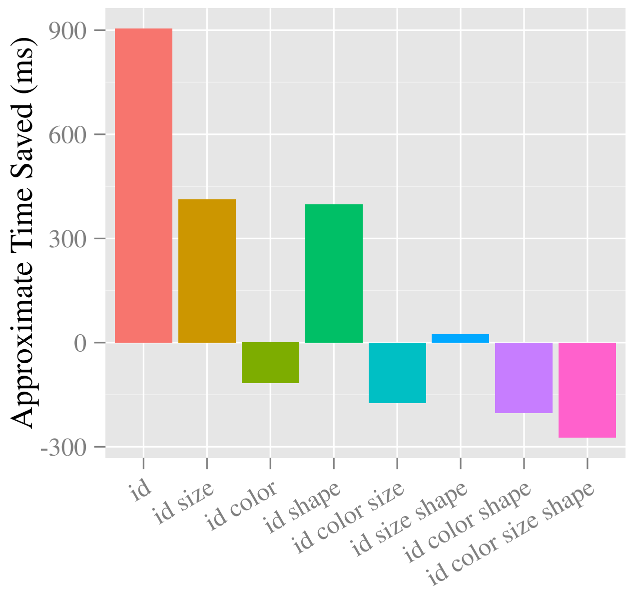

Figure 5. The estimated difference (T_{diff}) in time from using the parallel or serial microstrategy, computed using Equation 4. Positive values favor the parallel strategy, negative values favor the serial strategy.

have resulted in a return saccade. This would result in the lowering of the average fixation duration for trials with out a return saccade due to improper strategy classification. Additionally, a trial that shows a return saccade could still have contained many aborted early saccades inflating the average fixation duration of return saccade trials.

Conclusion

Our study has provided evidence that eye movements in visual search can be sensitive to millisecond level cost-benefit trade-offs. Whether or not people can actually optimize their eye movements to take advantage of these cost-benefit trade-offs is still not clear. In addition, the fact that the parallel strategy could under some circumstances appear as if it were actually a serial strategy allows for the possibility that the serial strategy does not even exist. This idea is consistent with the findings of Hornof and Kieras (1997) and Hornof and Halverson (2003). Further research on this topic will need to find better ways to quantify the prevalence of parallel processing in eye movement microstrategies.

Acknowledgments

The work was supported, in part, by grants N000140910402 and N000141310252 to Wayne Gray from the Office of Naval Research, Dr. Ray Perez, Project Officer.

References

Becker, W. & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision research*.

Deubel, H. & Schneider, W. X. (1996). Saccade target selection and object recognition: evidence for a common attentional mechanism. *Vision research*, *36*(12), 1827–37.

Engel, F. (1977). Visual conspicuity, visual search and fixation tendencies of the eye. *Vision research*, (June).

Findlay, J. M. & Gilchrist, I. D. (2003). *Active vision*. Oxford Univ. Press.

Gould, J. D. (1973). Eye movements during visual search and memory search. *Journal of experimental psychology*, *98*(1), 184–95.

Henderson, J. M. (1992). Visual attention and eye movement control during reading and picture viewing. In K. Rayner (Ed.), *Eye movements and visual cognition: scene perception and reading* (pp. 260–283). New York: Springer.

Hooge, I. T. C. & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Attention, Perception, & Psychophysics*, *58*(7), 969–976.

Hornof, A. J. & Halverson, T. (2003). Cognitive strategies and eye movements for searching hierarchical computer displays. In *Proceedings of ACM CHI 2003: Conference on Human Factors in Computing Systems*. New York, NY: ACM.

Hornof, A. J. & Kieras, D. E. (1997). Cognitive modeling reveals menu search is both random and systematic. In *Acm chi'97 conference on human factors in computing systems* (pp. 107–114). ACM.

Neisser, U. (1963). Decision-Time without Reaction-Time: Experiments in Visual Scanning. *The American Journal of Psychology*, *76*(3), 376–385.

Nyström, M. & Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. *Behavior Research Methods*.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Salthouse, T. & Ellis, C. (1980). Determinants of eye-fixation duration. *The American journal of psychology*.

SR Research Ltd. (2007). EyeLink 1000 User Manual (9/13/2007).

van Diepen, P. M., De Graef, P., & D'Ydewalle, G. (1995). Chronometry of foveal and peripheral information encoding during scene perception. In J. M. Findlay, R. Walker, & R. W. Kentridge (Eds.), *Eye movement research: mechanisms, processes and applications* (pp. 349–362). Amsterdam: North Holland.

Williams, L. G. (1966). The effect of target specification on objects fixated during visual search. *Perception and Psychophysics*, *1*(5), 315–318.