Journal of Experimental Psychology: Human Perception and Performance

Eye Movements Reveal how Task Difficulty Moulds Visual Search

Angela H. Young and Johan Hulleman

Online First Publication, May 28, 2012. doi: 10.1037/a0028679

CITATION

Young, A. H., & Hulleman, J. (2012, May 28). Eye Movements Reveal how Task Difficulty Moulds Visual Search. *Journal of Experimental Psychology: Human Perception and Performance*. Advance online publication. doi: 10.1037/a0028679

Journal of Experimental Psychology:
Human Perception and Performance

© 2012 American Psychological Association
0096-1523/12/$12.00 DOI: 10.1037/a0028679

Eye Movements Reveal how Task Difficulty Moulds Visual Search

Angela H. Young
University of Hull

Johan Hulleman
University of Hull and University of Manchester

In two experiments we investigated the relationship between eye movements and performance in visual search tasks of varying difficulty. Experiment 1 provided evidence that a single process is used for search among static and moving items. Moreover, we estimated the functional visual field (FVF) from the gaze coordinates and found that its size during visual search shrinks with increasing task difficulty. In Experiment 2, we used a gaze-contingent window and confirmed the validity of the size estimates. The experiment also revealed that breakdown in robustness against item motion is related to item-by-item search, rather than search difficulty per se. We argue that visual search is an eye-movement-based process that works on a continuum, from almost parallel (where many items can be processed within a fixation) to completely serial (where only one item can be processed within a fixation).

Keywords: visual search, moving items, eye movements, functional visual field (FVF), inhibitory tagging

Most theories of visual search (e.g., Heinke & Humphreys, 2003; Itti & Koch, 2000; Wolfe, 1994; Zelinsky, 2008) would predict search slopes to become steeper and error rates to increase when items move around in the search display. The underlying reason is that it should become more difficult to avoid reinspections of items when they are moving around, yielding longer RTs and more missed targets. However, when Horowitz and Wolfe (1998) tested this prediction, they found that search slopes on present trials were the same irrespective of whether items were static or randomly relocated every 111 ms. But there were some problematic aspects to the data, since Horowitz and Wolfe (1998) also found that target-present and target-absent slopes were identical in the dynamic condition and that there were more errors and longer RTs overall for the dynamic condition relative to the static condition. Because of this, von Mühlenen, Müller, and Müller (2003) argued that participants may have used a sit-and-wait strategy; looking at one section of the array and waiting a fixed period before responding "present" if the item had moved into the section by that time or "absent" if not. To test this hypothesis, von Mühlenen et al. (2003) forced participants to adopt a sit-and-wait strategy by only revealing one part of the search display. They found that performance for this condition was very similar to a dynamic control condition. Additionally, Geyer, von Mühlenen, and Müller (2007) replicated Horowitz and Wolfe's experiment with the addition of eye movement recording and found that in the moving item condition fewer fixations were made, fixation durations were longer, and saccade amplitudes were smaller than in the static condition. This pattern of eye movements supports the argument that, in this case, a sit-and-wait strategy was used. Horowitz and Wolfe's study, therefore, was rejected as evidence that visual search is robust against motion.

Recently, however, Hulleman (2009, 2010) reported evidence of robustness against item motion in visual search that could not be explained by a sit-and-wait strategy. Using static items and items that moved smoothly with speeds of up to 10.8°/s in search for T among Ls, Hulleman (2009) found that search slopes for static and moving items were almost the same. Contrary to what would have been expected had participants used a sit-and-wait strategy, in Hulleman's experiments search performance for static and moving item conditions was virtually identical in every aspect of the search process. Not only the present slopes, but also the absent slopes, the overall RTs for both target present and target absent trials and even the error rates. So, the observed robustness against motion cannot be explained by a sit-and-wait strategy. Robustness against motion only broke down for very difficult, deeply serial searches, with search slopes of around 100/210 ms per item for target-present and -absent trials, respectively (Hulleman, 2010). These search slopes are consistent with effortful item-by-item search, with each item requiring a fixation and saccades between items. It should be noted that performance in the static version of this difficult search task was very poor as well.

Robustness against item motion provides a challenge to any theory of visual search that assumes that the whole display is processed at once (either serially or in parallel), since all items need to be tracked for the entire duration of the trial. Robustness against motion comes much more naturally to models that define visual search in terms of eye movements and fixations. Estimates of fixation durations are in the region of 200–250 ms (Gilchrist & Harvey, 2000) and fairly stable during trials (Over, Hooge, Vlaskamp, & Erkelens, 2007) and across trial difficulty (Hooge & Erkelens, 1996). The relatively short duration of a fixation means that the items, at the speeds used in Hulleman (2009, 2010), will

Angela H. Young, Department of Psychology, University of Hull, Hull, United Kingdom; Johan Hulleman, Department of Psychology, University of Hull and School of Psychological Sciences, University of Manchester, Manchester, United Kingdom.

Angela H. Young is now at the School of Psychology, University of Nottingham.

This research was supported by a grant from The Leverhulme Trust (F/00 181/T) to Johan Hulleman.

Correspondence concerning this article should be addressed to Johan Hulleman, School of Psychological Sciences, University of Manchester, Oxford Road, Manchester M13 9PL, United Kingdom. E-mail: johan.hulleman@manchester.ac.uk

not have moved very far. Moreover, given the physiology of the retina, the number of items that has to be processed during a fixation will be reduced as well (cf. Geisler & Chou, 1995; Motter & Simoni, 2008). This kind of divide-and-conquer approach would become even more robust against item motion if items would be processed in parallel during a fixation, because items that move from one fixated area to another during a trial would be less detrimental to search performance. The case for parallel processing during a fixation has been made by many authors (e.g., Eckstein, Thomas, Palmer, & Shimozaki, 2000; Palmer, Verghese, & Pavel, 2000) and although other authors have proposed a rapid succession of serial deployments of attention (Wolfe, 2003), it has recently been argued that there is experimental evidence in favor of parallel processing during a fixation (Motter & Holsapple, 2007). Moreover, an extensive modeling study by Thornton and Gilden (2007) also concluded that most searches (including T among L) are conducted by a parallel limited-capacity process. In fact, the very robustness of search for a T among Ls against item motion (with up to 36 items in a display, Hulleman, 2010) might itself be considered another argument for parallel processing.

Although sequences of relatively short fixations and parallel processing within them explain why easier search is robust against motion, they do not explain why this robustness breaks down for very difficult search. However, even with parallel processing within a fixation, it should be noted that revisits of previously fixated areas will be detrimental, since every refixation adds 200–250 ms to a search. So, avoidance of areas of the display that have already been fixated is necessary for successful visual search. Gilchrist and Harvey (2000) and Peterson, Kramer, Wang, Irwin, and McCarley (2001) have reported that the number of refixations of items is indeed smaller than expected from chance. If memory for eye movements (either prospective or retrospective, Peterson et al., 2001) is restricted, then this would provide the limit to robustness against motion for difficult search. When every item is fixated individually, the number of fixations necessary to complete the task successfully will exceed memory capacity and revisits can no longer be avoided. One final ingredient is therefore needed to provide a complete theoretical framework: the functional visual field (FVF) should depend on the difficulty of the search task. It has already been pointed out that the physiology of the retina limits the number of items that can be processed in a single fixation. However, it has been shown that the size of the FVF (the number of items processed during a fixation) does not only depend on acuity, but also on the difficulty of the task (Motter & Simoni, 2008; Geisler & Chou, 1995). The FVF shrinks when the task becomes more difficult. For instance, Geisler and Chou (1995) reported that the FVF was smaller for a conjunction of color and orientation than for an orientation feature. In this context it is interesting to note that both Gilchrist and Harvey (2000) and Peterson et al. (2001) used very difficult search tasks, which might explain why they reported their avoidance of previous inspections in terms of items rather than of fixated areas.

Under this account, easier search is robust against motion because the FVF is so large that the number of fixations needed to process the entire search display is small enough to be able to avoid refixation of areas. Difficult search is not robust against motion because the FVF shrinks to such an extent that only one or two items are processed per fixation. As a consequence, the number of fixations needed to process the display is so large that refixations can no longer be avoided. Moreover, because only one or two items are processed per fixation, parallel processing no longer affords any protection and every refixated item increases the time needed to find the target.

In conclusion, a theoretical framework with four elements would provide a description of all aspects of the results of Hulleman (2009, 2010):

1) A fixed and short fixation duration

2) Parallel processing of items within a fixation

3) Limited avoidance of previously fixated areas

4) FVF depends on the difficulty of the search task.

One clear implication of this theoretical framework is that the eye movement characteristics in visual search should be driven by the difficulty of the search task. They should not depend on whether the display contains static or moving items, since there is only a single search process that is applied to both kinds of displays. This lack of influence of item motion should be clearest in target-present trials. As we have pointed out elsewhere (Hulleman, 2010), in target-absent trials we might expect some effect of item motion due to a shift in stopping criterion. Participants could feel compelled to postpone the strategic target-absent decision (Chun & Wolfe, 1996) when they notice that the items move around.

Because Hulleman (2009, 2010) did not record eye movements, it still remains to be established that the eye movement patterns are identical for moving and static items, even though there were no differences in RTs, search slopes, and error rates in the easier search conditions. In Experiment 1, we therefore recorded behavioral and eye movement data to establish that there is indeed no distinct eye movement pattern associated with search among moving items (e.g., a sit-and-wait strategy like that seen in Geyer et al., 2007). We then used the rich data from gaze coordinates to determine whether, as assumed by the framework, the FVF does indeed depend on the difficulty of the search task. We not only estimated the size of the FVF for the three different search difficulties, but also analyzed the drift during a fixation (the distance the gaze has shifted between two saccades) and the distance between saccadic endpoints and the nearest item, which both should vary with the size of the FVF. Finally, in Experiment 2 we used a gaze-contingent window paradigm to test the accuracy of the estimates of the FVF, as well as the implication of the framework that the breakdown in robustness against item motion is due to search being conducted item by item, rather than the difficulty of the search task per se.

Experiment 1

In Experiment 1 we used the medium and difficult search conditions from Hulleman (2010) in order to investigate the eye movements used for visual search tasks of varying difficulty, using static or moving items. We also added a feature search condition where participants searched for an oblique target among vertical distractors. Our framework predicts that the same search processes are active in both search among static and search among moving items. If this is indeed the case, then we should expect no clear difference in eye movements between static and moving item search, in terms of number of fixations, fixation durations, or saccade amplitude. One specific way in which it has been reported that search among moving items may differ from search among static items is if participants use a sit-and-wait strategy (e.g., Geyer et al., 2007). If participants focus on just one part of the screen throughout a trial, we should see an increase in fixation duration and a decrease in the number of fixations. Alternatively, there could be a reduction in the saccadic amplitude, so that the same number of fixations would cover a smaller area of the display. A sit-and-wait strategy, therefore, can be diagnosed by either a combination of longer fixation durations and fewer fixations or by shorter saccade amplitudes in the moving item condition.

Following Hulleman (2010), we expected an effect of item speed on manual performance only for difficult search, with an increase in errors in the moving item condition. According to our theoretical framework, there will be more fixations for more difficult search (due to a reduction in the number of items that can be processed in a single fixation), but little change in fixation duration or saccade amplitude. Since our framework suggests that parallel processing within a fixation provides robustness against reinspections of items, we would expect to find that individual items will be revisited in difficult search. There will be more revisits for larger display sizes, but even for small display sizes there will probably be some revisits.

We used gaze coordinates and item coordinates to estimate the FVF of the visual search process for each of the three search difficulties. According to our theoretical framework, processing of items is parallel within fixations and serial between fixations. So, during each fixation a part of the display is sampled, and the items in the sample are processed. Because of the random location of the target, we would expect that, on average, half of the items have to be processed before the target is found. Consequently, the size of the FVF can be determined by counting the number of items within a certain radius from the location of fixation and adding the counts for all fixations (each item is counted only once, so if it falls inside the radius of a further fixation it does not contribute to the overall count). The size of the FVF is then defined as that radius around the location of fixation where the sum of the counts across all fixations equals half the number of items in the display. Since fixation durations are typically in the 200–250 ms range and the position of moving items changes over this period, we used the end point of a saccade as a proxy for the location of fixation, because the end of a saccade is a discrete point in time.

There should be two further indicators for a change in size of the FVF. The first is provided by drift: the distance the gaze position has drifted within a single fixation. We estimated this distance by taking the difference in gaze position between the end of the previous saccade and the start of the current saccade. Drift is an indicator for the size of the FVF for the following reason. The direction of motion of each item can be represented by a vector. When all the individual motion vectors are averaged, a net motion vector results (e.g., if one item moves to the left and another to the right, the net motion vector would be 0, because the individual motion vectors cancel each other out). It is important that the more items contribute to the net motion vector, the more likely they are to cancel each other out, since each item has a randomly assigned motion direction. So, the size of drift should increase with decreasing FVF when the search items move. When they are static, we do not expect an effect of difficulty.

The second indicator for a size change in FVF should be the distance between the endpoint of a saccade and the nearest item. For large FVFs, it is not be necessary to fixate very close to any item in particular. One of the implications of parallel processing is that no single item is favored over others. However, when the FVF becomes so small that it only contains a single item, it follows that this item will be targeted by the saccade, and hence its distance to the saccade endpoint will be small.

Method

Participants. Fourteen University of Hull students (all female, one left-handed, 18–30 years, average age = 21 years) participated in this experiment. All participants received course credit for their participation and were naive to the purpose of the experiment.

Stimuli. Items were white on a black background and presented within a virtual rectangle (29.0° × 19.3°). Minimum distance between items was 1.45°. All items in a display moved with identical velocity. Depending on the condition velocity was 0.0° or 7.2°/s along a linear trajectory in a randomly chosen direction. Motion sequences consisted of 1,100 frames. Each frame was presented for 13.3 ms, yielding a maximum display duration of 14.6 seconds. In every frame, all items were shifted the appropriate number of pixels (0 and 2 for 0.0 and 7.2°/s, respectively). Whenever items reached minimum distance with another item or reached the edge of the virtual rectangle, they bounced, and their trajectory changed instantaneously according to an elastic collision. Examples of the easy, medium, and difficult displays are shown in Figure 1. Items in the easy condition consisted of vertical lines and a diagonal line target (0.05° × 0.96°). In the medium condition the target was a T among Ls (0.96° × 0.96°, with four possible orientations: upright or rotated −90°, 90°, or 180°). In the difficult condition, the target was a square (0.96° × 0.96°) with a smaller square (0.48° × 0.48°) in the top left corner among squares with a smaller square in one of the other three corners.

Procedure and design. Stimuli were presented and responses recorded using software custom written in C. The search displays were presented on a 19-in. Monitor (Iiyama Vision Master Pro 454; 800 × 600 pixels, 75 Hz) controlled by a GeForce 6800 graphics card. After a 1,000-ms blank display, a 0.5° × 0.5° fixation cross was presented for 500 ms in the center of the display. After offset of the fixation cross, the search display was presented. Displays contained 6, 12, or 18 items and item speed was either static (0.0°/s) or moving (7.2°/s). The participant's task was to search for a target item, which was present in 50% of the trials, pressing the right or left trigger key of a Sidewinder Gamepad to indicate whether the target was present or absent. Participants used their preferred hand for present responses. The four factors (difficulty, display size, item speed, and target) were fully crossed, yielding 3 × 3 × 2 × 2 = 36 cells for the analysis. Every cell contained 25 trials, giving 900 trials in total.

The trials were blocked by the factor difficulty, which was counterbalanced across participants. At the beginning of each part of the experiment, participants were given appropriate instructions for the next task and completed at least 10 practice trials. For the easy and medium difficulty tasks participants were presented with

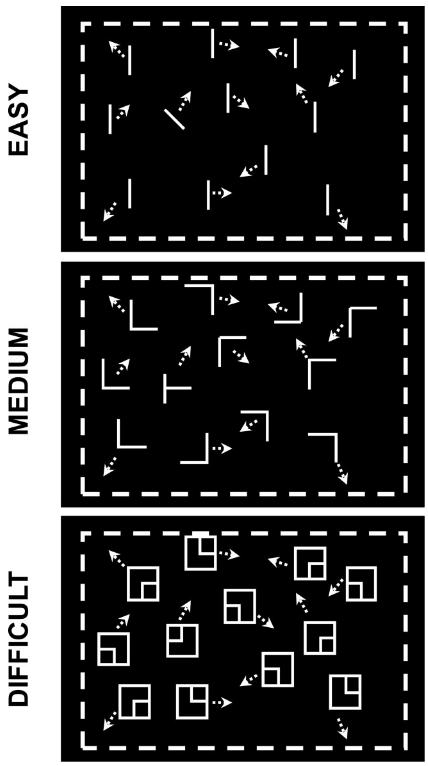

Figure 1. Illustration of the displays used in Experiment 1. In the easy, medium, and difficult search conditions, participants had to search for a diagonal line, a T, or a square with a smaller square in the top left corner, respectively. All items in a display moved with the same velocity in randomly chosen directions. Arrows and dashed line are for illustration of movement and virtual rectangle and were not visible during the actual trials.

six blocks of 50 trials, with an opportunity to rest between blocks. For the hard trials, 12 blocks of 25 trials were used.

Eye movement data was recorded using the SR Research Ltd Eyelink 1000 infrared eyetracker (500 Hz temporal resolution, $<0.01°$ RMS spatial resolution). Eye movements that exceeded an acceleration threshold of $8000°/s^2$ and a velocity threshold of 30°/s were classified as saccades. Participants viewed the stimuli with both eyes, but only the dominant eye was tracked. A chin and head rest were used throughout the experiment to stabilize head position and allow accurate eye tracking. At the start of each new level of difficulty, a calibration procedure was completed. Subsequently, the accuracy of the calibration was checked at the beginning of each block and a new calibration carried out if necessary.

Results

Trials where the participants failed to answer before the final frame (0.01%) were excluded from the analysis, as were RTs that were further than 2.5 *SD*s from the cell mean (1.8%). The data from one participant was removed, as they made incorrect responses on more than 35% of the trials for four different cells. All the remaining trials were used in the error analysis and only correct trials were used in the reaction time (RT) and eye movement analysis. A loss of the pupil during tracking can indicate a blink or a tracking error. Because of this, trials containing more than one episode of pupil loss were excluded from the eye movement analysis (1.9%) as potentially indicating tracking error or excessive blinking.

Summary of results. Both manual RTs and error rates (see Figure 2) were very similar to those found in Hulleman (2010): there were hardly any RT differences between moving and static on the present trials for any of the three difficulties. For the absent trials, RTs when items moved were a little slower for easy and medium search and this difference increased for hard search. Only in the hard condition were there more errors for moving. The three measures that tested for a sit-and-wait strategy did not detect any indication for it: the number of fixations (see Figure 3) closely followed the RT pattern, so there were always as many or even more fixations in the moving trials. Similarly, the fixation durations (Figure 4, top row) were only slightly longer for the moving items (maximally 53 ms) and the saccadic amplitudes (Figure 4, bottom row) were also very similar (saccades in the moving condition maximally 0.7° shorter). The three indicators for FVF did find a dependence on difficulty: drift for moving items (Figure 5, top row) increased with search difficulty. For both static and moving conditions, the distance to the nearest item (Figure 5, bottom row) decreased with increasing search difficulty, especially on absent trials. Finally, the direct estimate of FVF based on the combination of eye movements and item location (Figures 7 and 8) also showed a clear decrease in FVF size when task difficulty increased—both when items were moving and when they remained static.

Manual reaction times. Reaction time results for all three difficulty levels are shown in Figure 2. There were no main effects or interactions involving item speed for medium or hard conditions when the target was present and the main effect of speed for the easy condition involved a very small RT difference (17 ms). It is clear from Figure 2 that, for target present trials, RTs and RT slopes for moving and static items are almost identical. The RT pattern is very much like that found in Hulleman (2010).

This follows from the analysis of RTs where a $3 \times 2 \times 3 \times 2$ (difficulty $\times$ speed $\times$ display size $\times$ target) Greenhouse-Geisser corrected within-subjects ANOVA yielded a significant three-way interaction between difficulty, display size and target, $F(4, 48) = 34.7$, $p < .001$, $\eta^2 = .743$, and difficulty, speed and target, $F(2, 24) = 7.7$, $p < .02$, $\eta^2 = .391$. Because of these interactions, the analysis was split by difficulty, with separate $2 \times 3 \times 2$ (speed $\times$ display size $\times$ target) ANOVAs for the easy, medium and difficult tasks.

In the easy task, there was a significant interaction between all three factors, $F(2, 24) = 5.3$, $p < .02$, $\eta^2 = .305$, so the analysis was further split by target. For the absent trials a 2×3 (speed $\times$ display size) ANOVA found main effects of speed, $F(1, 12) = 8.8$, $p < .02$, $\eta^2 = .424$ (longer RTs for moving items), and display size, $F(2, 24) = 9.5$, $p < .01$, $\eta^2 = .442$ (longer RTs for larger display sizes) and a significant interaction between speed and display size, $F(2, 24) = 4.5$, $p < .04$, $\eta^2 = .274$ (the RT difference between static and moving increased with display size). However, for the present trials there was a main effect of speed only, $F(1, 12) = 8.0$, $p < .02$, $\eta^2 = .400$ with slightly longer RTs for moving items than static (static M = 532 ms, moving M = 549 ms).

A separate $2 \times 3 \times 2$ ANOVA on the medium task revealed a significant interaction between speed and target, $F(1, 12) = 25.5$,

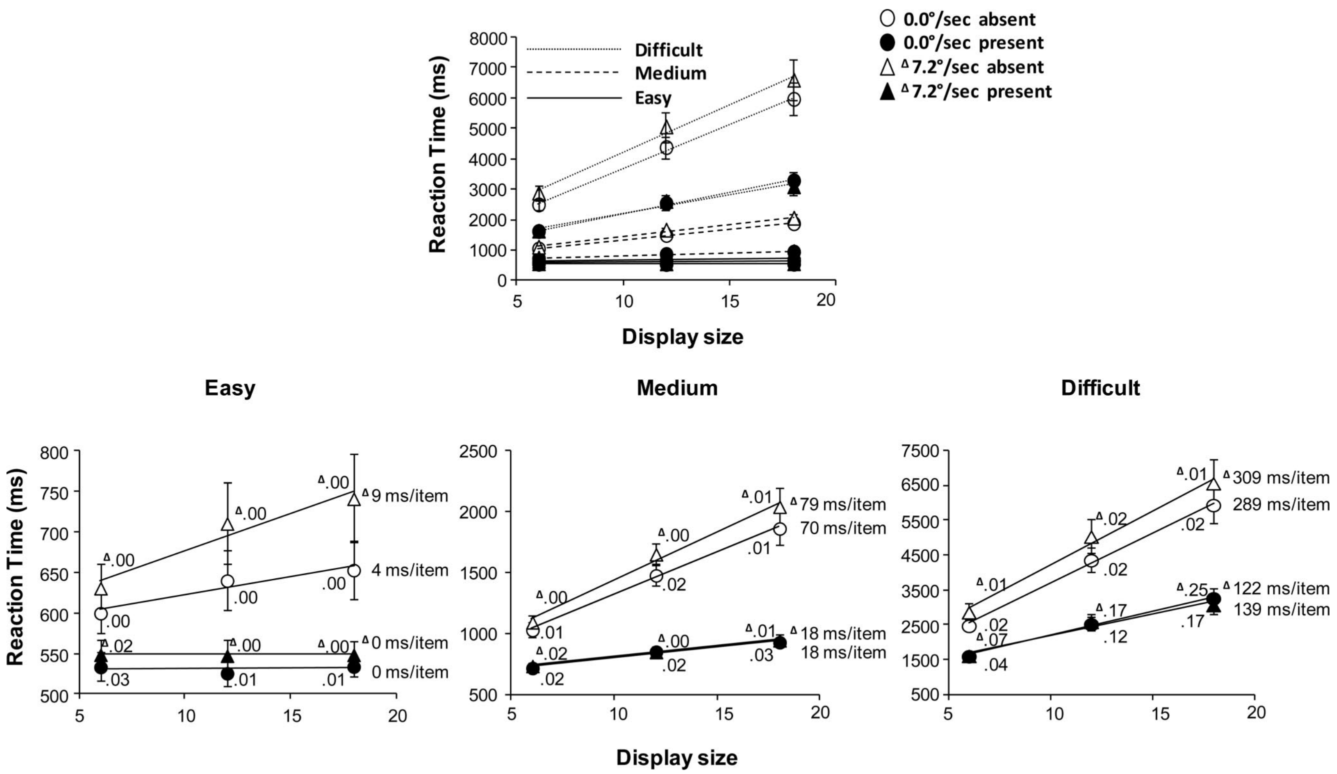

Figure 2. Reaction times as a function of display size (top), with separate plots for the easy, medium, and difficult tasks (left to right). Solid, dashed, and dotted lines represent the easy, medium, and difficult tasks respectively. Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). The error proportions for each condition are shown next to each symbol. The slope for each reaction time regression line is given to the right of the line. For error proportions and slopes, small triangles indicate that the data is from the moving item condition. Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

$p < .001$, $\eta^2 = .680$ and display size and target, $F(2, 24) = 46.2$, $p < .001$, $\eta^2 = .794$. Therefore, the analysis was further split by target. In the absent condition a separate ANOVA revealed main effects of speed, $F(1, 12) = 31.5$, $p < .001$, $\eta^2 = .724$ and display size, $F(2, 24) = 72.1$, $p < .001$, $\eta^2 = .857$ with longer RTs for moving items and larger display sizes. In the present condition a separate ANOVA revealed a main effect of display size only, $F(2, 24) = 40.8$, $p < .001$, $\eta^2 = .773$ with longer RTs for larger display sizes.

For the 2 × 3 × 2 ANOVA on difficult trials, the analysis also yielded significant interactions between speed and target $F(1, 12) = 10.6$, $p < .01$, $\eta^2 = .468$ and display size and target, $F(2, 24) = 52.6$, $p < .001$, $\eta^2 = .817$ so again the analysis was further split by target. For the absent trials there were main effects of speed $F(1, 12) = 13.8$, $p < .01$, $\eta^2 = .535$ (longer RTs for moving items) and display size $F(2, 24) = 76.4$, $p < .001$, $\eta^2 = .864$ (longer RTs for larger display sizes). For the present trials there was a main effect of display size only $F(2, 24) = 55.5$, $p < .001$, $\eta^2 = .822$ with longer RTs for larger display sizes.

Manual errors. The proportion of errors for all three difficulty levels are shown with RTs in Figure 2. Overall, the pattern of errors is similar to that found previously in Hulleman (2010), with more errors for moving trials only when the task was difficult. This impression was supported by a 3 × 2 × 3 × 2 (difficulty × speed × display size × target) Greenhouse-Geisser corrected within subjects ANOVA on error rates which found a significant three-way interaction between difficulty, speed and target $F(2, 24) = 8.9$, $p < .01$, $\eta^2 = .425$ and difficulty, display size and target, $F(4, 48) = 18.1$, $p < .001$, $\eta^2 = .601$. Because of these interactions, the analysis was again split along the difficulty dimension, with separate 2 × 3 × 2 (speed × display size × target) ANOVAs for each level of difficulty.

In the easy condition, there was a significant main effect of target, $F(1, 12) = 5.9$, $p < .04$, $\eta^2 = .329$, only. In the medium condition there were no significant main effects or interactions.

In the difficult condition, there was a dramatic increase in errors overall, with an error rate of 25% in the condition with most errors (moving, display size 18). Since there were significant interactions between speed and target, $F(1, 12) = 10.3$, $p < .01$, $\eta^2 = .462$, and display size and target, $F(2, 24) = 20.1$, $p < .001$, $\eta^2 = .627$, the analysis was split by target for the difficult condition. The absent trials yielded neither significant main effects nor a significant interaction. For present trials there was a significant main effect of display size, $F(2, 24) = 27.0$, $p < .001$, $\eta^2 = .693$, and a significant main effect of speed, $F(1, 12) = 8.9$, $p < .02$, $\eta^2 = .424$, with more errors at larger display sizes and more errors in the

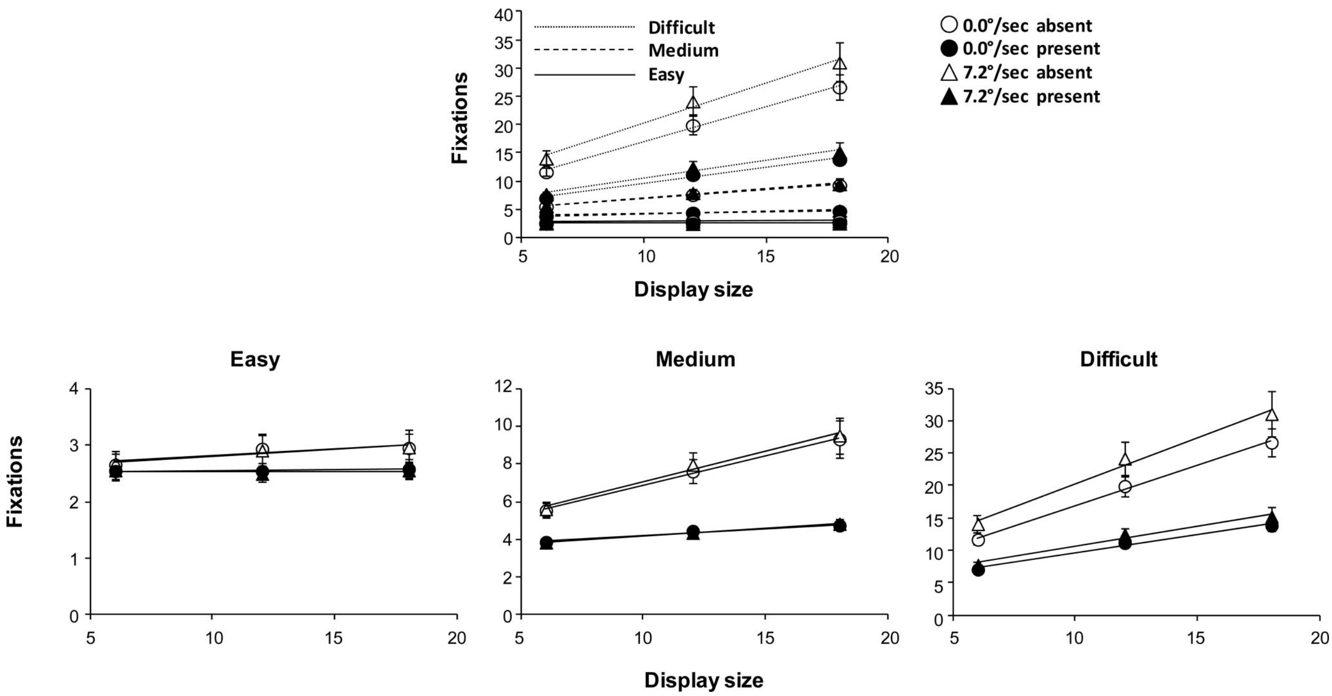

Figure 3. Average number of fixations per trial as a function of display size (top), with separate plots for the easy, medium, and difficult tasks (left to right). Solid, dashed, and dotted lines represent the easy, medium, and difficult tasks respectively Circles are static item and triangles are moving item trials (7.2°/sec). Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

moving item condition than the static item condition. Planned comparisons showed that there were significantly more errors for moving items than static in present trials with 18 items: $t(12) = 2.9$, $p < .02$, whereas this difference in error rate failed to reach significance for display size 6, $t(12) = 1.5$, $p < .17$, or 12, $t(12) = 1.6$, $p < .14$.

Number of fixations. Figure 3 shows the average number of fixations per trial for each condition. The pattern of results is very similar to that of RTs, with changes in RT typically accounted for by changes in number of fixations rather than fixation duration. Details of the statistical analysis can be found in the Appendix.

If participants used a sit-and-wait strategy, then we would expect fewer fixations for moving items than for static. However, we found no main effects of speed or any interaction involving speed when the target was present in the number of fixations during easy, medium, or difficult search tasks. Moreover, for the absent trials with moving items there were actually more fixations when the search task was difficult.

Fixation duration. Within each level of difficulty, the variation in fixation duration between conditions was fairly small (the maximum difference was 53 ms), as shown in the top row of Figure 4. There is therefore very little evidence of a sit-and-wait strategy in the fixation duration data. Fixation durations were only 9 to 47 ms longer for moving trials than static in the easy and medium absent conditions. This would hardly seem to qualify as waiting. Moreover, there was no main effect of speed in the absent trials of the difficult condition nor in the easy and medium target present conditions. In the target-present trials for difficult condition fixation durations were in fact 20 to 29 ms longer for static trials than moving.

This follows from a 3 × 2 × 3 × 2 (difficulty × speed × display size × target) Greenhouse-Geisser corrected within subjects ANOVA on average fixation durations which revealed significant interactions between difficulty and target, $F(2, 24) = 13.9$, $p < .01$, $\eta^2 = .536$, difficulty and speed, $F(2, 24) = 12.9$, $p < .01$, $\eta^2 = .518$, and speed and target, $F(1, 12) = 22.3$, $p < .001$, $\eta^2 = .650$. Again, the analysis was split by difficulty.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on easy trials revealed a significant interaction between speed and target, $F(1, 12) = 10.0$, $p < .01$, $\eta^2 = .455$. A separate ANOVA on the absent trials yielded a main effect of speed, $F(1, 12) = 7.8$, $p < .02$, $\eta^2 = .394$, with longer average fixation durations for moving items and a significant interaction between speed and display size, $F(2, 24) = 6.1$, $p < .02$, $\eta^2 = .338$, with average fixation durations for moving items remaining stable over display sizes, while average fixation durations for static items decreased with increasing display sizes. When the target was present there were no significant main effects and no interaction.

The results for medium difficulty trials were very similar to those for easy trials. A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA yielded an interaction between speed and target, $F(1, 12) = 11.4$, $p < .01$, $\eta^2 = .488$, so the analysis was split by target. When the target was absent there was a significant main effect of speed, $F(1, 12) = 49.3$, $p < .001$ $\eta^2 < .804$, with longer average fixation durations for moving items. There was also a significant interaction between speed and display size, $F(2, 24) = 3.4$, $p < .05$, $\eta^2 = .221$, with average fixation durations for moving items remaining stable over display sizes, while average fixation durations for static items decreased with increasing dis-

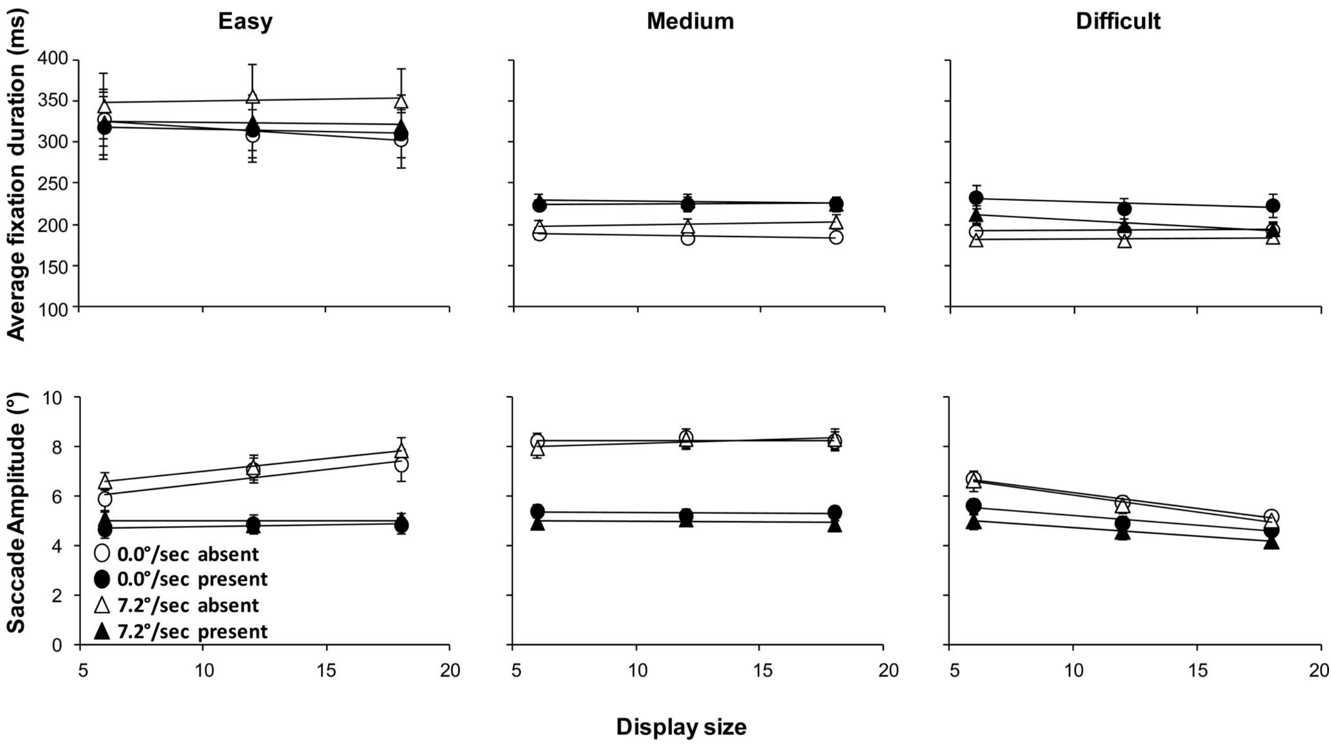

Figure 4. Average fixation duration (top) and saccade amplitudes (bottom) as a function of display size for the easy, medium, and difficult conditions (left to right). Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

play sizes. When the target was present there were no significant main effects and no interaction.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on difficult trials yielded significant interactions between speed and target, $F(1, 12) = 16.5$ $p < .01$, $\eta^2 = .579$, and between display size and target, $F(2, 24) = 19.7$, $p < .001$, $\eta^2 = .622$. Analysis of difficult target absent trials showed no significant main effects or interactions. Analysis of difficult target present trials showed a main effect of speed, $F(1, 12) = 9.9$, $p < .01$, $\eta^2 = .452$, with shorter fixations for moving items, and a significant main effect of display size, $F(2, 24) = 17.0$, $p < .001$, $\eta^2 = .587$, with fixation duration decreasing with increasing display size.

Saccade amplitude. Average saccade amplitudes for the three task difficulties are shown in Figure 4 (bottom row). From Figure 4 and the statistical analysis (see Appendix) it becomes clear that the effect of motion on saccade amplitude is minimal. Under a sit-and-wait strategy we might expect many short saccades in the same area for moving trials, compared to longer saccades, covering the whole display in static trials. The saccade amplitude data shows no evidence of this, with the condition with the largest difference between static and moving trials having saccade amplitudes only 0.7° shorter for moving trials.

Saccadic amplitude varied with the difficulty of the search task. For easy search, saccadic amplitude on absent trials increased for larger display sizes. For difficult search, saccadic amplitude on absent trials decreased as a function of display size. Furthermore, saccadic amplitudes were much more alike for present and absent trials in difficult search.

This, in fact, reflects the underlying individual saccades. For easy search, later saccades tended to be larger than earlier saccades in absent trials, whereas in present trials later saccades tended to be smaller. For difficult search, the dependence on rank of the saccade was much less.

Drift. The average drift in each condition is shown in Figure 5 (top row). From Figure 5 it becomes apparent that drift is larger for moving than for static displays, and that this difference increases with task difficulty. This is what our framework would lead us to expect. This was confirmed by a 3 × 2 × 3 × 2 (difficulty × speed × display size × target) Greenhouse-Geisser corrected within subjects ANOVA on drift data which yielded significant interactions between difficulty and speed, $F(2, 22) = 35.4$, $p < .001$, $\eta^2 = .763$, difficulty and display size, $F(4, 44) = 5.2$, $p < .02$, $\eta^2 = .319$, difficulty and target, $F(2, 22) = 4.0$, $p < .04$, $\eta^2 = .269$ and speed and display size, $F(2, 22) = 9.4$, $p < .01$, $\eta^2 = .462$. As before, the analysis was split by difficulty.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on easy trials yielded a significant interaction between all three factors, $F(2, 22) = 4.6$, $p < .03$, $\eta^2 = .294$. Since it is clear from Figure 5 that the relationships are quite different in the static and moving conditions, the analysis was split by speed. Two × three within-subjects ANOVAs revealed a main effect of display size when items were static, $F(2, 22) = 3.8$, $p < .05$, $\eta^2 = .256$, with slightly more drift at display sizes 12 and 18 than at display size 6, and no significant interaction or main effects when the items moved.

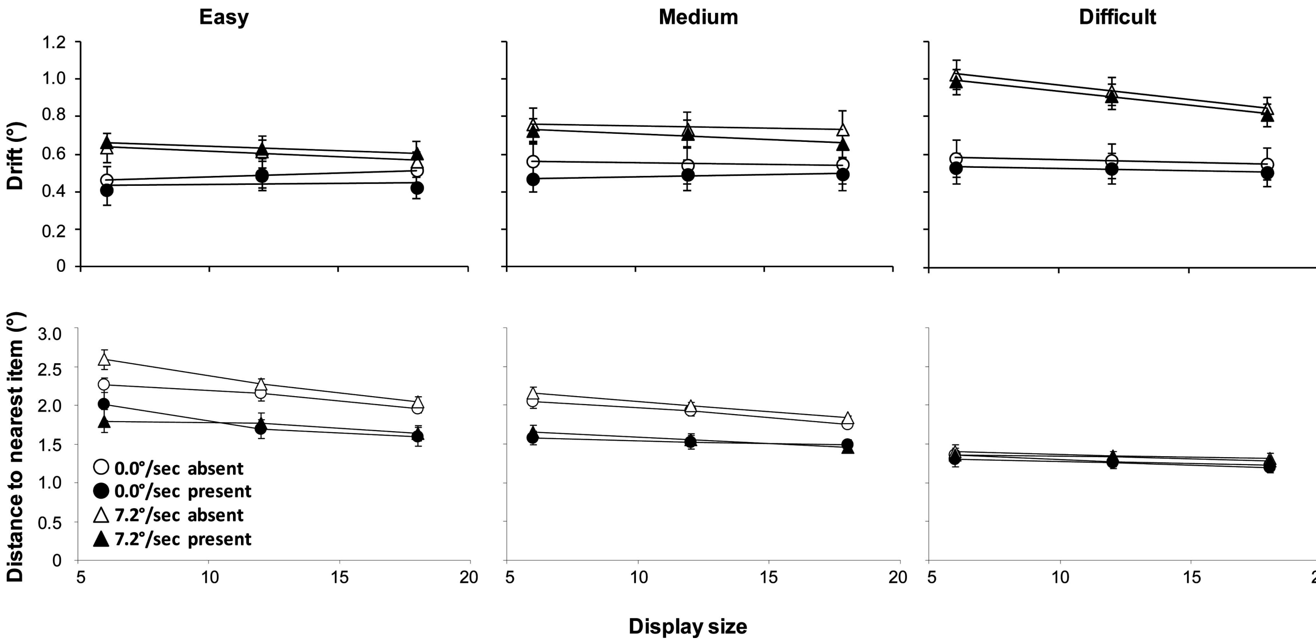

Figure 5. Average difference in gaze position between the end of one saccade and the start of the next saccade (drift: top) and average distance from the endpoint of a saccade to the nearest item (bottom) as a function of display size. Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on the medium difficulty trials yielded main effects of speed, $F(1, 12) = 111.4$, $p < .001$, $\eta^2 = .903$, and target, $F(1, 12) = 5.3$, $p < .04$, $\eta^2 = .308$, with more drift for moving items and less drift for larger display sizes.

A similar ANOVA on difficult trials yielded a significant interaction between speed and display size, $F(2, 24) = 11.1$, $p < .01$, $\eta^2 = .481$. Separate ANOVAs for static and moving items in the difficult condition revealed a significant main effect of target when items were static, $F(1, 12) = 5.4$, $p < .04$, $\eta^2 = .312$, with slightly more drift when the target was absent, and a significant main effect of display size when items were moving, $F(2, 24) = 31.5$, $p < .001$, $\eta^2 = .724$, with less drift for larger display sizes.

Distance to nearest item. We computed the average distance between the saccadic endpoint and the nearest item (see bottom row of Figure 5). Figure 5 shows that the distance to the nearest item decreases with increasing difficulty. This is especially clear on the absent trials. On present trials the effect is slightly weaker, but it should be kept in mind that participants might actually have fixated the target. This would lead to an underestimation of the distance; in particular for the easy and medium conditions, where only few saccades were made to begin with.

This was confirmed by a 3 × 2 × 3 × 2 (difficulty × speed × display size × target) Greenhouse-Geisser corrected within subjects ANOVA on distance to nearest item which yielded significant interactions between difficulty and display size, $F(4, 48) = 9.9$, $p < .001$, $\eta^2 = .452$, difficulty and target, $F(2, 24) = 19.9$, $p < .001$, $\eta^2 = .624$ and speed and target, $F(1, 12) = 6.4$, $p < .03$, $\eta^2 = .347$. A three-way interaction between difficulty, speed, and target was also close to significance, $F(2, 24) = 4.1$, $p = .052$, $\eta^2 = .225$.

In a separate 2 × 3 × 2 (speed × display size × target) ANOVA for easy trials there was a significant interaction between speed and target, $F(1, 12) = 5.3$, $p < .05$, $\eta^2 = .307$, with a closer nearest item for static trials than moving when the target is absent. There were also significant main effects of display size, $F(2, 24) = 25.5$, $p < .001$, $\eta^2 = .680$, and target, $F(1, 12) = 31.7$, $p < .001$, $\eta^2 = .725$, with a closer nearest item for larger display sizes and when the target was present.

In a similar ANOVA for medium difficulty trials, there was a significant interaction between display size and target, $F(2, 24) = 12.9$, $p < .001$, $\eta^2 = .518$, with nearest item distance decreasing more with display size when the target is absent than when it is present. There were also significant main effects of all three factors: speed, $F(1, 12) = 4.7$, $p = .05$, $\eta^2 = .283$, display size, $F(2, 24) = 22.5$, $p < .001$, $\eta^2 = .652$, and target, $F(1, 12) = 72.9$, $p < .001$, $\eta^2 = .859$, with a closer nearest item for static items, smaller display sizes and when the target is present.

In the difficult condition there were no significant interactions, but once again there were main effects of all three factors: speed, $F(1, 12) = 32.4$, $p < .001$, $\eta^2 = .730$, display size, $F(2, 24) = 5.7$ $p < .03$, $\eta^2 = .323$, and target, $F(1, 12) = 5.3$, $p < .05$, $\eta^2 = .308$. As with the medium difficulty search, these main effects involved a closer nearest item for static items, smaller display sizes and when the target is present.

Functional visual field. In order to estimate the size of the FVF for the three different search difficulties, gaze position and item position data were used (see Figure 6). At the end of each

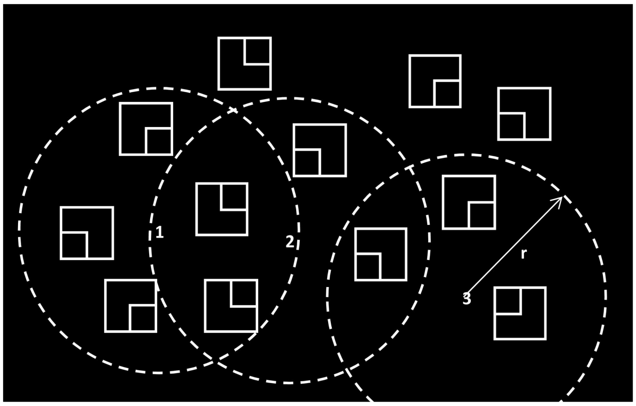

Figure 6. Illustration of FVF analysis technique. Dashed circles show one example of a radius around each saccade endpoint across the course of a trial. Each item can only be counted once, so the total count for this radius in this example would be 9. In the actual analysis, many different radii were used.

saccade, the gaze coordinates were recorded and the number of items in the display falling within a given radius around the gaze coordinates was counted. The radii used started at 1.0° visual angle and increased in steps of 1.0°. Across each trial, for each circle radius, the number of items inside the radius was summed (every item was counted only once), to give the total number of items processed across the trial if the FVF were to have that particular radius. Figure 6 illustrates how items were counted at a given radius (r). In the case of Figure 6, a total of nine items would be counted across the three fixations made during the trial.

To establish the radius of the FVF, we used the principle that, on average, the target will be found at the point where half of the items have been processed. Figure 7 shows the proportion of items, across an average trial, falling within each tested circle radius for target present trials. The 50% line shows that the FVF decreased when the task was more difficult. For the easy task the estimated radius of the FVF was 8.7 to 9.7°. For the medium task, the estimate was 5.8 to 6.8°. For the difficult task, the boundaries of the FVF did not sit clearly between two circles separated by 1°, so a finer grained analysis was carried out and the radius if the FVF was found to be between 1.7 and 2.2°. It is important to note that the estimates of the FVF are very similar for both moving and static displays.

Figure 8 shows an analysis for absent trials that used the estimates of the FVF. For each difficulty it is worth noting that the cumulative number of items inside each FVF fell short of 100%, with only 70 to 72% of items counted on easy trials, 85 to 92% on medium trials, and 70 to 88% for difficult trials. The low upper limit for the easy trials suggests that 8.7 to 9.7° might actually be an underestimation of the FVF for the easy trials (which would occur if participants fixate the target once they have found it). Using a criterion of 90% of items visited in absent trials would yield an estimate of 12 to 13°.

Finally, Figure 9 shows that for every level of difficulty, there is a substantial number of items that repeatedly falls within the FVF. This is even the case for static displays that contain only six items. This is consistent with the assumption of parallel processing within

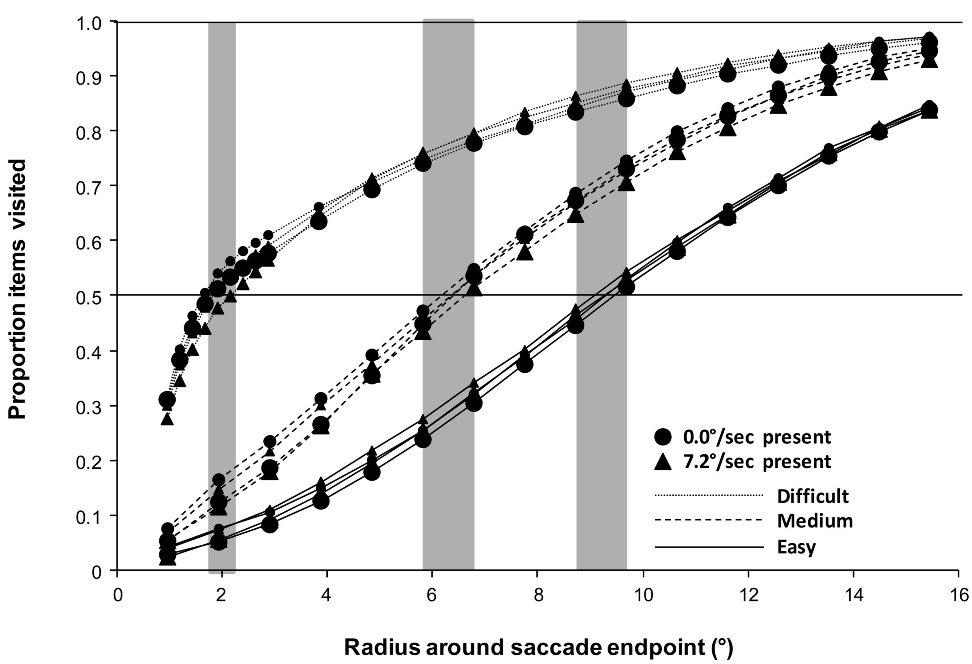

Figure 7. Average proportion of unique items counted in a target-present trial as a function of the radius of the circle within which items were counted. Circles are static item and triangles are moving item trials (7.2°/sec). Dotted, dashed, and solid lines show difficult, medium, and easy tasks, respectively. Large and small symbols represent 18 and 12 item displays, respectively. Gray bars represent the radii at which each of the easy, medium, and difficult tasks reach 50%: the estimated FVF.

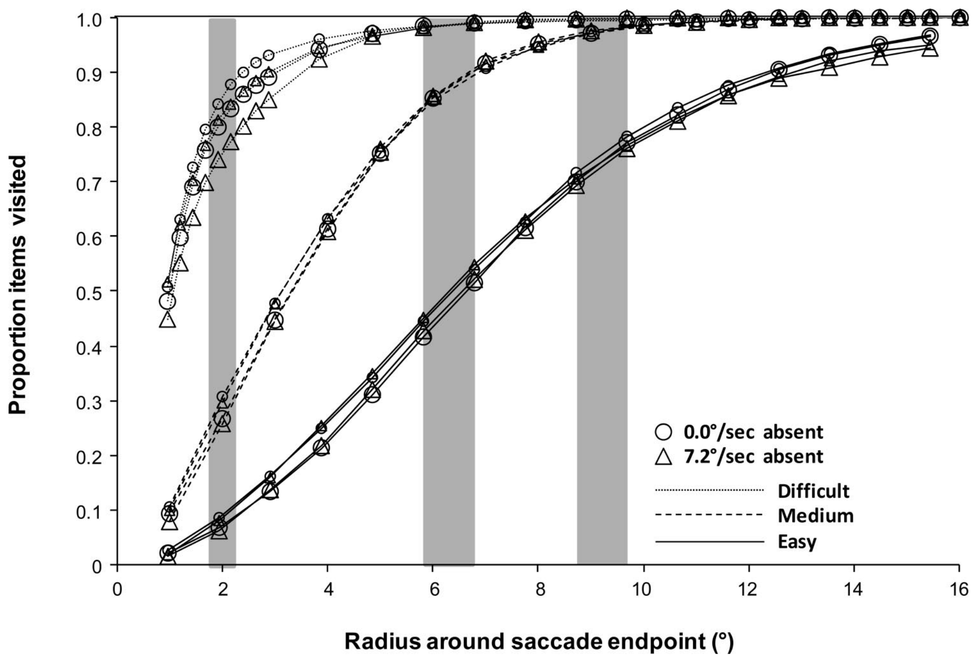

Figure 8. Average proportion of unique items counted in a target-absent trial as a function of the radius of the circle within which items were counted. Circles are static item and triangles are moving item trials (7.2°/sec). Dotted, dashed, and solid lines show difficult, medium, and easy tasks respectively. Large and small symbols represent 18 and 12 item displays, respectively. Gray bars represent the estimated FVFs for the easy, medium, and difficult tasks.

fixations made by our framework. When items are processed in parallel, reinspections of individual items should not be detrimental to search performance and do not have to be avoided.

Discussion

The results of Experiment 1 are in line with those of Hulleman (2009, 2010). Robustness against item motion was found for easy and medium difficulty search. Although the difference in RT between static and moving items was significant for present trials in the easy condition, this difference was very small, with a maximum increase of less than 3 ms per item and no interaction with display size. As seen in Hulleman (2010), robustness began to break down in the case of difficult search, where there was still no effect of speed on RTs, but a significant increase in errors for displays with 18 items.

It is important to note that the eye movement data showed no sign of a different search strategy being used for search among moving items. In particular, the hallmarks of a sit-and-wait strategy were not in evidence. There was no systematic increase in fixation durations or decrease in number of fixations for search among moving items; nor was there a decrease in saccade amplitude for search among moving items.

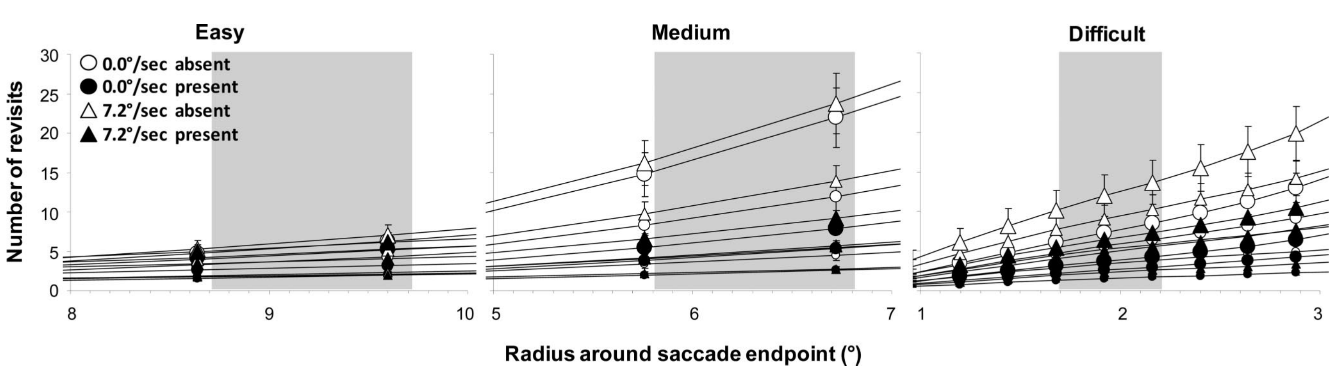

Figure 9. Average number of times within a trial that an item fell within the FVF after having been in the FVF at the end of any earlier saccade as a function of task difficulty and display size. Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). Symbol size indicates display size, with smaller symbols indicating smaller display sizes.

Consistent with our theoretical framework, the number of fixations increased with increasing task difficulty, with little variation in fixation duration or saccade amplitude. It is interesting that the size of the drift in eye position between the end of the previous and the start of the current saccade was not only larger for moving items than for static items, the difference also increased for more difficult search. This dovetails with an account of visual search in which fewer items are monitored within a fixation as search becomes more difficult. For easier search, the FVF is large. So, there are many items, all moving in different directions. As a result, the net motion vector would be rather small. For a difficult search, where the FVF only encompasses one or two items, the net motion vector will become larger, since there is less opportunity for the motion of the individual items to cancel each other out. The shrinking of the FVF with increased task difficulty was also reflected in the distance between saccadic end point and nearest item. Because of the smaller FVF in the difficult search task, saccades were more closely targeted to specific items. This drop in distance to the nearest item with increased difficulty was especially pronounced in absent trials. It is interesting that there is a much larger drop in distance to the nearest item on absent trials when the difficulty of the search task goes from medium to hard than when it goes from easy to medium. This is consistent with the assumption of our framework that there is parallel processing in the easy and medium task, whereas the difficult task is purely serial.

Finally, we estimated the size of the FVF for the three search difficulties. As suggested by our theoretical framework, the results showed smaller FVFs for more difficult searches. The estimates of the FVF were very similar for moving and static search displays. This is in keeping with the assumption of our framework that there is a single search process active during search among static and search among moving displays. (Convergence in the estimates of an attentional property based on moving and static display has been reported before, Intriligator & Cavanagh, 2001).

When the estimates of the FVF are applied to the eye movement data, it becomes clear that there are always revisits to items. In fact, there are more revisits during a medium difficulty search than during a hard search even though the latter is much more vulnerable to motion than the former. This is consistent with the assumption of our framework that because of parallel processing inside an FVF, revisits to individual items are not very detrimental to performance. Only when the FVF has become very small do revisits reduce search performance. Because only a single item is processed per fixation, each revisit increases the RT substantially.

Even in the easiest search condition some fixations were made, and despite the task being typical of what would be considered to be a parallel task, the FVF does not encompass the entire display (even if we use the more generous estimate based on the absent trials), suggesting that very easy search may not be purely parallel when items fall well outside of the parafoveal area.

However, for the moment, our estimates of the FVFs are little more than mathematical constructs. In fact, it could be argued that the reduction in the radius of the FVF as a function of the difficulty of the search task is completely driven by the increase in number of fixations. Because we have defined the size of the FVF as that radius around the location of fixation that will accrue half of the items in the display across all fixations, the FVF has to become smaller when the number of fixations increases. One aspect in the data of Experiment 1 that would speak against this argument is that the estimated FVFs were fairly constant across display sizes 12 and 18 for each of the three difficulties. Since there are more fixations in the 18-item displays than in the 12-item displays, this constancy would seem to go against a purely mathematical scaling effect. Moreover, the increase in drift and the reduced distance between saccade endpoint and nearest item could also be interpreted as evidence for a reduction in FVF due to increased search difficulty. However, if we want to make the case that the FVF depends on search difficulty more solidly, we will have to show that in easier search, information about the items is extracted from a larger part of the search display.

Experiment 2

To this end, we used a gaze-contingent window paradigm in Experiment 2 with three window sizes (2.4°, 4.9°, and 9.7°). In this paradigm, only the items that were within these distances of the current gaze coordinates were shown. The items that fell outside were masked. If the FVF is indeed larger for easier search tasks, then a gaze contingent window that is larger than our estimate of the difficult FVF (1.7–2.2°), but smaller than our estimates of the easy and medium FVFs (8.7–9.7° and 5.8–6.8°, respectively) should interfere only with the search process for the easy and medium difficulty task, leading to a differential effect on search efficiency. Moreover, when the FVF is estimated from the eye movement data in this gaze contingent experiment, its size should scale with the size of the window as soon as the gaze contingent window becomes smaller than the "natural" FVF for the task.

Furthermore, our theoretical framework explains robustness against item motion as the consequence of being able to process several items per fixation in parallel and the breakdown in that robustness for difficult search as the result of being forced to process only one item per fixation. If this is the case, then we would expect that the robustness against item motion should break down even for easier search, whenever there is only one visible item per fixation. Specifically, we would expect the same pattern of errors for easy and medium search when gaze is limited to only one or two items (2.4° radius around the point of fixation), as we typically find for difficult search, with many errors for static and even more errors for moving items.

Method

Participants. Fifty students (12 male, four left-handed, 18–36 years, average age = 20) participated in this experiment. They were naive to the purpose of the experiment.

Stimuli. The stimuli used in Experiment 2 were the same as those used in Experiment 1.

Procedure and design. The apparatus was the same as in Experiment 1 and the tasks were the same as those used in the easy, medium, and difficult conditions of Experiment 1. The eye tracker fed back gaze location data to the presentation computer in real time allowing us to limit the visible area to the region being fixated. The size of the visible area of the search display was limited to a radius around fixation of either 2.4°, 4.9°, or 9.7° (small, medium, and large, respectively). Figure 10 shows an example of an easy display with a large visible window, a medium difficulty display with a small visible window and a difficult

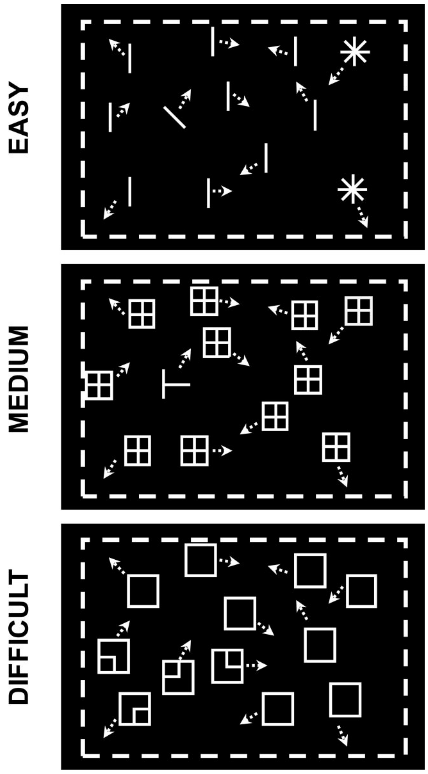

Figure 10. Illustration of the displays used in Experiment 2. The top panel shows the easy condition with the largest window (9.7°), the middle panel shows the medium difficulty condition with the smallest window (2.4°), and the bottom panel shows the difficult condition with the medium window (4.9°). Arrows illustrate movement and the dashed line illustrates the virtual rectangle; these were not visible during the actual trials.

display with a medium visible window. Display size was 12 or 18 items. Task difficulty was a between subjects factor, with 18 participants carrying out the difficult task, 16 the medium difficulty task and 16 carrying out the easy task. The procedure was identical for all three groups. The four within-subjects factors (window size, display size, speed and target) were fully crossed, yielding $3 \times 2 \times 2 \times 2 = 24$ cells for analysis, with 20 trials per cell, a total of 480 trials per participant. The trials were blocked by window size, which was counterbalanced across participants. For each window size, participants were presented with 8 blocks of 20 trials.

Results

Trials where the participants failed to answer before the final frame (0.2%) were excluded from the analysis, as were RTs that were further than 2.5 *SD*s from the cell mean (1.5%). The data from one participant from the medium difficulty condition and two participants from the difficult condition was removed from the analysis, as they made incorrect responses on more than 35% of the trials for three or more cells. All the remaining trials were used in the error analysis and only correct trials were used in the RT and eye movement analysis. Trials containing more than one episode of pupil loss were excluded from the eye movement analysis (1.9%) as potentially containing tracking errors or excessive blinking.

Summary of results. For all three difficulties, the manual RTs (see Figure 11) increased with decreasing window size. But this increase was much larger for the easy and medium difficulty tasks than for the difficult search task. Similarly, the error proportions (see Figure 11) for easy and medium difficulty search increased with decreasing window size. The number of fixations (see Figure 12) followed the RT pattern: there was a much larger increase for easy and medium difficulty search when the window size decreased. The fixation durations (see Figure 13) remained relatively constant across the different window sizes, although there were some slight variations. Drift (see Figure 14) clearly depended on window size for easy and medium difficulty search, with more drift for smaller windows. For difficult search this increase was much less. As the window size decreased, the direct estimate of FVF (see Figure 15) for easy and medium difficulty search became smaller as well. For difficult search, the direct estimate of the FVF did not depend on the window size.

Manual reaction times. Figure 11 shows RTs for the easy task in the top row, the medium difficulty task in the middle row, and for the difficult task in the bottom row. For medium and difficult search tasks, the pattern of results for the largest window is similar to that in Experiment 1. For the easy task there is a difference, since there is now an effect of display size for the absent trials. This is consistent with the observation made above that the larger FVF-estimate (12°–13°) for the easy condition based on the absent trials in Experiment 1 might be more appropriate.

For all three tasks, RTs increase when the visible window decreases. However, the effect of window size is much larger for the easy and medium difficulty tasks than for the difficult search task. Experiment 1 estimated the easy task FVF between 8.7 and 13°; the medium difficulty FVF between 5.8 and 6.8° and the difficult FVF between 1.7 and 2.2°. When the visible window was 2.4°—larger than the upper limit of the estimated FVF for difficult search, but much smaller than the lower limit of the FVF's for easy and medium search—RTs increased by only between 333 and 762 ms for difficult search from those for the largest window size. This is considerably smaller than the increase from the largest window for easy and medium search which was by between 1362 and 4723 ms.

A $3 \times 3 \times 2 \times 2 \times 2$ (difficulty $\times$ window size $\times$ speed $\times$ display size $\times$ target) Greenhouse-Geisser corrected mixed design ANOVA on the RTs yielded—among others—significant three-way interactions between difficulty, windows size and speed, $F(4, 88) = 2.8$, $p < .05$ $\eta^2 = .111$; difficulty, window size and target, $F(4, 88) = 23.6$, $p < .001$, $\eta^2 = .518$; difficulty, window size, and display size $F(4, 88) = 19.5$, $p < .001$, $\eta^2 = .469$, and difficulty, display size, and target $F(2, 44) = 3.4$, $p < .05$, $\eta^2 = .133$. Because of these interactions, separate within-subjects ANOVAs were carried out for the easy, medium, and difficult search tasks.

For the easy task, there were significant interactions between window size, speed, and target, $F(2, 30) = 6.6$, $p < .02$, $\eta^2 = .306$ and between window size, display size, and target, $F(2, 30) = 20.2$, $p < .001$, $\eta^2 = .574$. We therefore split the analysis further into target present and target absent ANOVAs.

The ANOVA for the absent trials found significant main effects of window size, display size, and speed $F(2, 30) = 240.9$, $p <$

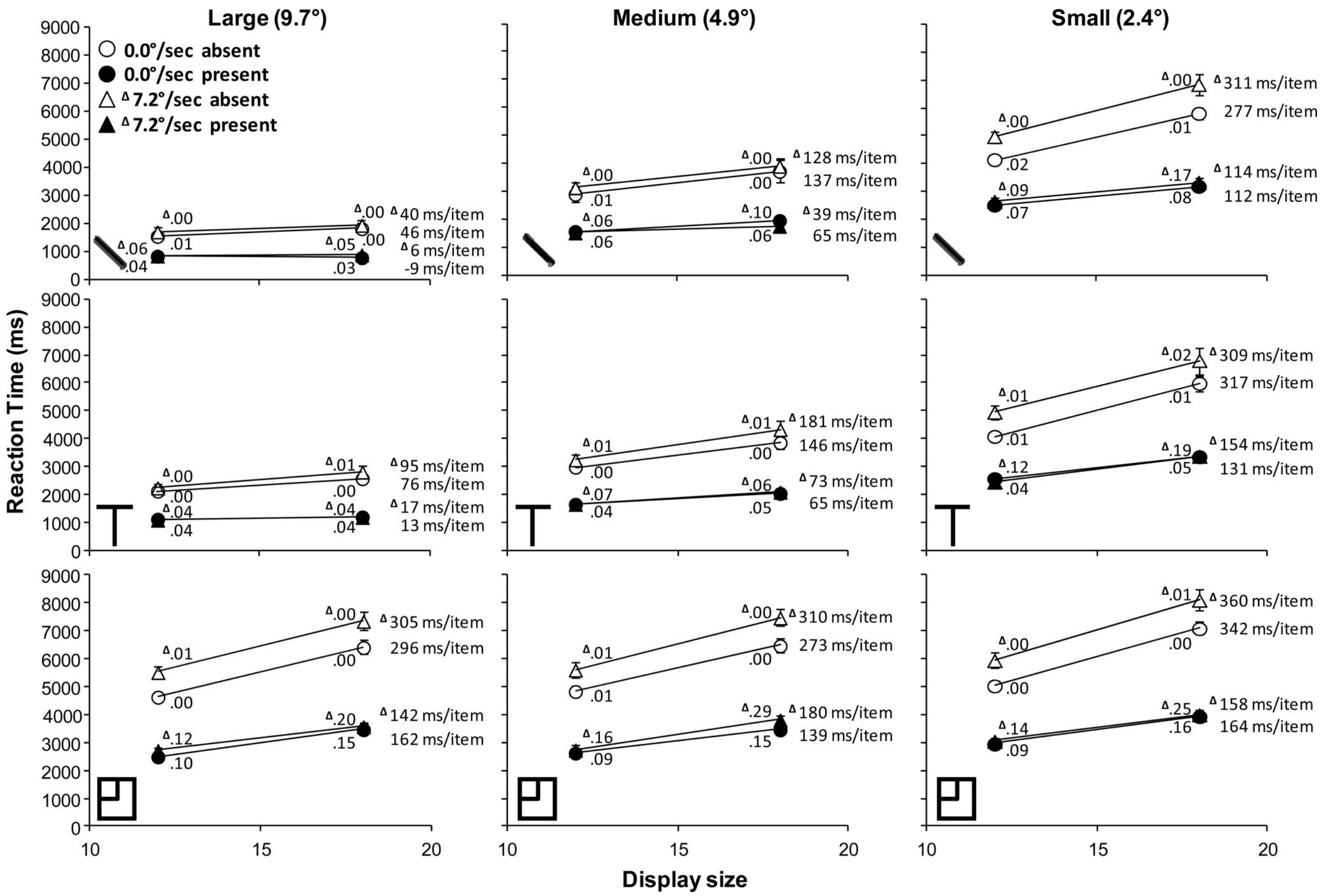

Figure 11. Reaction times as a function of display size for the easy, medium, and difficult tasks (top row to bottom row) with a large, medium, or small window (left to right). Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). The error proportions for each condition are shown next to each symbol. The slope for each reaction time regression line is given to the right of the line. For error proportions and slopes, small triangles indicate that the data is from the moving item condition. Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

.001, η^2 = .941, $F(1, 15) = 131.0$, $p < .001$, η^2 = .871 and $F(1, 15) = 17.2$, $p < .001$, η^2 = .534, respectively, with RTs increasing for smaller windows, larger display size, and moving items. There were also two-way interactions between window size and display size $F(2, 30) = 110.0$, $p < .001$ η^2 = .880 indicating that the effect of display size increased for smaller windows, and between window size and speed $F(2, 30) = 7.3$, $p < .015$, η^2 = .326, showing that the effect of motion increased for smaller window sizes. No other interactions involving speed were significant (*p*s > .20).

The ANOVA for the present trials found no effects involving speed (all *p*s > .12). The main effects of window size and display size and the interaction between the two were significant though: $F(2, 30) = 276.0$, $p < .001$, η^2 = .948, $F(1, 15) = 83.4$, $p < .001$, η^2 = .848 and $F(2, 30) = 23.2$, $p < .001$, η^2 = .607, respectively, with the effects going in the same direction as for the absent trials.

For the medium task, there were significant interactions between window size, speed, and target, $F(2, 28) = 9.4$, $p < .01$, η^2 = .402, and between window size, display size, and target, $F(2, 28) = 12.0$, $p < .001$, η^2 = .463. Because of these interactions, the analysis of the medium difficulty task was split into target present and target absent ANOVAs. The ANOVA for absent trials revealed significant main effects of window size, display size, and speed, $F(2, 28) = 244.9$, $p < .001$, η^2 = .946, $F(1, 14) = 105.5$, $p < .001$, η^2 = .883 and $F(1, 14) = 14.7$, $p < .01$, η^2 = .511, respectively, with RTs increasing for smaller windows, larger display size, and moving items. Interactions between window size and display size, $F(2, 28) = 80.5$, $p < .001$, η^2 = .852, and window size and speed, $F(2, 28) = 11.3$, $p < .01$, η^2 = .446, were also significant, with larger effects of display size and speed for smaller windows. The target present ANOVA yielded main effects of window size, $F(2, 28) = 297.7$, $p < .001$, η^2 = .955, and display size, $F(1, 14) = 72.6$, $p < .001$, η^2 = .838, as well as a significant interaction between window size and display size, $F(2, 28) = 29.2$, $p < .001$, η^2 = .676, but no interaction between speed and window size or a main effect of speed.

For the difficult task, there was a significant interaction between speed and target, $F(1, 15) = 29.8$, $p < .001$, η^2 = .665, and between display size and target, $F(1, 15) = 80.6$, $p < .001$, η^2 =

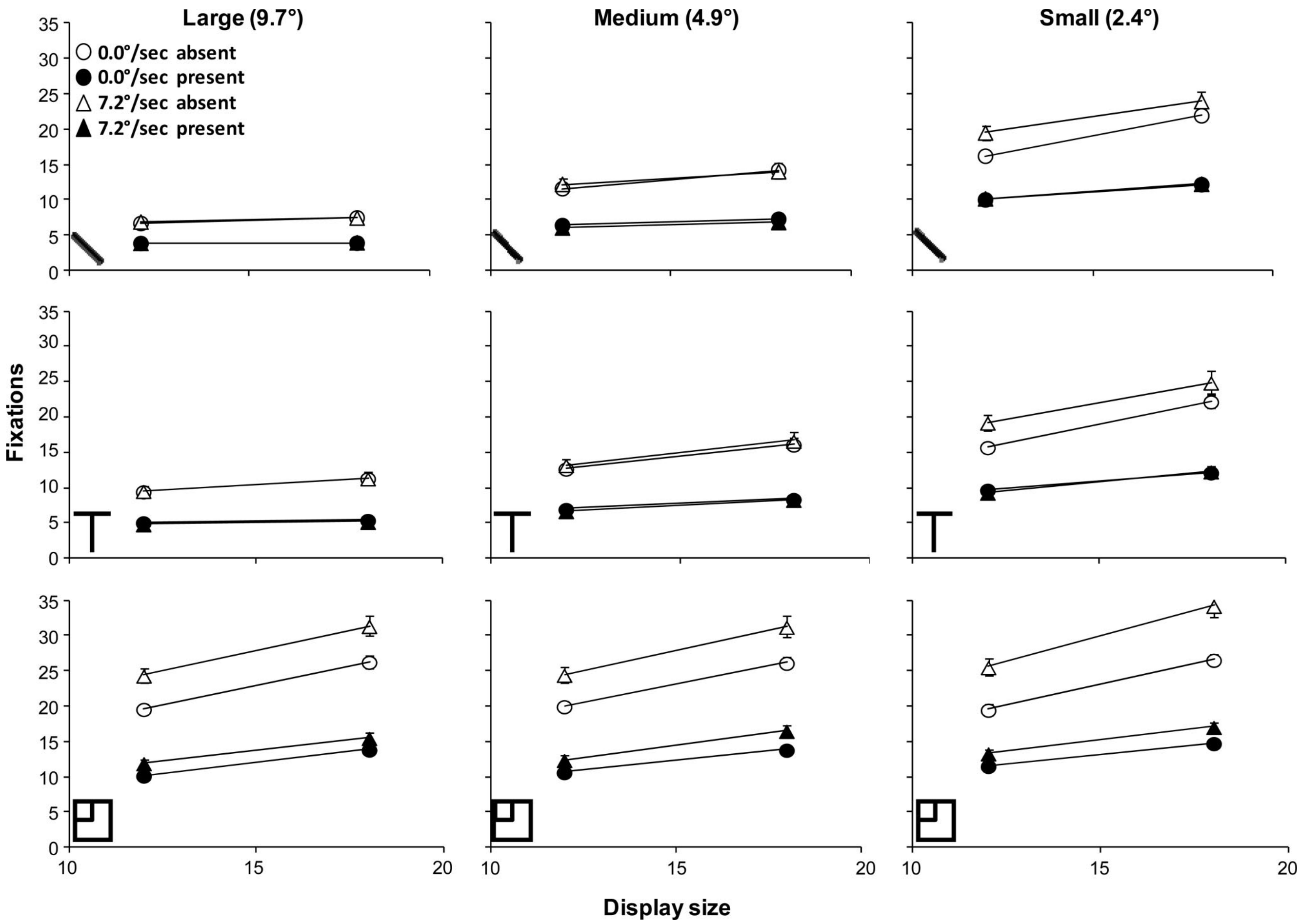

Figure 12. Average number of fixations per trial as a function of display size for the easy, medium, and difficult tasks (top row to bottom row) with a large, medium, or small window (left to right). Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

.843. Separate analyses were performed for target absent and target present trials. For target absent trials there was a significant interaction between window size and display size, $F(2, 30) = 5.3$, $p < .02$, $\eta^2 = .261$, with a greater effect of window size when the display size was 18, and significant main effects of window size, $F(2, 30) = 3.6$, $p < .05$, $\eta^2 = .194$, speed, $F(1, 15) = 32.1$, $p < .001$, $\eta^2 = .681$, and display size, $F(1, 15) = 480.7$, $p < .001$, $\eta^2 = .970$, with longer RTs for smaller windows, moving items, and larger display size. For target present trials there were significant main effects of window size, $F(2, 30) = 5.6$, $p < .01$, $\eta^2 = .274$, speed, $F(1, 15) = 6.7$, $p < .03$, $\eta^2 = .309$, and display size, $F(1, 15) = 277.6$, $p < .001$, $\eta^2 = .949$, with differences in the same direction as for target absent trials.

Manual errors. The proportion of errors in each condition is shown with the RTs in Figure 11, with each proportion next to the appropriate symbol. As before, almost all errors were made on target present trials. It is important to note that there were many more errors for moving than for static items in the easy and medium difficulty search tasks when the window size was smallest. For the difficult search task, the difference in errors between moving and static did not depend on window size. It did depend on display size though; only for the largest display size were there more errors moving items than for static.

This was illustrated by a $3 \times 3 \times 2 \times 2 \times 2$ (difficulty $\times$ window size $\times$ speed $\times$ display size $\times$ target) Greenhouse-Geisser corrected mixed design ANOVA on the error rates which yielded a significant four way interaction between difficulty, window size, speed and target, $F(4, 88) = 4.4$, $p < .005$, $\eta^2 = .166$. As a consequence, separate within-subjects ANOVAs were carried out for the easy, medium, and difficult search tasks.

For the easy task, there was a significant three-way interaction between window size, display size, and target $F(2, 30) = 7.7$, $p < .005$, $\eta^2 = .340$. Because of this interaction, $2 \times 2 \times 2$ (speed $\times$ display size $\times$ target) ANOVAs were carried out for each window size.

For the largest window, there was a main effect of target $F(1, 15) = 18.2$, $p < .001$, $\eta^2 = .549$ (more errors on present trials). The main effect of display size $F(1, 15) = 4.3$, $p < .06$, $\eta^2 = .224$

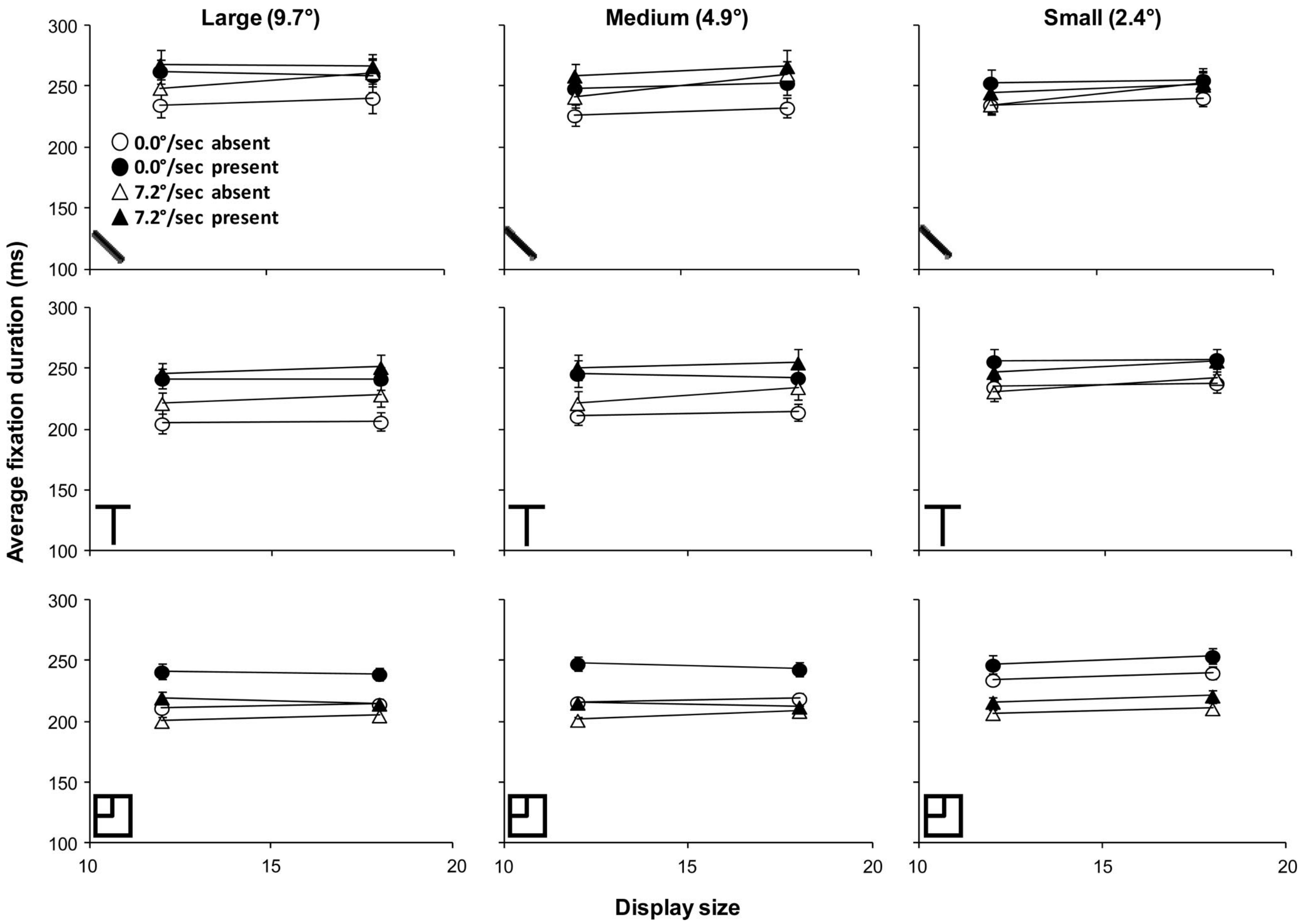

Figure 13. Average fixation duration as a function of display size for the easy, medium, and difficult tasks (top row to bottom row) with a large, medium, or small window (left to right). Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

(more errors for the smallest display size) and the interaction between speed and target $F(1, 15) = 4.1$, $p < .07$, $\eta^2 = .213$ (more errors for moving trials only on present trials) were almost significant.

For the medium window, there was a main effect of target $F(1, 15) = 18.0$, $p < .001$, $\eta^2 = .546$ (more errors on present trials) and significant interactions between display size and speed $F(1, 15) = 5.0$, $p < .05$, $\eta^2 = .252$, (more errors for moving trials only for the largest display size) and between display size and target $F(1, 15) = 9.1$, $p < .01$, $\eta^2 = .377$ (the difference between present and absent trials was larger for the largest display size). The interaction between speed and target was not significant $F(1, 15) = 3.3$, $p < .10$, $\eta^2 = .178$.

For the smallest window, there were now main effects of speed $F(1, 15) = 5.2$, $p < .04$, $\eta^2 = .256$ (more errors on motion trials); display size $F(1, 15) = 12.3$, $p < .005$, $\eta^2 = .450$ (more errors for the largest display size) and target $F(1, 15) = 24.8$, $p < .001$, $\eta^2 = .624$ (more errors on absent trials). As well as a clear interaction between speed and target $F(1, 15) = 8.5$, $p < .015$, $\eta^2 = .361$ (only more errors for motion on the present trials) and between display size and presence $F(1, 15) = 15.2$, $p < .002$, $\eta^2 = .504$ (the difference between present and absent trials was larger for the largest display size). A planned comparison showed that there were more errors for moving items than static for display size 18, $t(15) = 2.51$, $p < .025$. For display size 12, the difference between moving and static, although in the expected direction, was not significant $t(15) = 1.59$, $p < .15$.

For the medium task, there were significant interactions between window size, speed, and display size, $F(2, 28) = 4.5$, $p < .04$, $\eta^2 = .244$, between window size, speed, and target, $F(2, 28) = 15.4$, $p < .01$, $\eta^2 = .524$ and between window size, display size and target, $F(2, 28) = 4.8$, $p < .02$, $\eta^2 = .254$. As before, the analysis was split along window size. For the large and medium windows, there was a main effect of target only, $F(1, 14) = 7.3$, $p < .02$, $\eta^2 = .343$ and $F(1, 14) = 6.4$, $p < .03$, $\eta^2 = .314$, respectively, with more errors in target present trials. For the small window, there were many more errors for moving than static trials when the target was present, particularly at the larger display size. The ANOVA showed significant interactions between speed and target, $F(1, 14) = 16.9$, $p < .001$, $\eta^2 = .547$, between speed and

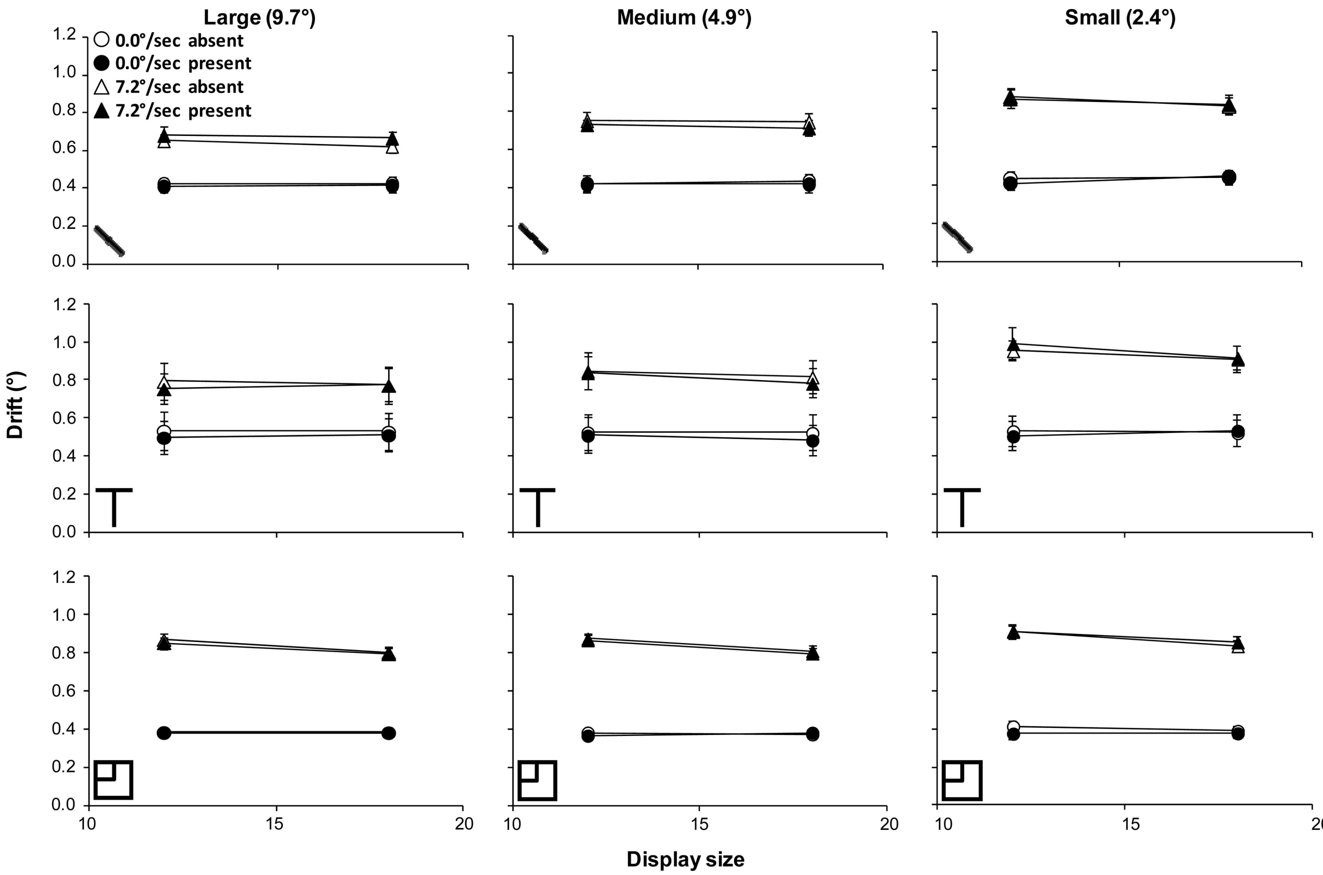

Figure 14. The average difference in gaze position between the end of one saccade and the start of the next saccade (drift) as a function of display size for the easy, medium, and difficult tasks (top row to bottom row) with a large, medium, or small window (left to right). Open symbols are target-absent trials; black symbols are target-present trials. Circles are static item and triangles are moving item trials (7.2°/sec). Error bars indicate standard errors; whenever error bars seem missing, they are covered by the data point.

display size, $F(1, 14) = 7.5$, $p < .02$, $\eta^2 = .350$ and between display size and target, $F(1, 14) = 5.3$, $p < .04$, $\eta^2 = .275$. Planned comparisons showed there were significantly more errors for moving items than static at both display size 12, $t(14) = 2.3$, $p < .04$ and display size 18, $t(14) = 4.9$, $p < .001$.

For the difficult task, there were no main effects or interactions involving window size. However, there were significant interactions between speed and target, $F(1, 15) = 34.7$, $p < .001$, $\eta^2 = .698$, with more errors for moving trials than static in target present trials only, and between display size and target, $F(1, 15) = 54.6$, $p < .001$, $\eta^2 = .785$, with more errors at the larger display size in target present trials only.

Number of fixations. Figure 12 shows the average number of fixations for each condition. As before, the number of fixations matched the RT data closely (see Appendix for the statistical analysis).

Fixation duration. Figure 13 shows the average fixation duration for each condition. Fixation durations broadly stayed between 200 and 250 ms, irrespective of the size of the window. However, a 3 × 3 × 2 × 2 × 2 (difficulty × window size × speed × display size × target) Greenhouse-Geisser corrected mixed design ANOVA on the fixation durations yielded a number of significant two way interactions, including interactions between display size and target, $F(1, 44) = 16.7$, $p < .001$, $\eta^2 = .298$, and between speed and target, $F(1, 44) = 33.3$, $p < .001$, $\eta^2 = .431$. Because of these interactions, separate 3 × 3 × 2 × 2 (difficulty × window size × speed × display size) ANOVAs were carried out for the target absent and target present conditions.

For target-absent, a significant interaction between speed, display size, and difficulty, $F(2, 44) = 5.0$, $p < .02$, $\eta^2 = .158$, meant that the analysis was further split by difficulty. In the easy condition, a significant interaction between window size and speed, $F(2, 30) = 4.25$, $p < .04$, $\eta^2 = .221$, was found, with little difference between static and moving trials for the small window, but longer fixations for moving trials than static for the medium and large window. There was also a significant interaction between speed and display size, $F(1, 15) = 34.6$, $p < .001$, $\eta^2 = .697$, with a bigger difference in fixation duration between static and moving at the larger display size. Similar relationships were observed in the medium difficulty target-absent condition, with significant interactions between window size and speed, $F(2, 28) = 16.0$, $p < .001$, $\eta^2 = .533$, and speed and display size, $F(1, 14) = 12.0$, $p < .01$, $\eta^2 = .463$. There was also a significant interaction between window size and display size in the medium difficulty target-

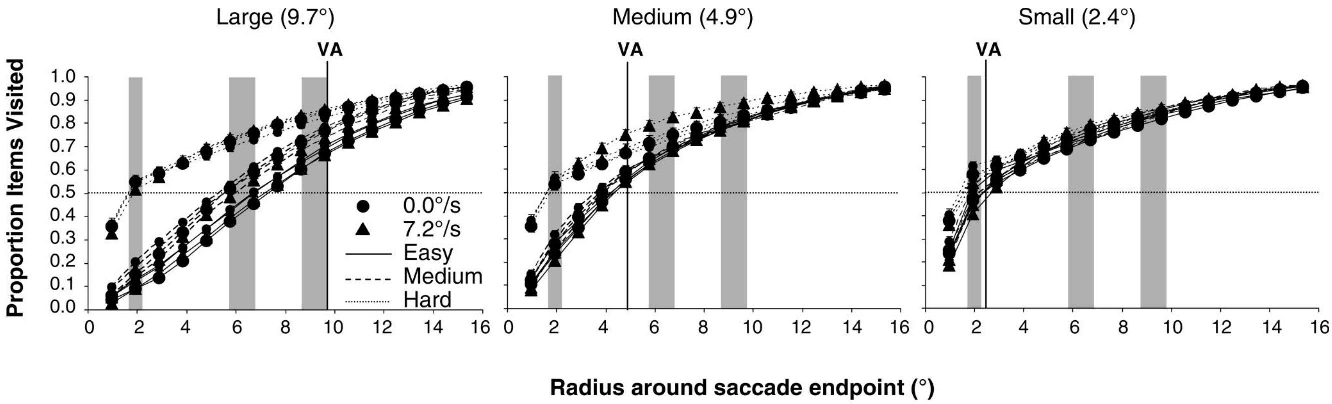

Figure 15. Average proportion of unique items counted in a target-present trial as a function of the radius of the circle within which items were counted with a large, medium, or small window (left to right). Solid, dashed, and dotted lines show easy, medium, and difficult tasks, respectively. Circles are static item and triangles are moving item trials (7.2°/sec). Large and small symbols represent 18 and 12 item displays, respectively. Gray bars represent the radii at which the easy, medium, and difficult tasks reached 50% in Experiment 1 (i.e. estimated FVF's). Lines marked VA show the window radius.

absent condition, $F(2, 28) = 3.4$, $p < .05$, $\eta^2 = .194$, with an effect of display size only for the medium window. In the difficult target-absent condition, a significant interaction between window size and speed was found, $F(2, 30) = 17.9$, $p < .001$, $\eta^2 = .544$, with longer fixations for static trials than for moving and the length of fixation for static trials increasing more than moving trials as the window becomes smaller.

In the target-present condition significant interactions between speed and difficulty, $F(2, 44) = 20.8$, $p < .001$, $\eta^2 = .486$, window size and difficulty, $F(4, 88) = 3.2$, $p < .02$, $\eta^2 = .126$, and window size and speed, $F(2, 88) = 4.6$, $p < .02$, $\eta^2 = .095$, meant that this analysis was also split by difficulty. In the easy and medium target-present condition there were no significant main effects or interactions. In the difficult target-present condition there was a main effect of speed, $F(1, 15) = 51.5$, $p < .001$, $\eta^2 = .774$, with longer fixations for static trials than moving, and an interaction between window size and display size, $F(2, 30) = 3.7$, $p < .05$, $\eta^2 = .197$ with longer fixations for display size 18 when the window was small, but longer fixations for display size 12 when the window was medium or large.

Drift. Figure 14 shows the drift in each condition. Most important, when the window size decreased, drift increased when the items moved in the easy and medium difficulty condition. To a much lesser extent, this was also the case for difficult search. This follows from a 3 × 3 × 2 × 2 × 2 (difficulty × window size × speed × display size × target) of the Greenhouse-Geisser corrected mixed design ANOVA on average drift which yielded three-way interactions between difficulty, window size and speed $F(4, 88) = 10.9$, $p < .001$, $\eta^2 = .332$ and between window size, speed and present $F(4, 88) = 3.4$, $p < .05$, $\eta^2 = .017$. The four-way interaction between difficulty, window size, speed, and target, $F(4, 88) = 2.4$, $p < .08$, $\eta^2 = .10$ also came close to significance. Because of these interactions, the analysis was split by difficulty.

In the easy condition, there was a significant three-way interaction between window size, speed, and target $F(2, 30) = 4.3$, $p < .04$, $\eta^2 = .222$. Because of this interaction, and because drift clearly differs between static and moving conditions, the analysis was further split by speed. For the moving item condition, there were significant main effects of window size $F(2, 30) = 22.4$, $p < .001$ $\eta^2 = .599$ (drift increases when the window becomes smaller) and display size $F(1, 15) = 5.4$, $p < .035$ $\eta^2 = .266$ (more drift for the smallest display size). The interaction between window size and target $F(2, 30) = 3.2$, $p < .06$ $\eta^2 = .177$ was nearly significant (due to the largest window size, where there is less drift for the absent trials). For the static condition, there were no significant effects (all *p*s > .10)

In the medium difficulty condition, there were significant interactions between window size and display size, $F(2, 28) = 3.9$, $p < .04$, $\eta^2 = .216$, and between window size and speed, $F(2, 28) = 47.4$, $p < .001$, $\eta^2 = .772$. As before, the analysis was further split by speed. In the medium difficulty moving item condition there was a main effect of window size, $F(2, 28) = 6.9$, $p < .02$, $\eta^2 = .331$, and a main effect of display size, $F(1, 14) = 10.2$, $p < .01$, $\eta^2 = .421$, with more drift for smaller windows and less drift at larger display sizes. In the medium difficulty static item condition there was a main effect of target only, $F(1, 14) = 6.7$, $p < .03$, $\eta^2 = .323$, with slightly more drift for absent trials.

For the difficult search, there was a significant interaction between window size and speed, $F(2, 30) = 3.6$, $p < .05$, $\eta^2 = .192$ and between speed and display size $F(1, 15) = 31.9$, $p < .001$, $\eta^2 = .680$ so the analysis was again split by item speed. In the moving item condition, there was a significant main effect of window size, $F(2, 30) = 5.0$, $p = .02$, $\eta^2 = .251$, and display size, $F(1, 15) = 67.7$, $p < .001$, $\eta^2 = .819$, with more drift when the window was smaller and when display size was smaller. In the difficult static item condition, there were no significant effects (all *p*s > .14).

Functional visual field. Figure 15 shows the proportion of items visited across the course of an average target present trial at each given radius around the end of each saccade for the large, medium, and small windows. In the large window condition, FVFs for the medium difficulty task were similar to those in Experiment 1, with 50% of items visited between 5.8° and 6.8° for the medium task. For the easy task, the size was estimated between 6.7° and 7.7°. Although this is larger than for the medium difficulty task, it is somewhat smaller than the estimate from Experiment 1 (8.7–9.7°). For the difficult task FVFs in Experiment 1 were 1.7 to 2.2°. In this experiment they were similar, but slightly smaller, at 1.2 to 1.7°. Figure 15 clearly shows that the FVF for both the easy and medium difficulty task shrinks as the visible window is reduced. For the medium window (4.9° radius) the FVF for both the easy and medium difficulty search task was reduced to 3.9 to 4.8°. For the small window (2.4° radius) the FVF for both was further reduced to 1.9 to 2.9°. However, FVFs remain fairly constant for the difficult task at around 1.2 to 1.7° irrespective of window size.

Discussion

Our assumption of larger FVFs for easy and medium search than for difficult search is supported by the results from our gaze-contingent window experiment. When the visible window became 4.9°, there was an immediate effect on RTs, search slopes, and error rates for easy and medium difficulty search, whereas there were no such effects for difficult search. When the gaze contingent window became 2.4° and only one item could be processed within a fixation, the robustness of both easy and medium difficulty search against item motion broke down, with more errors for moving item trials. This suggests that it is not the difficulty of the task per se that renders search vulnerable to item motion (since the difference between a T and an L or between / and | does not depend on the size of the gaze-contingent window), but the fact that only one item can be processed per fixation when it is hard to distinguish between target and distractors.

The finding that the FVF, as computed from the eye movement data, did not exceed the gaze-contingent window size confirms that our approach to determining the FVF in Experiment 1 yields more than a mathematical artifact, and captures an essential element of the visual search process. It also provides an argument against the suggestion that our method of computing the FVF will underestimate the number of items processed in the moving condition. One consequence of our method is that items that later move into the radius around the saccadic endpoint will be ignored. However, an underestimation of the number of items processed should result in an overestimation of the size of the FVF (a larger radius around the saccadic endpoint is needed to reach 50% visited items). Given the close match between the FVFs and the smallest gaze-contingent window, it would seem that any underestimation of the number of items is only minimal.

The data for the moving items in the easy and medium difficulty conditions shed an interesting light on the control of fixation duration during search. On the one hand, the drift increased with decreasing window size. This would seem to be caused by the reduced size of the FVF, which led to search being restricted to fewer items. On the other, there was no sign of a shortening in the duration of fixations. So, although it is possible to process multiple items per fixation when searching for a T among Ls or for a / among |, there seems to be no scope for a reduction in the fixation duration when only a single item has to be processed. This chimes with the observation by Hooge and Erkelens (1996) that there is only a limited relation between fixation duration and target discriminability.

It is interesting to note that the error rates remain highest for the difficult search task, even at the smallest window size. This indicates that there are two sources of miss errors in visual search. The first kind of error is that the target does not fall within the FVF, the second that the target does fall within the FVF but is not recognized as such. Almost all errors in the easy and medium difficulty task would seem to be of the first kind, whereas the difficult task also yields errors of the second kind. This would also provide an answer to the question of what it is that determines whether a search task is easy or difficult: the discriminability of the target.

General Discussion

In both Experiment 1 and 2, we have shown that there is very little difference in eye movement behavior between search among static or among moving items. Fixation counts, fixation durations, and saccade amplitudes vary little between static and moving item search, as long as the discrimination between target and distractor does not require inspection of individual items. The eye movement data certainly do not give any indication that there are different search strategies for static and for moving displays. On the contrary, the data is much more consistent with a single search process being applied to both kinds of display. Any theory of visual search must therefore account for the robustness of visual search against item motion in easier search and the limits to this robustness in very difficult search.

Our proposed theoretical framework, an eye-movement-based system, with parallel processing within fixations and serial processing between fixations would be able to do this. (Although our results are probably also consistent with models that assume rapid serial processing within a fixation). We suggested that an important property of this system has to be an FVF whose size depends on the difficulty of the task. Both Experiment 1 and Experiment 2 have yielded support for this contention. In Experiment 1, analysis of the FVFs revealed that their estimated size indeed decreases as the difficulty of the search increases. Experiment 2 provided important empirical support for the validity of these estimates. There was an increase in RTs for both easy and medium difficulty search when the size of the gaze-contingent window became smaller than the estimated FVF from Experiment 1. Moreover, the FVFs estimated from the eye movement data in Experiment 2 closely matched the area of the search display visible through the gaze contingent window.

Our framework explains robustness against item motion as the product of processing several items in parallel within a fixation, thereby limiting the effect of item movement in two ways. First, it affords a divide-and-conquer approach. Rather than processing the entire display in a single step, parts of the display are processed sequentially. This reduces the complexity of the problem posed by the motion of the items. Second, since items are processed in parallel, items that are processed more than once have little or no additional cost. The assumption of parallel processing within a fixation is supported by the results of Thornton and Gilden (2007) who reported that, with a few exceptions, most searches (including T among L) are best described with a parallel model.

In the case of very difficult search, the FVF is limited to only one or two items, so robustness breaks down both because parallel processing does not offer any protection against reprocessing of items and because revisits become more likely due to the many fixations needed to search the display. According to our account, the breakdown in robustness against item motion is not the product of the task difficulty per se, but rather the result of the fact that within each fixation only a few items are processed. Search for a T among Ls is typically robust against item motion (Hulleman, 2009, 2010) However, in Experiment 2 it was shown that this robustness disappeared when only one or two items were visible per fixation. Many more errors were made in moving item search than static. This provides prima facie evidence that it is the number of items that can be processed within a fixation that determines whether search is robust against item motion.

The Role of Eye Movements

Whereas eye movements are integral to our proposed framework, most theories of visual search typically have not taken the overt attentional shifts involved in eye movements into account (Treisman & Gelade, 1980; Wolfe, 1994). The main argument for disregarding eye movements has come from evidence that visual search can be carried out without eye movements and yields similar search latencies. But closer inspection of this evidence suggests that this might have been the consequence of the size of the search items and the difficulty of the task. For instance, Klein and Farrell (1989) had participants carry out serial or parallel search of a symmetrical array arranged in a circle, asking participants to avoid eye movements. Trials where eye movements occurred were removed from the analysis. They found that there was no effect of restricting eye movements for the parallel task, while for the serial task there was only a small, albeit significant, increase in error rates. However, the radius of their circular search array was quite small, subtending a visual angle of only 2.4°, while the radius of the circular search items was .56°. Additionally, the maximum display size was 10 and the serial task was relatively easy. This suggests that the FVF might have been large enough to perform the task adequately without eye movements. Similarly, Zelinsky and Sheinberg (1997) reported that visual search was actually more efficient when eye movements were restricted. Even though in this case the search display had a maximum radius of 6°, the items themselves were quite large, with a radius of .66°. So, again, the FVF might have been sufficiently large. Moreover, despite the fact that participants might be able to search without eye movements, Zelinsky and Sheinberg (1997) also demonstrated that even with large stimuli participants did make eye movements when free to do so. In fact, the number of saccades accounted for up to 67% of the variation in RTs. Other evidence for the importance of eye movements in visual search comes from Scialfa and Joffe (1998). They used a more difficult search task and found that performance did worsen when eye movements were restricted and that this effect increased with increasing eccentricity of the target.

It seems, therefore, that eye movements are actually an integral part of the visual search process (cf. Zelinsky, 2008). Only when the search items are rather large, the search display is relatively small, and the search task is not too difficult does restriction of eye movements not impede performance.

Inhibitory Tagging of Individual Items

Klein (1988) was the first to suggest a role for IOR in serial visual search, with the locations of items that have been inspected and rejected as the target being tagged to prevent reinspection. Evidence to support this theory has come predominantly from probe detection studies (for a review see Wang & Klein, 2010). In probe detection studies a visual search task is followed by a detection task, with a probe presented either on a nontarget item or at an empty location. A RT cost is associated with probe presentation on a nontarget item relative to a probe on an empty location, with a greater cost for serial search than parallel. The interpretation of these results is that detection of the probe is delayed by inhibition of return (IOR) applied to location of the item, because it has been inspected and rejected as the target. Other evidence for IOR comes from eye movement studies that have shown that the number of refixations of nontarget items is lower than would be expected due to chance (Gilchrist & Harvey, 2000; Peterson et al., 2001), although there is some debate over whether avoidance of refixations may be explained by scan path planning, rather than inhibition of return (Peterson et al., 2001).

While the inhibitory tagging was location based under Klein's (1988) original proposal, evidence suggests that IOR can be both location based and object based (Gibson & Egeth, 1994; Jordan & Tipper, 1998; Tipper, Weaver, Jerreat, & Burak, 1994). After some failures to replicate Klein's (1988) original result (e.g., Wolfe & Pokorny, 1990), several authors have proposed that the inhibitory tagging is actually object-based (Müller & von Mühlenen, 2000). Moreover, Ogawa, Takeda, and Yagi (2002) found a probe effect in visual search with moving items, using stimuli almost identical to the difficult condition used in our experiments.

Ogawa et al. (2002) took this as evidence that the inhibitory tag can travel with the item. But we would suggest that although Ogawa et al. (2002) might have found evidence for inhibition, it might simply have been due to a recent refixation of the item, rather than to a traveling inhibitory tag. Ogawa et al. (2002) used four and eight items in their experiment, and from the difficult condition in Experiment 1 it becomes clear that the number of revisits is already quite substantial for both static and moving items even when there are only six items in the search display. This would seem to go against Ogawa et al.'s (2002) suggestion that inhibitory tagging operates on at least eight items.

We would contend that what has been labeled as inhibitory tagging of individual items is actually the avoidance of parts of the display to which eye movements already have been made (a suggestion also made by Beck, Peterson &Vomela, 2006, who reported search biases away from inspected items based on their locations). But, due to the difficulty of the search task, the FVF has shrunk to the size of a single item.

Under our theoretical framework, this is a special case of a more general process that attempts to process as many items as possible in a single fixation, while avoiding previously fixated areas. In this context, it is interesting to note that some of the most robust evidence for item-based IOR comes from studies where there was item-by-item inspection, either due to the difficulty of the task (Müller & von Mühlenen, 2000; Ogawa et al., 2002), or due to the nature of the presentation of the display (Snyder & Kingstone, 2007).

Experiment 1 also gives an insight into the number of fixations that can be used to avoid revisits: memory for fixations seems severely limited and is probably smaller than 6. Especially given that searches were terminated before all items were fixated. This would be close to the estimate of 3–4 that McCarley et al. (2003) gave.

Termination of Search

When the estimate of the FVF is combined with the location of fixation and the location of the items in the search display, it becomes clear that participants terminate their search before they have processed all items. As already pointed out by Chun and Wolfe (1996), this suggests that the "absent" response contains a strategic component. It is interesting to note that for both medium and difficult search and irrespective of display size, the cut-off point seems to lie around 85–90% of items processed. The presence of a strategic component might also explain why there is sometimes a small effect of motion in absent trials. Participants might convince themselves that the moving item condition should be more difficult than the static item condition. As a result they might be more reluctant to terminate their search. Under this scenario the effect of motion on absent trials would represent this conservative shift in the stopping rule, rather than any perceptual effect of motion itself.

It could be argued that this reluctance should result in an increase of the number of fixations for the motion conditions (which only happened during difficult search). However, it also possible that motion decreased the signal-to-noise ratio for the moving item displays. Being surrounded by more distractors makes it easier to detect the target (Duncan & Humphreys, 1989), but this effect will be diminished when the items are moving. This by itself might lead to longer fixation durations for moving, especially for larger display sizes (which happened during easy search). When the FVF became smaller for medium and difficult search, this effect of motion on fixation duration disappeared.

The Nature of Serial Search

In the original version of guided search Wolfe, Cave and Franzel (1989) suggested that search for a T among Ls is serial, since none of the feature maps would be able to guide attention toward the target. In the latest incarnation of the Guided Search Model (e.g., Wolfe, Horowitz, Palmer, Michod, & Van Wert, 2010), this type of search is still classified as unguided. Similarly, Zelinsky (2008) modeled search for an O among Qs as a purely serial process, even though Treisman and Souther (1985) reported search slopes of 25.2–33.5 ms for target-present trials. From the FVFs observed in our experiments, it would seem that there is still a large parallel component in both these types of searches. This chimes with Wolfe, Palmer, and Horowitz (2010), who argued that the RT distributions of searches with present slopes of around 40 ms/item do not conform to a purely serial model.

Our theoretical framework is better capable of encompassing this kind of result, because it proposes that visual search is a continuum that stretches from mostly parallel with some seriality for very easy search to almost completely serial for very difficult search. Given that eye movements were made even in the orientation search task, it would seem fair to suggest that there might not be purely parallel search. However, if the search task is made hard enough, it is possible to make search completely serial. It is important to note, however, that this only happens for searches with present slopes around 125 ms/item, where every item has to be fixated individually. In that sense, the qualitative distinction between feature search and conjunction search that Treisman and Gelade (1980) made, and which has proved hugely influential, might have to be reconsidered (something that was already observed by Wolfe (1998), who plotted the frequency of the search slopes measured in his lab and found the distribution to be unimodal). According to our theoretical framework, the real qualitative change actually occurs at the other end of the search difficulty spectrum, at the point where search becomes so difficult that the last vestiges of parallel processing are lost, and every individual item has to be fixated to find the target.

References

Beck, M. R., Peterson, M. S., & Vomela, M. (2006). Memory for where, but not what, is used during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 235–250. doi:10.1037/0096-1523.32.2.235

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30,* 39–78. doi:10.1006/cogp.1996.0002

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458. doi:10.1037/0033-295X.96.3.433

Eckstein, M. P., Thomas, J. P., Palmer, J., & Shimozaki, S. S. (2000). A signal detection model predicts effects of set size on visual search accuracy for feature, conjunction, triple conjunction and disjunction displays. *Perception & Psychophysics, 62,* 425–451. doi:10.3758/BF03212096

Geisler, W. S., & Chou, K. (1995). Separation of low-level and high-level factors in complex tasks: Visual search. *Psychological Review, 102,* 356–378. doi:10.1037/0033-295X.102.2.356

Geyer, T., von Mühlenen, A., & Müller, H. J. (2007). What do eye movements reveal about the role of memory in visual search? *The Quarterly Journal of Experimental Psychology, 60*(7), 924–935. doi:10.1080/17470210600831119

Gibson, B. S., & Egeth, H. (1994). Inhibition of return to object-based and environment-based locations. *Perception & Psychophysics, 55,* 323–339. doi:10.3758/BF03207603

Gilchrist, I. D., & Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology, 10,* 1209–1212. doi:10.1016/S0960-9822(00)00729-6

Heinke, D., & Humphreys, G. W. (2003). Attention, spatial representation, and visual neglect: Simulating emergent attention and spatial memory in the selective attention for identification model (SAIM). *Psychological Review, 110,* 29–87. doi:10.1037/0033-295X.110.1.29

Hooge, I. T. C., & Erkelens, C. J. (1996). Control of fixation duration in a simple search task. *Perception & Psychophysics, 58,* 969–976. doi:10.3758/BF03206825

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394,* 575–577. doi:10.1038/29068

Hulleman, J. (2009). No need for inhibitory tagging of locations in visual search. *Psychonomic Bulletin & Review, 16,* 116–120. doi:10.3758/PBR.16.1.116

Hulleman, J. (2010). Inhibitory tagging in visual search: Only in difficult search are items tagged individually. *Vision Research, 50,* 2069–2079. doi:10.1016/j.visres.2010.07.017

Intriligator, J., & Cavanagh, P. (2001). The spatial resolution of visual attention. *Cognitive Psychology, 43,* 171–216. doi:10.1006/cogp.2001.0755

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489–1506. doi:10.1016/S0042-6989(99)00163-7

Jordan, H., & Tipper, S. P. (1998). Object-based inhibition of return in static displays. *Psychonomic Bulletin & Review, 5,* 504–509. doi:10.3758/BF03208829

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334,* 430–431. doi:10.1038/334430a0

Klein, R., & Farrell, M. (1989). Search performance without eye-movements. *Perception & Psychophysics, 46,* 476–482. doi:10.3758/BF03210863

McCarley, J. S., Wang, R. X. F., Kramer, A. F., Irwin, D. E., & Peterson, M. S. (2003). How much memory does oculomotor search have? *Psychological Science, 14,* 422–426. doi:10.1111/1467-9280.01457

Motter, B. C., & Holsapple, J. (2007). Saccades and covert shifts of attention during active visual search: Spatial distributions, memory, and items per fixation. *Vision Research, 47,* 1261–1281. doi:10.1016/j.visres.2007.02.006

Motter, B. C., & Simoni, D. A. (2008). Changes in the functional visual field during search with and without eye movements. *Vision Research, 48,* 2382–2393. doi:10.1016/j.visres.2008.07.020

Müller, H. J., & von Mühlenen, A. (2000). Probing distracter inhibition in visual search: Inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1591–1605. doi:10.1037/0096-1523.26.5.1591

Ogawa, H., Takeda, Y., & Yagi, A. (2002). Inhibitory tagging on randomly moving objects. *Psychological Science, 13,* 125–129. doi:10.1111/1467-9280.00423

Over, E. A. B., Hooge, I. T. C., Vlaskamp, B. N. S., & Erkelens, C. J. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research, 47,* 2272–2280. doi:10.1016/j.visres.2007.05.002

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

Peterson, M. S., Kramer, A. F., Wang, R. X. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12,* 287–292. doi:10.1111/1467-9280.00353

Scialfa, 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. doi:10.3758/BF03211940

Snyder, J. J., & Kingstone, A. (2007). Inhibition of return at multiple locations and itsimpact on visual search. *Visual Cognition, 15,* 238–256. doi:10.1080/13506280600724892

Thornton, T. L., & Gilden, D. (2007). Parallel and serial processes in visual search. *Psychological Review, 114,* 71–103. doi:10.1037/0033-295X.114.1.71

Tipper, S. P., Weaver, B., Jerreat, L. M., & Burak, A. L. (1994). Object-based and environment-based inhibition of return of visual-attention. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 478–499. doi:10.1037/0096-1523.20.3.478

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

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114,* 285–310. doi:10.1037/0096-3445.114.3.285

von Mühlenen, A., Müller, H. J., & Müller, D. (2003). Sit-and-wait strategies in dynamic visual search. *Psychological Science, 14,* 309–314. doi:10.1111/1467-9280.14441

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research, 50,* 220–228. doi:10.1016/j.visres.2009.11.013

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238. doi:10.3758/BF03200774

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9,* 33–39. doi:10.1111/1467-9280.00006

Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in Cognitive Sciences, 7,* 70–76. doi:10.1016/S1364-6613(02)00024-4

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the Feature Integration Model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433.

Wolfe, J. M., Horowitz, T. S., Palmer, E. M., Michod, K. O., & Van Wert, M. J. (2010). Getting into guided search.In V. Coltheart (Ed.), *Tutorials in visual cognition* (pp. 93–119). Hove, UK: Psychology Press.

Wolfe, J. M., Palmer, E. M., & Horowitz, T. S. (2010). Reaction time distributions constrain models of visual search. *Vision Research, 50,* 1304–1311. doi:10.1016/j.visres.2009.11.002

Wolfe, J. M., & Pokorny, C. W. (1990). Inhibitory tagging in visual search: A failure to replicate. *Perception & Psychophysics, 48,* 357–362. doi:10.3758/BF03206686

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115,* 787–835. doi:10.1037/a0013118

Zelinsky, G. J., & Sheinberg, D. (1997). Eye movements during parallel-serial visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 244–262. doi:10.1037/0096-1523.23.1.244

(*Appendix follows*)

Appendix

Details of Analyses

Experiment 1

Number of Fixations

A 3 × 2 × 3 × 2 (difficulty × speed × display size × target) Greenhouse-Geisser corrected within subjects ANOVA on number of fixations yielded a significant three-way interactions between difficulty, speed, and target, $F(2, 24) = 7.1$, $p < .03$, $\eta^2 = .372$, and difficulty, display size and target, $F(4, 48) = 33.6$, $p < .001$, $\eta^2 = .737$. As before, the analysis was split by difficulty.

A 2 × 3 × 2 (speed × display size × target) ANOVA on easy trials revealed a significant interaction between display size and target, $F(2, 24) = 6.5$, $p < .02$, $\eta^2 = .352$. In separate ANOVAs for absent and present trials there was a significant main effect of display size when the target was absent, $F(2, 24) = 6.6$, $p < .01$, $\eta^2 = .354$, with more fixations for larger display sizes, but no main effect of display size when the target was present. For the medium task, a similar ANOVA revealed significant interactions between speed and target, $F(1, 12) = 5.0$, $p < .05$, $\eta^2 = .294$, and between display size and target, $F(2, 24) = 21.0$, $p < .001$, $\eta^2 = .637$. When the analysis was split by target, there was a significant main effect of display size for absent trials, $F(2, 24) = 30.5$, $p < .001$, $\eta^2 = .718$, and present trials, $F(2, 24) = 31.3$, $p < .001$, $\eta^2 = .723$, but no effect of speed in either case. In the difficult condition there were significant interactions between speed and target, $F(1, 12) = 8.1$, $p < .02$, $\eta^2 = .403$, and between display size and target, $F(2, 24) = 48.8$, $p < .001$, $\eta^2 = .802$. Separate ANOVAs for absent and present trials revealed main effects of speed, $F(1, 12) = 10.7$, $p < .01$, $\eta^2 = .471$, and display size, $F(2, 24) = 73.0$, $p < .001$, $\eta^2 = .859$, when the target was absent, but when the target was present, there was only a main effect of display size $F(2, 24) = 53.0$, $p < .001$, $\eta^2 = .815$.

Saccadic Amplitudes

A 3 × 2 × 3 × 2 (difficulty × speed × display size × target) Greenhouse-Geisser corrected within subjects ANOVA on average saccade amplitude revealed a three-way interaction between difficulty, speed and display size, $F(4, 48) = 3.7$, $p < .02$, $\eta^2 = .234$, and difficulty, display size, and target, $F(4, 48) = 7.6$, $p < .01$, $\eta^2 = .387$. As previously, the analysis was split along the difficulty dimension.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on easy trials yielded significant interactions between speed and display size, $F(2, 24) = 3.6$, $p < .05$, $\eta^2 = .229$, and between display size and target, $F(2, 24) = 6.3$, $p < .01$, $\eta^2 = .343$. In the absent trials there were significant main effects of speed, $F(1, 12) = 8.9$, $p < .02$, $\eta^2 = .425$, and display size, $F(2, 24) = 14.6$, $p < .001$, $\eta^2 = .548$, with slightly longer saccades for moving items and increased saccade length for larger display sizes. In easy present trials there were no main effects or interactions.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on medium difficulty trials yielded a significant interaction between speed and target, $F(1, 12) = 9.1$, $p < .02$, $\eta^2 = .431$. For the absent trials there were no significant main effects and no interaction, while for present trials there was a significant main effect of speed only, $F(1, 12) = 20.1$, $p < .001$, $\eta^2 = .627$, with slightly longer saccades for static items than for moving items.

A 2 × 3 × 2 (speed × display size × target) within subjects ANOVA on difficult trials revealed significant interactions between speed and target, $F(1, 12) = 23.1$, $p < .001$, $\eta^2 = .658$, and between display size and target, $F(2, 24) = 38.9$, $p < .001$, $\eta^2 = .764$. For target absent trials there was a main effect of display size only, $F(2, 24) = 126$, $p < .001$, $\eta^2 = .913$, with saccade amplitude decreasing with increasing display size. For target present trials there was a main effect of speed, $F(1, 12) = 61.4$, $p < .001$, $\eta^2 = .837$, and display size, $F(2, 24) = 46.1$, $p < .001$, $\eta^2 = .793$, with slightly longer saccades for static items and saccade amplitude decreasing with increasing display size.

Experiment 2

Number of Fixations

A 3 × 3 × 2 × 2 × 2 (difficulty × window size × speed × display size × target) Greenhouse-Geisser corrected mixed design ANOVA on the number of fixations yielded significant three-way interactions between difficulty, window size, and target $F(4, 88) = 27.5$, $p < .001$, $\eta^2 = .556$, difficulty, speed and target $F(2, 44) = 9.3$, $p < .001$, $\eta^2 = .296$, difficulty, window size, and display size $F(4, 88) = 16.0$, $p < .001$, $\eta^2 = .422$, difficulty, display size, and target, $F(2, 44) = 8.2$, $p < .001$, $\eta^2 = .271$, window size, display size, and target, $F(2,88) = 19.5$, $p < .001$, $\eta^2 = .307$, and window size, speed, and target, $F(2,88) = 23.3$, $p < .001$, $\eta^2 = .346$. Because of the many interactions, separate within-subjects ANOVAs were carried out for easy, medium, and difficult search tasks.

A 3 × 2 × 2 × 2 (window size × speed × display size × target) ANOVA for the easy trials revealed significant interactions between window size, speed, and target $F(2,30) = 5.4$, $p < .05$, $\eta^2 = .264$ and window size, display size, and target, $F(2,30) = 15.3$, $p < .001$, $\eta^2 = .505$. Because of these interactions, the analysis of the easy task was split by window size. For each window size there was an interaction between display size and target, $F(1, 15) = 31.2$, $p < .001$, $\eta^2 = .675$, $F(1, 15) = 32.9$, $p < .001$, $\eta^2 = .687$ and $F(1,15) = 37.5$, $p < .001$, $\eta^2 = .714$ (large, medium, and small, respectively), with more fixations for the larger display size in target-absent trials only. For the small window, there was also a significant interaction between speed and target, $F(1,15) = 9.4$, $p < .01$, $\eta^2 = .386$, with more fixations in moving item trials when the target was absent.

(*Appendix continues*)

In an equivalent ANOVA for medium difficulty trials, very similar relationships were observed. Significant interactions between window size, speed and target $F(2, 28) = 16.0$, $p < .001$, $\eta^2 = .534$ and window size, display size, and target, $F(2, 28) = 9.4$, $p < .01$, $\eta^2 = .403$ led to the analysis of the medium difficulty trials being split by window size. For each window size there was an interaction between display size and target, $F(1, 14) = 15.7$, $p < .01$, $\eta^2 = .529$, $F(1, 14) = 25.4$, $p < .001$, $\eta^2 = .644$ and $F(1, 14) = 45.3$, $p < .001$, $\eta^2 = .764$ (large, medium, and small, respectively), with more fixations for the larger display size in target -absent trials only. For the small window, there was also a significant interaction between speed and target, $F(1, 14) = 19.9$, $p < .001$, $\eta^2 = .587$, with more fixations in moving item trials when the target was absent.

For the difficult task, there was no main effect of window size, but a significant interaction between window size, speed, and target, $F(2, 30) = 6.8$, $p < .01$, $\eta^2 = .311$. Because of this interaction the analysis for difficult trials was split by target. There was no main effect of window size for absent trials. However, there was a significant interaction between window size and speed, $F(2, 30) = 5.4$, $p < .02$, $\eta^2 = .266$ with more fixations for moving trials than for static overall, but more fixations for moving trials with a small window relative to a large window and no effect of window size on static trials. There was also an interaction between window size and display size, $F(2, 30) = 5.9$, $p < .02$, $\eta^2 = .283$, with the number of fixations increasing more with decreasing window size at the display size 18.

In present trials there were no significant interactions, but there were significant main effects of window size, $F(2, 30) = 3.7$, $p < .04$, $\eta^2 = .199$, speed, $F(1, 15) = 38.5$, $p < .001$, $\eta^2 = .720$, and display size, $F(1, 15) = 203.7$, $p < .001$, $\eta^2 = .931$, with more fixations for with the smaller windows, moving items, and the larger display size.

Received July 19, 2011
Revision received April 3, 2012
Accepted April 5, 2012 ■