One visual search, many memory searches: An eye-tracking investigation of hybrid search

One visual search, many memory searches: An eye-tracking investigation of hybrid search

Trafton Drew University of Utah, Salt Lake City, UT, USA,

Sage E. P. Boettcher University of Oxford, Oxford, UK

Jeremy M. Wolfe Brigham and Women's Hospital, Boston, MA, USA
Harvard Medical School, Boston, MA, USA

Suppose you go to the supermarket with a shopping list of 10 items held in memory. Your shopping expedition can be seen as a combination of visual search and memory search. This is known as "hybrid search." There is a growing interest in understanding how hybrid search tasks are accomplished. We used eye tracking to examine how manipulating the number of possible targets (the memory set size [MSS]) changes how observers (Os) search. We found that dwell time on each distractor increased with MSS, suggesting a memory search was being executed each time a new distractor was fixated. Meanwhile, although the rate of refixation increased with MSS, it was not nearly enough to suggest a strategy that involves repeatedly searching visual space for subgroups of the target set. These data provide a clear demonstration that hybrid search tasks are carried out via a "one visual search, many memory searches" heuristic in which Os examine items in the visual array once with a very low rate of refixations. For each item selected, Os activate a memory search that produces logarithmic response time increases with increased MSS. Furthermore, the percentage of distractors fixated was strongly modulated by the MSS: More items in the MSS led to a higher percentage of fixated distractors. Searching for more potential targets appears to significantly alter how Os approach the task, ultimately resulting in more eye movements and longer response times.

Introduction

From searching for the snooze button when the alarm rings to a midnight hunt for a snack, we engage in hundreds of visual search tasks each day. A vast literature devoted to the visual search paradigm has yielded a deep understanding of how we find targets or decide that the target we are searching for is not present (Duncan & Humphreys, 1989; Najemnik & Geisler, 2005; Wolfe, Cave, & Franzel, 1989; Zelinsky, 2008). The majority of this literature has focused on situations with a single, well-defined target: for instance, a T among Ls or the snooze button in the example above.

We know a great deal less about how search is accomplished when the observer (O) is searching for one of a number of possible targets. Under these circumstances, the O must search through both the visual scene and his or her memory of the potential targets in order to determine whether a target is present. We refer to this as "hybrid search," a term originally coined by Shiffrin and colleagues (Cunningham & Wolfe, 2014; Drew & Wolfe, 2014; Shiffrin & Schneider, 1977; Wolfe, 2012; Wolfe, Boettcher, Josephs, Cunningham, & Drew, 2015). For example, if we were looking exclusively for cookies during a late-night search for a snack, this would be considered a "simple visual search" with only one target. On the other hand, a search for peanut butter, jelly, and bread would be a "hybrid search" in which we must hold several targets in memory. Hybrid search is an important task for expert searchers. Consider radiologists, who need to determine whether signs of any of a number of ailments are present in a medical image, or baggage screeners, who must continuously search for many different possible threats that might be smuggled aboard an airplane. The present work uses eye tracking to examine the interaction of visual and memory search in a hybrid search task.

Previous work has shown, as might be expected, that increasing the number of possible targets results in slower response times (RTs), less efficient search, and slightly elevated error rates (Wolfe, 2012). More interestingly, that work has shown that RT increases linearly with increases in the visual set size (VSS), but RT increases linearly with the *log* of the number of possible targets held in memory (memory set size [MSS]). As a result, Os

Citation: Drew, T., Boettcher, S. E. P., & Wolfe, J. M. (2017). One visual search, many memory searches: An eye-tracking investigation of hybrid search. *Journal of Vision*, *17*(11):5, 1–10, doi:10.1167/17.11.5.

doi: 10.1167/17.11.5 Received February 19, 2017; published September 11, 2017 ISSN 1534-7362

were capable of searching for any of 100 possible targets much faster than would have been predicted by a linear increase of RT with MSS. More concretely, if memory search, like visual search, moved at a rate of ~25–50 ms/item, it would take on the order of 40 min to search a single photograph of 100 people for the presence of any of 1,000 friends. The logarithmic compression of the memory search allows such a search to be accomplished in seconds.

Several mechanisms could produce the logarithmic relationship of RT to MSS. For instance, consider the child's game of guessing a number between 1 and 100. If you use the strategy of dividing the set in half on each guess ("Is the number bigger than 50? No? Is it bigger than 25?," etc.), the number of steps to reach an answer will be log2 of 100. Similarly, if half of the memory set can be excluded on each step of a memory search, the number of required steps will be log2 of the MSS. Such a process could proceed with Os selecting item after item in the visual display. Alternatively, the O might somehow search the entire visual display in parallel for half the remembered items, then for half of the remaining items in the memory set, and so forth. In either case, this would seem to require several passes through the visual display in order to identify whether any of the possible targets are present.

Leite and Ratcliff (2010) have shown logarithmic increases in RT can be a by-product of a diffusion process. When an item in the visual display is compared to the contents of memory, Leite and Ratcliff envision separate diffusion processes accumulating information about whether the current item matches each of the items in the memory set. The visual item is identified as a match to a specific item in memory if the diffuser reaches the decision threshold for that item. If that threshold is set too low, the item might be misidentified (a false-alarm error). As the MSS and, thus, the number of diffusers rises, the chance of a false-alarm error also rises. To keep the error rate constant, the decision thresholds must increase with the MSS. A higher decision threshold produces a longer RT, and the increase of RT with MSS in such a model turns out to be logarithmic.

As these different models produce similar patterns of RT data, it is difficult for behavioral measures alone to distinguish between them. However, these models predict very different patterns of eye movements. In order to test the predictions of these models, we performed a conceptual replication of Wolfe's (2012) previous work while tracking eye movements. The data from this experiment provide strong evidence that Os tend to search through the memory set while fixating individual targets rather than, for example, searching through part of the memory set on each of several searches through the visual set.

In addition to providing a means by which to evaluate different models of how hybrid search is accomplished, the eye-tracking data from this paradigm provide a method to examine how manipulating the number of potential targets alters the way in which we approach a scene. Previous work has demonstrated that manipulating search difficulty results in a series of reliable changes in how search is accomplished. For instance, Young and Hulleman (2013) found that increasing search task difficulty via increasing target–distractor similarity led to an increased number of fixations, but fixation dwell time and saccadic amplitude were not strongly influenced. The most striking effects of task difficulty were observed in measures of *function visual field* (FVF). The FVF, sometimes also referred to as *perceptual span*, is a popular metric that is proposed to quantify the spatial range over which information can be effectively processed in a single fixation. Young and Hulleman found that FVF generally decreases as the complexity of the task increases (Rayner & Fisher, 1987, 2013; Young & Hulleman, 2013).

In a search task, the FVF is typically calculated by placing virtual circles around each fixation and asking how big the radius of those circles needs to be in order for the circles to cover some percentage of the items in a display. There are multiple constraints on the ability to process items away from the point of fixation. There are basic limits due to the decline in acuity away from the fovea and due to peripheral crowding (Levi, 2008). These are captured by performance measures such as the useful field of view (Ball, Beard, Roenker, Miller, & Griggs, 1988). Attention also plays a role (Williams, 1989). The FVF attempts to measure the effective number of items that can be processed on each fixation as that number varies with the search task. Thus, if Os are searching for a red item among green, all items can be processed at once, and the resulting FVF would be very large. Harder tasks produce smaller FVFs. The FVF is tightly coupled to the number of fixations required in a task. As the number of fixations goes up, FVF typically declines. One exception to this general rule would be instances in which there are many refixations on the same location. Under these circumstances, one could see a decoupling of these two metrics as more fixations could theoretically lead to no change in the FVF estimate. Whether one sees the FVF as causal (Hulleman & Olivers, 2017) or as side effect of RT and fixation (Wolfe, 2017) is a matter of some ideological debate. Nevertheless, because the measure is of current interest, we will report those analyses here.

Materials and methods

Fourteen Os (mean age: 32.6, *SD*: 9.4, six women) were each tested on five blocks, each with a different MSS of one, four, 16, 32, or 100 unique objects. Sample size was chosen to exceed the sample size of 10 from

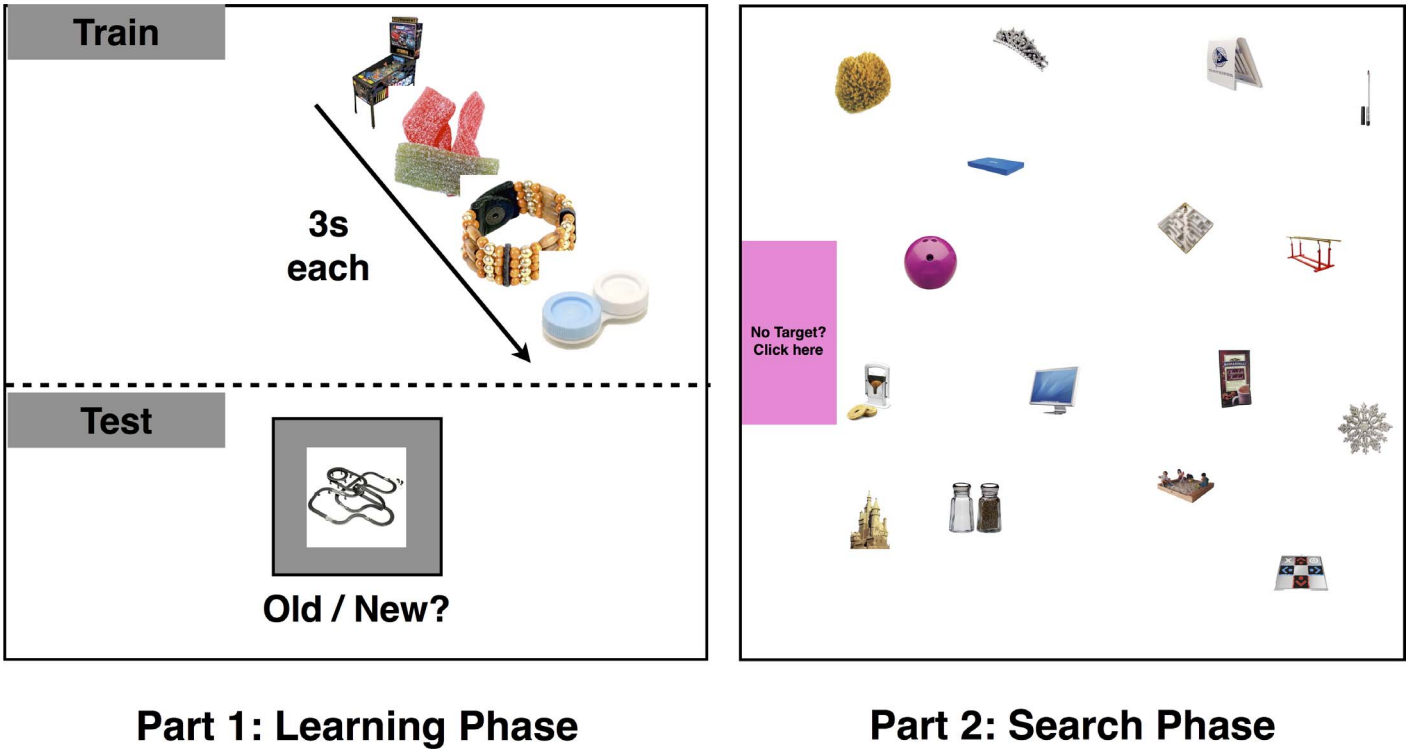

Figure 1. Experimental design and sample stimulus.

Wolfe's (2012) hybrid search experiments. Block order was randomized across Os. Each block of the experiment began with a memorization procedure in which Os were asked to memorize the set of targets. Each target object was presented in isolation for 3 s in the center of the screen (Figure 1a). During the memory test phase (Figure 1b), Os saw a total of 2X objects (where is X is the MSS for the current block). Half of these objects were targets. Os identified each item as being a member of the memory set or a novel item. Os had to perform this recognition test twice with an accuracy of more than 90%. Failing to do so resulted in a repeat presentation of the memory set and another recognition test. Once above threshold, Os could move onto the next, hybrid search portion of the experiment.

On each hybrid search trial, Os searched for any one of their memorized targets in visual displays of eight or 16 items. Targets were present on 50% of trials. Os were asked to localize a target, if present, by using the mouse to click on it directly. If no target was present, they were instructed to click on the "absent" box along the left side of the screen (see Figure 1). Os completed 12 practice trials and 160 experimental trials (640 total) for each of the five memory blocks.

All Os gave informed consent, were compensated $10 per hour, had at least 20/25 acuity with correction, and passed the Ishihara color blindness test. During the experiments, Os sat at a chin rest positioned 60 cm from a 20-in. CRT monitor (Mitsubishi Diamond Pro 91TXM) with resolution set to 1280 × 960 pixels and an 85-Hz refresh rate. Experimental sessions were carried out on a Macintosh G4 computer running Mac OS 10.5. Experiments were written in Matlab 7.5 (The Mathworks) using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997), Version 3. Stimuli were photographs of objects that subtended 2.39°.

Eye-tracking analysis

Eye tracking was carried out using a desktop-mounted Eyelink 1000 (SR Research, ON, Canada) which sampled the x- and y-positions of the eye at 500 Hz. We calibrated the eye tracker using a nine-point calibration procedure. An eye movement was classified as a saccade when its distance exceeded 0.5° and its velocity reached 35 °/s (or acceleration reached 9500 °/s^2). Viewing was binocular, but data from only one eye was recorded.

Eye movements were analyzed by placing 122 × 122 pixel areas of interest around each item in the search array. Based on these regions of interest, we measured fixations, refixations, cumulative dwell time, and the percentage of objects visited as a function of MSS.

FVF analyses

To calculate FVF, we used the gaze coordinates of each fixation and subsequently calculated the percentage of distractors contained within a variable circular

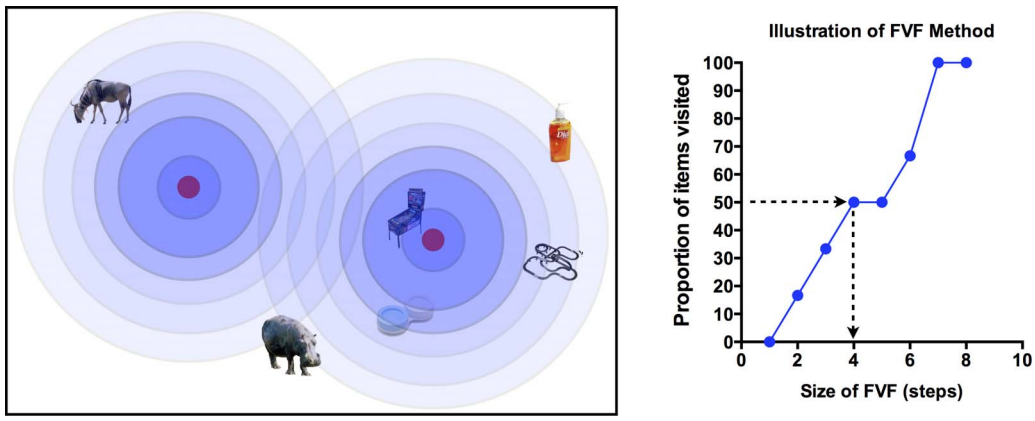

Figure 2. FVF calculations.

window around this point. Items falling within this circular window were considered attended. As previously stated, Young and Hulleman (2013) systematically varied the size of the radius around each fixation until at least 50% of the items were viewed on target-present trials. If search were random over the course of an experiment the O would need to visit, on average, 50% of the items prior to finding the target on present trials. Therefore, the threshold is set to half of the items. We conducted this analysis for each MSS as outlined in Figure 2. We increased the radius around each fixation from 0.38 degrees of visual angle (DVA) to 13.2 DVA, then found the point at which 50% of the distractors were visited for each O as function of MSS. Following Young and Hulleman, we then used the FVF estimate generated from present trials to estimate the proportion of items visited (i.e., coverage) on absent trials.

Results

Behavioral results

Figure 3 shows RT as a function of VSS (Figure 3a) and MSS (Figure 3b). Throughout the Results section, the Greenhouse–Geisser correction was applied in cases in which violations of sphericity were detected, and generalized eta squared (g-η^2) is reported as a measure of effect size. The behavioral data provide a conceptual replication of Wolfe's (2012) previous work. As in previous work, error rates were quite low. Even with a MSS of 100 objects, Os were correct 86% of the time. There was a main effect of both the presence of a target item and MSS on response accuracy: target presence, $F(1, 13) = 31.19$, $p < 0.001$, generalized eta squared (g-η^2, a measure of effect size) = 0.26; MSS, $F(1.45, 18.86) = 88.81$, $p < 0.001$, g-$\eta^2 = 0.52$, and the two factors interacted significantly, $F(1.48, 19.31) = 23.78$, $p < 0.001$, g-$\eta^2 = 0.05$. RT data on correct trials followed the same pattern of large main effects for both factors: target presence, $F(1, 13) = 28.72$, $p < 0.001$, g-$\eta^2 = 0.25$; MSS, $F(1.38, 17.89) = 83.82$, $p < 0.001$, g-$\eta^2 = 0.52$, and a significant interaction between the two, $F(2.46, 31.94) = 24.26$, $p < 0.001$, g-$\eta^2 = 0.04$.

In order to evaluate whether the RT increase associated with increasing MSS followed a linear or log-linear function, we used the data from MSS one, four, and 16 to predict data from 100 items in memory (Drew & Wolfe, 2014). As can be observed in Figure 3, the log-linear model does a much better job of predicting actual behavior than the linear model in all conditions. The log-linear estimate was reliably closer to the observed value than the linear estimate in all cases (all t values > 7, all p values < 0.001). For example, with 100 items in memory, it took the average O 4.5 s to finish an absent trial with eight visual items to search. Although the log-linear model predicted these trials would take 4.9 s, the linear model predicted they would take 13.8 s. The slight overestimate of RT by the log-linear model probably reflects a modest speed–accuracy trade-off because errors are greatest at the largest MSS.

Turning to the slope of the RT × VSS functions, slope was significantly influenced by target presence, $F(1, 13) = 37.26$, $p < 0.001$, g-$\eta^2 = 0.28$. In addition, there was a significant effect of MSS on search slope, $F(1.35, 17.49) = 45.45$, $p < 0.001$, g-$\eta^2 = 0.47$. In sum, the behavioral results are in strong agreement with previous work by Wolfe and subsequent work by Drew and Wolfe using a rapid serial visual presentation stream of objects and letters (Drew & Wolfe, 2014).

Eye-tracking results and models of hybrid search

All Os searched the same arrays of random items, which were generated prior to the experiment—the only exception being the target item, which was unique for

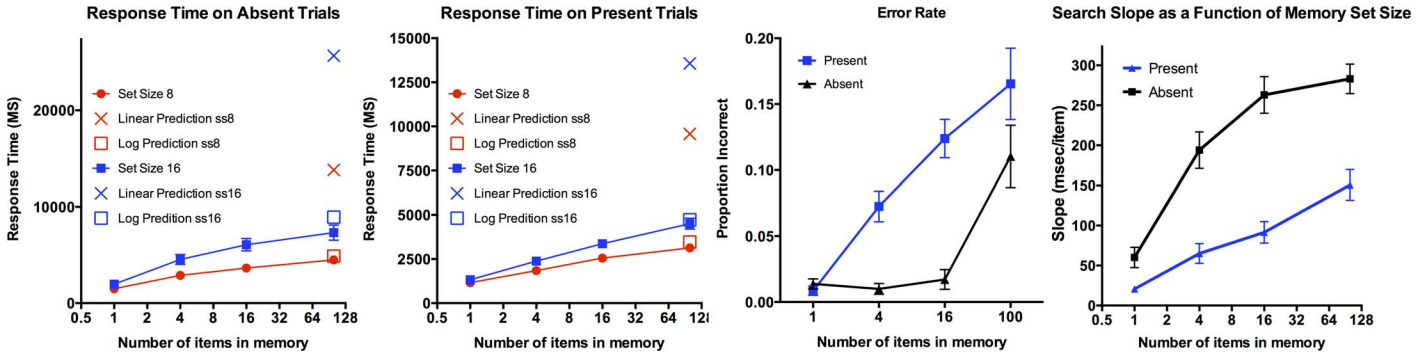

Figure 3. Behavioral results showing RT, error rate and RT × VSS slope as a function of MSS. Note logarithmic scaling of the *x*-axis.

each O. Thus, target-absent trials were identical across Os, and target-present trials were identical except for the actual target. This allowed us to create heat maps, which provide a useful depiction of how the pattern of fixations changed as a function of MSS. Figure 4 shows samples of absent trials. It is clear Os do not need to fixate every item on the absent trials of this task, particularly when the MSS is small. Below, we report the data from all correct trials as a function of both target presence and MSS. However, we focus most of discussion on target-absent trials because the interpretation of these trials is more straightforward for present purposes. Target-present trials can be terminated when the target is found. Target absent trials do not have this early exit option and are, thus, better suited to testing models of hybrid search that posit repeated searches through the same display.

In the eye-tracking data, we examined the following measures: proportion of distractors fixated, dwell time on all distractors, dwell time on fixated distractors, number of visits per distractor, and rate of refixation for distractors. For all of these measures, there is an unequivocal effect of MSS (all F values > 7, all p values < 0.001). There was a significant effect of target presence (all F values > 7, $p < 0.05$) on all measures except dwell time on fixated distractors, $F(1, 13) = 2.01$, $p = 0.18$, g-$\eta^2 = 0.02$. In all cases, the two factors interacted such that the effect of target presence became larger as MSS increased (all Fs > 6, all ps < 0.005).

What is clear to the eye in Figure 4 is born out in the eye-tracking results shown in Figure 5. Figure 5a shows that the percentage of distractors fixated on absent trials increases significantly from an average of 23% per

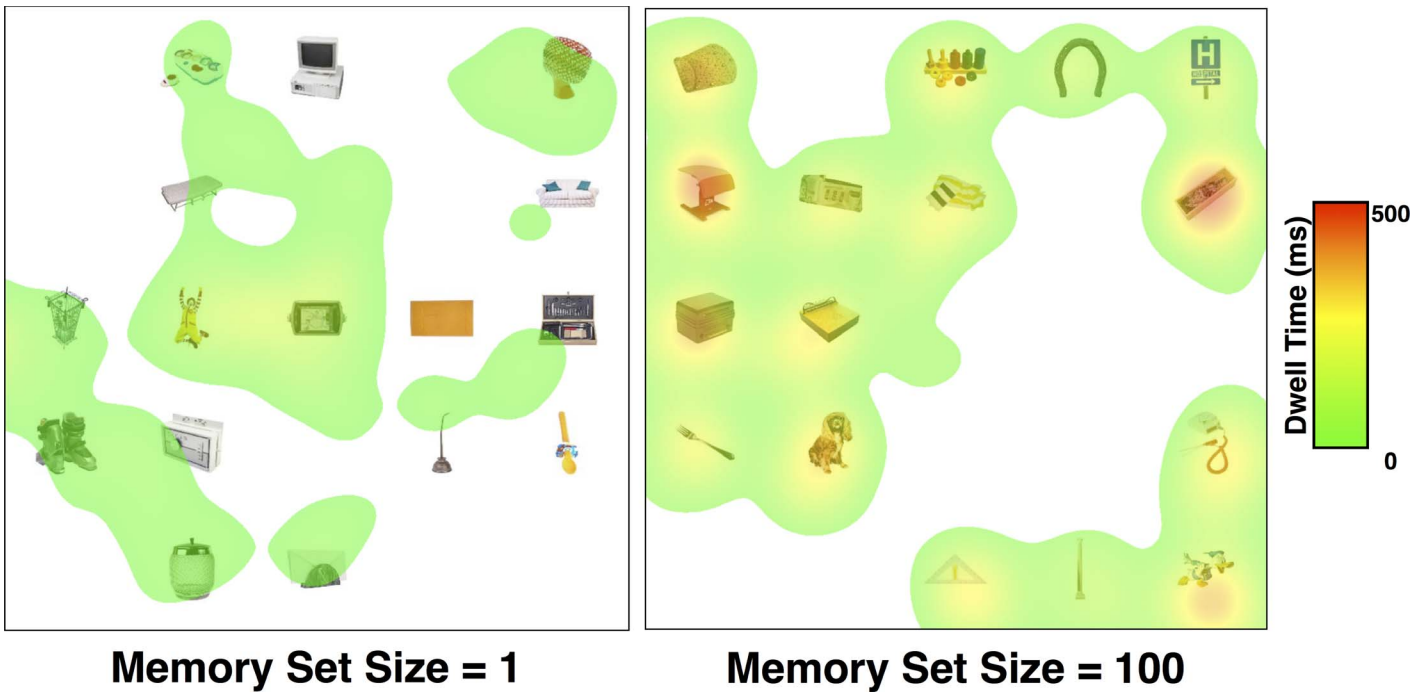

Figure 4. Heat maps for two target-absent trials averaged across all 14 Os. In this example, this indicates that 0/14 Os fixated the computer in the search array on the left. Note that these heat maps assume a 2 DVA FVF for both conditions.

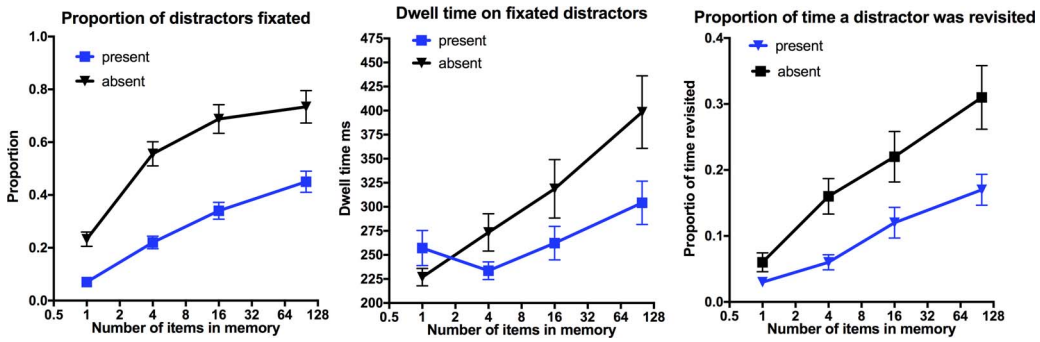

Figure 5. Eye-tracking measures as a function of MSS. Error bars represent standard error of the mean. Note logarithmic scaling of the *x*-axis.

O per trial to 73% as the MSS increases from one to 100 items, $F(1.75, 22.87) = 56.9$, $p < 0.001$, g-$\eta^2 = 0.55$. Dwell time on each distractor increased from 53 to 291 ms, $F(1.25, 16.24) = 27.98$, $p < 0.001$, g-$\eta^2 = 0.37$. This result replicates if we restrict the analysis to only distractors that were fixated, $F(1.28, 16.58) = 13.56$, $p < 0.005$, g-$\eta^2 = 0.24$; see Figure 5b). As shown in Figure 5c, Os are more likely to refixate an item when the MSS is larger. Total fixations per item on absent trials rose from 0.26 with one item in memory to 1.27 with 100 items in memory, $F(1.26, 16.32) = 29.23$, $p < 0.001$, g-$\eta^2 = 0.36$. If we restrict our analysis to items that have been fixated at least once, the chance that an item will be fixated again rises from 10% for a MSS of one to 60% at an MSS of 100, $F(1.17, 15.23) = 13.54$, $p < 0.001$, g-$\eta^2 = 0.19$.

One might have proposed that Os would need to perform one visual search for each of the members of the memory set; e.g., is there a frying pan present? Is there a truck present? Such a model would predict a linear increase in RT as a function of MSS, and as noted, that is not the case in the current data or previous work. Once again, RT increases with the log of MSS. Certainly, nothing in the eye-movement data suggests that a MSS of 100 requires 100 searches through the visual display. On target-absent trials, the mean rate of visits per distractor did increase from 0.26 with one item in memory to 1.27 with 100 items in memory, $F(1.26, 16.32) = 29.23$, $p < 0.001$, g-$\eta^2 = 0.36$. However, this increase of ~4.8× is far from that hundredfold increase that would have been expected if each additional item in memory led to an additional visual search.

Given the increase in RT with the log of MSS, it could be proposed that Os perform a smaller set of sequential searches. Logarithmic RTs can be obtained, for example, by a process that eliminates some percentage (e.g., half) of items on each iteration. Does the display contain any of these 50 memory set items? If yes, the search can be terminated and the O can respond to the item; if not, check with 25 of the remaining memory items and continue until the item is found or you have exhausted all your items. The low revisitation rates in the eye-tracking data argue against this account. If an item was fixated, the chance that it would be refixated on that trial increased from 10% to 60% as MSS increased from one to 100 items. Although this is a significant increase, $F(1.17, 15.23) = 13.54$, $p < 0.001$, g-$\eta^2 = 0.19$, items were not even fixated twice, on average, even at the highest MSS. The difference between this model of how hybrid search is accomplished and the data is particularly evident on target-absent trials. Here, a model based on eliminating 50% of the potential targets on each trial predicts more than five revisits on each distractor for a memory set of 100. In our data, each distractor was revisited less than one time per trial. In addition, although overall error rate does increase with MSS, the relatively low (~14%) error rate at MSS 100 is not consistent with performance in which the O never looked for a large percentage of the potential targets. If we assume that there is perfect memory for rejected distractors, an item that is fixated more than once could be taken as evidence for a search strategy that involves comparing a visual item to a subset of mnemonic representations. If memory for distractors is imperfect, revisits could be due to reselection of a forgotten distractor. In fact, Os do not even fixate each item twice, even when searching for one of 100 potential targets. We believe that this strongly argues against the idea that hybrid search is accomplished by looking first for one object or set of objects and then for others. Thus, the data appear to be most consistent with a single search through the display with imperfect memory for rejected distractors. Other data suggest that memory for rejected distractors is imperfect (Horowitz & Wolfe, 1998).

FVF analyses

At lower MSS, far fewer items are fixated. This clearly indicates that some distractors can be processed and rejected without fixation. Otherwise, we would need to predict a miss error rate of about 74% when

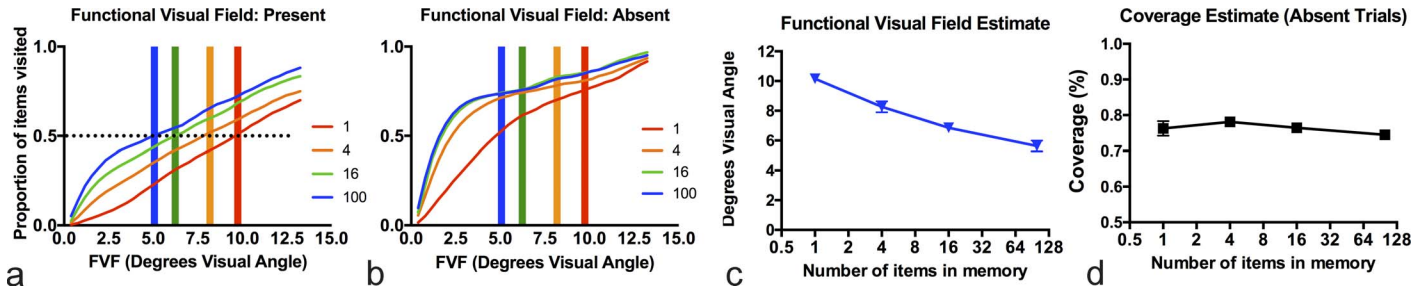

Figure 6. FVF analyses. (A) FVF was estimated by measuring the size of the FVF when 50% of distractors were visited on present trials. Solid bars indicate this level for each MSS. (B) Plot of the FVF estimates from panel A on absent trial data. (C) Plot of estimated FVF derived from panel A. (D) Plot of coverage estimates on absent trials based on FVF trials from present trials. Error bars in panels C and D represent standard error of the mean.

MSS is one because only 26% of items are fixated on absent trials in that case. The actual miss error rate is just 1%. This indicates that the *FVF* (Rayner & Fisher, 1987; Young & Hulleman, 2013) changes as a function of MSS. Our estimates of FVF demonstrate a strong effect of MSS, $F(3, 39) = 100$, $p < 0.001$, g-$\eta^2 = 0.81$. Using this metric, we estimate the size of the FVF decreased from ~10 DVA when searching for a single item to ~6 DVA when searching for one of 100 items. Following Young and Hulleman (2013), we then estimated the proportion of distractors visited (i.e., coverage) on absent trials based on our estimate of FVF (see Figure 6b, d). In sharp contrast to more traditional methods of coverage that assume a constant FVF (See Figures 4 and 5), this metric of coverage did not vary as a function of MSS, $F(3, 39) = 2.03$, $p = 0.126$, g-$\eta^2 = 0.08$; See Figure 6d). Based on this metric, roughly 76% of the distractors were processed on absent trials irrespective of MSS.

These results underline the "functional" aspect of the functional field of view. Because the FVF changes dramatically with MSS in the absence of any change in the actual stimulus, the size of the FVF cannot be simply attributed to simple effects of acuity or crowding. The task modulates the FVF.

Discussion

In sum, the eye-tracking data support an account in which Os make a single search through the visual image in a hybrid search. The amount of time required to process each visual item increases as a function of the log of the MSS. At low MSS, the demands of the memory search are moderate enough to allow the observer to determine that some objects near the current fixation can be rejected without additional foveal processing. Thus, high performance can be obtained despite Os having fixated only about one quarter of the items on the screen. As a corollary, the FVF is relatively large. At large MSS, it takes longer to process each item. Indeed, by the time MSS is 100, the memory search burden is so great that determining whether a single item is a target occupies the entire dwell time. We did not test if Os could determine if an item was one of 100 target types while not fixating on the object. Left to their normal devices, Os undoubtedly fixated the current object of attention while determining if it was in the memory set. Some items are refixated, perhaps because memory for rejected distractors is imperfect. MSS increases the time required to handle each visual item that is selected. Other changes in the visual search appear to follow from changes in that memory search time.

This work is related to a number of recent studies that have investigated the role of the target template in visual search. According to this line of research, Os determine where to look for potential targets primarily on the basis of three sources of information: low-level salience, scene context, and target template information (Malcolm & Henderson, 2010). Target templates are held in memory, then compared to visual information in order to determine whether a given item is a target (Olivers, Peters, Houtkamp, & Roelfsema, 2011; Zelinsky, 2008). As a result, distractors that are more similar to the target are more likely to be fixated (Findlay, 1997; Zelinsky, 2008). In the current study, low-level salience and scene context were carefully equated across conditions, so in this view, only differences in the target template would modulate behavior. Previous research has demonstrated that searching for more than one category of targets results in less efficient guidance toward target features (Menneer et al., 2012). Similarly, Godwin, Hout, and Menneer (2014) manipulated the fidelity of the target template by providing a target cue picture that either exactly or approximately matched the target. They found that RT, scan-path ratio (which measures the efficiency of the eye-movement path to the target), and decision time (which measures the time that elapses between first fixating a target and an affirmative button

press) all increased when targets were less well defined. Thus, increasing the number of potential targets appears to result in similar changes in behavioral and eye-tracking differences observed when the target template is deliberately weakened.

As noted above, these results demonstrate that the FVF is not a simple product of the physical stimulus. In the same display, given a particular fixation, an object away from the point of fixation may be successfully processed when MSS is one but not processed when MSS is 100. This can be seen as an example of a form of tunnel vision produced by cognitive load—in this case, a memory load (de Haas, Schwarzkopf, Anderson, & Rees, 2014; Mackworth, 1965; Williams, 1985). Based on this interpretation, one might imagine that something like the dramatic inattentional blindness effects of Mack and Rock (1998) could be obtained with the larger MSS.

The large changes observed in our estimates of FVF illustrate an important aspect of eye-tracking research that is often overlooked. Many studies of visual attention necessarily make assumptions about the size of the FVF although it is unusual to see these assumptions formalized. Any time we estimate the percentage of an image that has been covered by the eyes and any time we estimate the dwell time on an object, we are making an assertion about how far away an object can be from fixation and still be processed. Indeed, the heat maps in Figure 4 use this same line of reasoning, which is why coverage for MSS one appears so much lower than MSS 100. However, by estimating coverage based on the functional field of view size in target-present trials, we saw that the coverage estimate did not vary with the MSS in target-absent trials. Closely examining the data with FVF in mind reveals that this simple shortcut of assuming a constant FVF across conditions may lead to misleading conclusions if taken at face value. One illustrative example of this comes from the medical image perception literature. Many studies have shown that experts make far fewer fixations and longer saccades than novices evaluating the same medical image (e.g., Bertram et al., 2016; Kundel & La Follette Jr., 1972). It is therefore likely that simple measures of coverage would suggest that experts examine less of the image yet find more abnormalities. We predict that a FVF analysis would produce larger FVFs in experts than in novices. This could reflect an expert's ability to pull more meaningful information from the periphery thanks to years of experience. Alternatively, it could reflect the expert's knowledge of where *not* to look. Either way, traditional methods of estimating coverage based on a static estimate of FVF may lead to misleading results, particularly when analyzing between-subject differences in expertise.

Using the methods of Young and Hulleman (2013), we could use our estimate of FVF from present trials to estimate the proportion of items that were processed on these trials as a function of MSS. Young and Hulleman estimated that between ~71% and ~89% of the items were processed, depending on the difficulty of their search tasks. If the remainder of items were not processed, one would expect overall error rates to be approximately 100% minus the percentage coverage. Thus, when our FVF calculation yields a coverage of ~76%, this means that 24% of items were not processed and that 24% of targets should be missed. Moreover, our estimate of this rate did not vary with MSS. This prediction is at odds with the actual error data. Error rates are generally too low and the false-alarm rate clearly increases with MSS. Young and Hulleman note that such apparent discrepancies suggest that the methods of deriving FVF may be underestimating the actual size of the FVF. Clearly more work needs to be done to map out the implications of this issue.

Conclusions

Although the bulk of the visual search literature has been focused on the intricacies of searching for a single potential target, in the real world we frequently engage in hybrid searches in which there is more than one potential target. The current study is one of the first to use eye tracking to examine how the relatively well-understood mechanisms that underlie search for a single target are modulated by manipulating the size of the O's target set.

The growing hybrid search literature makes it clear that increasing the number of potential targets decreases the efficiency of visual search. As the first study to explore this paradigm using eye tracking, the current data provide strong evidence that this manipulation yields changes in eye-movement behavior that are consistent with a search process in which each item in the visual display is compared to all the targets in the memory set. With more items in that memory set, more time is required; however, the search through memory is not a linear function of the set size. Instead the memory search time is a logarithmic function of the set size such that the memory search comparing an item in the visual field to one of 16 possible targets in memory takes roughly 4× (not 16×) as long as the comparison of one visual item to one item in memory. The current study contributes to our growing understanding of this process by demonstrating that both dwell time on distractors and proportion of distractors fixated reliably increased with MSS. We interpret these findings to provide strong evidence that hybrid search is accomplished via a single visual search with multiple

memory searches in which increased dwell time reflects increased time searching memory.

Keywords: visual search, eye tracking, memory search, target templates

Acknowledgments

This work was supported via funding for TD from the DOD via Cooperative Agreement W911NF-07-2-0055 and NIH – NEI EY017001 to JMW.

Commercial relationships: none.
Corresponding author: Trafton Drew.
Email: Trafton.Drew@psych.utah.edu.
Address: University of Utah, Salt Lake City, UT, USA.

References

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America, A*, *5*(12), 2210–2219, doi:10.1364/JOSAA.5.002210.

Bertram, R., Kaakinen, J., Bensch, F., Helle, L., Lantto, E., Niemi, P., & Lundbom, N. (2016). Eye movements of radiologists reflect expertise in CT study interpretation: A potential tool to measure resident development. *Radiology*, *281*(3), 805–815, doi:10.1148/radiol.2016151255.

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

Cunningham, C. A., & Wolfe, J. M. (2014). The role of object categories in hybrid visual and memory search. *Journal of Experimental Psychology: General. 143*(4), 1585.

de Haas, B., Schwarzkopf, D. S., Anderson, E. J., & Rees, G. (2014). Perceptual load affects spatial tuning of neuronal populations in human early visual cortex. *Current Biology*, *24*(2), R66–R67, doi:10.1016/j.cub.2013.11.061.

Drew, T., & Wolfe, J. M. (2014). Hybrid search in the temporal domain: Evidence for rapid, serial logarithmic search through memory. *Attention, Perception & Psychophysics 76*(2), 296–303.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.

Findlay, J. (1997). Saccade target selection during visual search. *Vision Research*, *37*(5), 617–631.

Godwin, H. J., Hout, M. C., & Menneer, T. (2014). Visual similarity is stronger than semantic similarity in guiding visual search for numbers. *Psychonomic Bulletin & Review*, *21*(3), 689–695.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *394*(6693), 575–577.

Hulleman, J., & Olivers, C. N. (2017). The impending demise of the item in visual search. *Behavioral and Brain Sciences, 40*; e132.

Kundel, H. L., & La Follette Jr., P.S. (1972). Visual search patterns and experience with radiological images. *Radiology*, *103*(3), 523–528, doi:10.1148/103.3.523.

Leite, F. P., & Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Attention, Perception & Psychophysics*, *72*(1), 246–273, doi:10.3758/APP.72.1.246.

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

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science*, *3*(1–12), 67–68, doi:10.3758/BF03343023.

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision*, *10*(2):4, 1–11, doi:10.1167/10.2.4. [PubMed] [Article]

Menneer, T., Stroud, M. J., Cave, K. R., Li, X., Godwin, H. J., Liversedge, S. P., & Donnelly, N. (2012). Search for two categories of target produces fewer fixations to target-color items. *Journal of Experimental Psychology: Applied*, *18*(4), 404–418, doi:10.1037/a0031032.supp.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391, doi:10.1038/nature03390.

Olivers, C. N., Peters, J., Houtkamp, R., & Roelfsema, P. R. (2011). Different states in visual working memory: When it guides attention and when it does not. *Trends in Cognitive Sciences*, *15*(7), 327–334.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442, doi:10.1163/156856897X00366.

Rayner, K., & Fisher, D. L. (1987). Letter processing during eye fixations in visual search. *Perception and Psychophysics*, *42*(1), 87–100, doi:10.3758/BF03211517.

Rayner, K., & Fisher, D. L. (2013). Eye movements and the perceptual span during visual search. In *Eye movements from physiology to cognition: Selected/editing proceedings of the Third European Conference on Eye Movements, Dourdan, France, September 1985* (p. 293). Amsterdam: Elsevier.

Shiffrin, R., & Schneider, W. (1977). Controlled and automatic human information processing. II: Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190.

Williams, L. J. (1985). Tunnel vision induced by a foveal load manipulation. *Human Factors: the Journal of the Human Factors and Ergonomics Society*, *27*(2), 221–227, doi:10.1177/001872088502700209.

Williams, L. J. (1989). Foveal load affects the functional field of view. *Human Performance*, *2*(1), 1–28.

Wolfe, J. M. (2012). Saved by a log. *Psychological Science*, *23*(7), 698–703, doi:10.1177/0956797612443968.

Wolfe, J. M. (2017). "I am not dead yet!"–The Item responds to Hulleman & Olivers. *Behavioral and Brain Sciences, 40*, e161, doi:10.1017/S0140525X16000303.

Wolfe, J. M., Boettcher, S. E. P., Josephs, E. L., Cunningham, C. A., & Drew, T. (2015). You look familiar, but I don't care: Lure rejection in hybrid visual and memory search is not based on familiarity. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(6), 1576–1587, doi:10.1037/xhp0000096.

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.

Young, A. H., & Hulleman, J. (2013). Eye movements reveal how task difficulty moulds visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 168–190, doi:10.1037/a0028679.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787–835, doi:10.1037/a0013118.