CHAPTER

43

Visual Search: The Role of Memory for Rejected Distractors

Todd S. Horowitz

ABSTRACT

Serial models of attentional deployment in visual search have traditionally assumed sampling without replacement. Each item in a search array was thought to be selected only once, and rejected distractors were never revisited. This efficient search pattern requires some form of high-capacity memory for every deployment of attention. Several investigators suggested that inhibition of return served to implement this memory (see Chapter 16). Recently, this assumption has been challenged on several fronts. Experiments using the randomized search paradigm, search for multiple targets, and attentional reaction time methods suggest that search often uses sampling with replacement. Studies of eye movements have yielded a range of findings, from perfect memory to substantial evidence for resampling. Simple strategies can reduce the probability of resampling without positing high-capacity memory. Such strategies carry a cost in terms of speed and may explain the oculomotor data. Memory for rejected distractors in visual search is likely to be of limited capacity, similar to other aspects of visual memory.

I. SAMPLING IN VISUAL SEARCH

How is attention deployed during search of a cluttered scene? We know that the salience of objects in the present drives attention, whether salience derives from intrinsic properties (bottom-up, see Chapters 38, 39, 43, 93), or task relevance (top-down, see Chapters 12, 14, 18, 40, 96). Furthermore, it has long been assumed that past attentional deployments determine future deployments, in that attention is assumed to avoid objects that have already been searched. In the language of probability theory, visual attention is supposed to use sampling without replacement from the search array. We can use a simple urn model to illustrate the difference between sampling with replacement and sampling without replacement. Imagine we are looking for the red marble in an urn that contains 1 red marble and 20 blue marbles. Using sampling without replacement, we would set aside each marble after determining its color. If we were sampling with replacement, however, we would return the marble to the urn before pulling the next marble, thus making it likely that we would pull the same marble from the urn more than once. Sampling without replacement is clearly more efficient, which presumably explains why it was the default assumption for theories of search.

A. The Standard Model

This assumption is explicit in a wide range of search models that employ a serial component (see Chapter 43). Despite their differences, all of these models state that attended items are never resampled. Moreover, the assumption is implicit in many other papers. For instance, the routine claim that a serial model should produce a 2:1 ratio of target absent to present slopes is based on this assumption. Therefore, we will call the sampling without replacement option the Standard Model.

Under the Standard Model, as each distractor is rejected, it is somehow "marked" or inhibited. Consequently, the probability of detecting the target in a given epoch increases during the course of the search. For example, assume a standard search task in which there is one target among d distractors. Further assume that all items are equally salient. In the first epoch, the probability of detecting the target is $1/d$. In the second epoch, the probability is $1/(d-1)$, and so forth. Thus, the "hazard function" rises exponentially during the course of the search, reaching 1.0 in the dth epoch.

Models that do not employ a serial attentional component are formally equivalent to the Standard Model for present purposes if they produce an increasing hazard function.

The time to find targets using sampling without replacement is governed by the negative hypergeometrical distribution, and the average number of samples $E(s)$ to find a target in a display containing t targets and d distractors is given by Eq. (43.1):

$$E(S) = \frac{t+d+1}{t+1} \quad (43.1)$$

The common claim that "half of the items are searched before the target is found" derives from this equation. If we set $t = 1$, and N = set size $(d + t)$ $E(S) = (N + 1)/2$; on average, $(N + 1)/2$ items will be sampled before the target is found.

B. The Amnesic Model

The obvious alternative to the Standard Model is a model based on sampling with replacement. In such a model, items would be selected without regard to whether they had been selected before. Free recall, for instance, is widely held to proceed by sampling with replacement from memory. Under sampling with replacement, only the current salience would determine the probability of selection. If all objects were equally salient, selection would be essentially random. In the human factors literature (Arani, Karwan, and Drury, 1984; Chan and Courtney, 1998), this has been termed the *random search model*, as opposed to the *systematic search model*, which corresponds to what we have been calling the Standard Model here. However, we prefer the term *amnesic search*, which emphasizes the memoryless character of the model.

In contrast to the Standard Model, amnesic models produce flat hazard functions; the probability of finding the target is constant over time. In the case of a single target among d distractors, in each epoch the model has a $1/d$ probability of detecting the target. As a result, there is a small but finite probability that the target will never be attended.

Sampling with replacement is governed by the negative binomial distribution, and the number of samples required to find a target in a display containing t targets and d distractors is given by Eq. (43.2):

$$E(S) = 1 + \frac{d}{t} \quad (43.2)$$

In the case of a single target, $E(S) = d + 1$, so the number of samples needed to find the target is equal to the set size. Note that this does not mean that every item in the array is examined; this would be an exhaustive, memory-driven search. Instead, it is likely that some items would be examined several times, while other items might not be examined at all.

C. Other Roles for Memory in Visual Search

In statistics, a process using sampling with replacement is said to be "memoryless" because the state of the system (here the focus of attention) at any given time is not affected by previous states of the system. When we say that visual search is memoryless (Horowitz and Wolfe, 1998), we do not mean that no memory is involved in the process of search. On the contrary, memory is known to be critical to search. Memory processes and search processes interact on many levels. These include working memory (see Chapter 100), as well as implicit contextual guidance of visual attention (Chapter 40). The claim that search is memoryless is strictly an argument about whether the deployment of attention is determined by the history of previous deployments (or, to put it more broadly, about the form of the hazard function). Note that this view is also not incompatible with the idea that information about distractors is acquired during search. Subjects may have an accurate memory for the characteristics of distractors, yet still be unable to prevent attention from returning to them.

At the same time, memory-driven search does require some mechanism to prevent attention from returning to rejected distractors. The dominant hypothesis in the literature is that inhibition of return (see Chapter 16) serves this purpose. One could postulate other mechanisms more like conventional memory systems. Whatever "memory" structure is held to underlie sampling without replacement, it must have a high capacity to keep track of all distractor locations in a standard visual search experiment.

II. EMPIRICAL TESTS OF THE STANDARD MODEL

Researchers have studied the question of memory-driven versus amnesic search (aka systematic versus random search strategies) since Krendel and Wodinsky (1960). However, until recently, this debate has taken place within the applied literature and has had little impact on the development of mainstream cognitive theories of search. Recent work in the basic vision science literature was sparked by the development of the randomized search procedure by Horowitz and Wolfe (1998).

A. Randomized Search

One strategy to determine whether subjects use sampling with or without replacement under normal circumstances is to devise conditions under which the sampling strategy is known and to compare performance in such a forced condition to performance under standard conditions. The logic behind the randomized search procedure is to force subjects to use sampling with replacement. In the randomized search procedure, each trial consists of a series of search frames. The stimuli (targets and distractors) are held constant across frames. In the *dynamic* condition, however, the positions of all items are randomly shuffled (within certain constraints, see below) from frame to frame, while in the *static* condition, locations are also held constant, so that there is in fact no stimulus change from frame to frame. In the static condition, subjects are free to use either sampling without replacement or sampling with replacement. In the dynamic condition, however, sampling without replacement is impossible because there is no way for the visual system to know which distractors in the current frame correspond to distractors rejected on previous frames. Therefore, on each frame, search starts anew.

Predictions for the standard and amnesic models can be derived in a straightforward fashion from Eqs. (43.1–43.2). Under both models, performance in the dynamic condition is governed by Eq. (43.2). If there are n items in the search array, subjects have to examine an average of n items before finding the target. In the static case, the standard model holds that subjects can sample without replacement, which is much more efficient. To be precise, subjects need to examine only $(n+1)/2$ items on average. Thus, the RT $\times$ set size slope in the dynamic condition should be steeper than that in the static condition by a ratio of $2n:(n+1)$, or roughly 2:1. According to the amnesic model, however, subjects in the static condition are using sampling with replacement anyway, so there should be no difference in search efficiency (RT $\times$ set size slope) between the two conditions. Of course, the two conditions are not identical from a stimulus point of view. The dynamic condition by definition contains a variety of dynamic events, onsets and offsets, which may affect RT. However, these factors should have their influence on the nonsearch aspects of the task, and show up in the intercept of the RT $\times$ set size function. Thus, the critical datum is the slope ratio.

Horowitz and Wolfe (1998) found that the slopes for static and dynamic conditions were not reliably different from each other in a range of conditions. These included versions intended to thwart "sit and wait" strategies where the subject might fixate at one location and wait for the randomly replotted target to appear at fixation (von Muhlenen, Muller, and Muller, 2003). The basic finding was replicated in a number of labs (Gibson, Li, Skow, Salvagni, and Cooke, 2000; Kristjánsson, 2000). Kristjánsson (2000) produced a notable failure to replicate when he used large set sizes. However, his stimulus configuration may have led to more masking in the dynamic condition. Horowitz and Wolfe (2003) repeated the experiment with large set sizes under conditions intended to reduce masking in the dynamic condition. As before, they found that dynamic and static slopes were essentially identical. Moreover, RT distributions were similar across conditions. This should not be the case if static and dynamic search used different sampling mechanisms.

Although these data argue against the standard model, Gibson et al. (2000) used a similar method to show that, while subjects may not use memory for rejected distractors to guide subsequent search, they remember the location of a target when searching for multiple targets (see also Horowitz and Wolfe, 2001; Takeda, in press).

Although the substantial similarity between dynamic and static search is impressive, the randomized search method has a number of drawbacks. First, as noted earlier, the two conditions can never be equated for stimulus quality, so performance will never be equalized on all measures. Second, the argument for amnesic search, at least, relies on accepting the null hypothesis. Third, the logic of the method will fail if the search can be completed during a single frame. Fourth, experiments must be designed in order to thwart subjects from adopting different search strategies in the two conditions. In particular, if the target location is selected randomly on every frame, subjects may simply monitor a single location or set of locations and wait for the target to come to them. Horowitz and Wolfe attempted to circumvent this problem by restricting the locations at which the target could appear, though this may not have been entirely successful (von Muhlenen et al., 2003). Finally, subjects appear to have difficulty deciding when to give up on target-absent trials in dynamic conditions, so switching to a forced-choice design (i.e., one of two targets is always present) will significantly reduce dynamic error rates (Horowitz and Wolfe, 2003).

B. Multiple-Target Search

In the classic visual search task, where subjects have to find a single target, both the standard model and the amnesic model make similar predictions for the central tendency of RTs. (Distributions are a different matter, see next section.) However, when subjects must find multiple targets during a single trial, the two models make quite different predictions. Compare Eq. (43.1) and Eq. (43.2). As each target is found, the value of t in the denominator of both equations decreases. However, the value of d in the numerator of Eq. (43.1) also decreases as distractors are marked off, so that the interval between finding targets is constant for the Standard Model. In contrast, the value of d in Eq. (43.2) is constant, so that in the amnesic model, the intervals increase nonlinearly.

Of course, measuring multiple RTs during a single trial is problematic, so Horowitz and Wolfe (2001) devised a method in which subjects are asked to report as soon as they know that there are at least n targets in a display containing m targets. Both n and m are varied across trials (n across blocks in this case). The RT $\times$ n function for constant m is a proxy for the RTs to m successive targets.

Horowitz and Wolfe, using alphanumeric stimuli (digit targets among letters), found that this function is highly accelerated, disconfirming the standard model. However, Takeda (in press), noted that the search rate may depend on n, possibly due to memory load. The interaction of working memory (see Chapter 102), and visual search is not entirely clear.

C. Cumulative Distribution Functions

Since the Standard Model and the amnesic model have different hazard functions (section I), they also have different cumulative distribution functions (CDFs). Sampling without replacement will yield linear CDFs, which reach 1.0. Sampling with replacement will produce exponential CDFs, which approach but never quite reach 1.0, since there is always a chance that the target will not be located (Krendel and Wodinsky, 1960). Visual search studies in the mainstream cognitive literature rarely report CDFs. However, fitting CDFs has been a popular method to distinguish between random and systematic search in the human factors literature. A number of papers have reported exponential CDFs in visual search (e.g., Chan and Courtney, 1998), though linear CDFs have also been reported.

As Chan and Courtney (1998) noted, RT CDFs have two components: the search time (which may be distributed linearly or exponentially), and a nonsearch time (including response time), which is normally distributed. They found that search RTs are well described by the ex-Gaussian distribution, in which a normal distribution (representing nonsearch time) is convolved with an exponential (representing search time). Horowitz, Wolfe, and Alvarez (see discussion in Horowitz and Wolfe, 2003) attempted to measure the search CDF directly, eliminating the nonsearch component. Subjects searched for a mirror-reversed letter and reported its color. The colors of all items changed at some time T during the search. The dependent variable was the probability $p(T)$ that the subject reported the initial color of the target. Plotting $p(T)$ against T yielded the CDF for the search process. This function was exponential, supporting an amnesic account of search.

III. MEMORY IN THE OCULOMOTOR DOMAIN

Although the issue is not settled, the balance of the data suggest that covert deployments of attention are not guided by memory for prior deployments. What about overt deployments of the eyes? As we noted above (see Chapter 16) IOR has been proposed as an inhibitory tagging mechanism to support sampling without replacement in search, and has been tightly linked to the oculomotor system. Therefore, one might expect that eye movements would be memory-driven. Indeed, eye movements tend to be biased away from the previous fixation locations (see Chapter 16). On the other hand, Horowitz and Wolfe (2003) employed the randomized search method using small stimuli that required fixation and found no difference between the slopes of static and dynamic functions. Measurements of CDFs for oculomotor search support a memoryless account (Scinto, Pillalamarri, and Karsh, 1986).

The most obvious way to assess whether eye movements sample a visual search stimulus with or without replacement might be to record fixations and see whether or not items are revisited. A number of researchers have taken just such a straightforward approach. Peterson et al. (2001) reported that the hazard function to fixate a search target in an oculomotor search task was increasing rather than flat, indicating as their title stated that "visual search has memory," at least in the oculomotor domain. The question is not as easily resolved as one might think. Gilchrist and Harvey (2000) pointed out that much depends on the algorithm used to determine which fixations are counted as refixations. They arrived at a figure of 50 percent refixations for their task, but suggested that this may have been an overestimation. Moreover, the search tasks used in these experiments seem to be particularly inefficient and may have encouraged subjects to use an ordered pattern of eye movements; "reading" a display is one example.

IV. THE COST OF SYSTEMATIC SEARCH?

Why don't all searches use a systematic pattern of attentional deployments? This could provide the benefits of the standard model without the demand for perfect memory. If sampling without replacement could be accomplished by shifting attention or the eyes in a predetermined scanpath, then the only memory required would be memory for which direction to move in. You are using such a scanpath in reading this article. Why don't subjects simply scan search arrays left-to-right, in an expanding spiral, or some other efficient search pattern?

Wolfe, Alvarez, and Horowitz (2000) have demonstrated that when subjects are required to search systematically (e.g., clockwise around a circular array), search rates are slow (·300 ms/item), compared to control conditions in which search strategy is unconstrained. Although systematic search is efficient in terms of the number of samples required, the cost in terms of search rate appears to be too great for displays in which rapid, covert deployments of attention are possible. When search rate is limited to the slower rate of overt deployments of the eyes, systematic search may become a viable option.

V. LIMITED-CAPACITY MEMORY?

Memoryless search and the standard model are really two ends of a continuum. It is possible that there is some, imperfect memory for rejected distractors. The experiments that falsify perfect sampling without replacement do not falsify limited or fallible memory systems. Indeed, a successful limited-capacity model might resolve many controversies in this area. Two versions of this solution have been proposed so far. Horowitz and Wolfe (2001) introduced a buffer model in which attention was prevented from orienting to the *c* most recently examined items, but the remaining arrays items were sampled randomly, even those that had been sampled in the past. After each item was attended, whichever item had been attended $c + 1$ samples back was dropped off the stack, and the current item was added. For the first *c* samples, this model followed the Standard Model and afterward behaved as an amnesic model with the set size reduced by *c* items. This compromise model was a better fit to their data. The average estimate of *c* was three items for their subjects. (Note that substantially larger estimates were obtained by Takeda, in press.) A similar model was used to fit CDFs by Horowitz, Wolfe, and Alvarez (see discussion in Horowitz and Wolfe, 2003), who arrived at an estimate of one to two items. A similar account of memory for overt shifts of attention was proposed by McCarley et al. (2003), who observed a buffer size of three to four items. Moreover, recent work suggests that inhibition of return can be measured for the last five to six attended items.

An alternative conception is the variable memory model of Arani, Karwan, and Drury (1984). Instead of allocating a fixed capacity, Arani et al. proposed that subjects might fail either to encode or to retrieve attended locations, and that re-fixation or re-attending might occur through such forgetting.

VI. GENERAL CONCLUSIONS

It seems likely that the Standard Model is wrong. Current research is developing a more nuanced and precise understanding of the degree to which rejected distractors are remembered in visual search and the mechanisms that might underlie such a memory.

References

Arani, T., Karwan, M. H., and Drury, C. G. (1984). A variable-memory model of search. *Human Factors* **26**, 631–639.

Chan, A. H., and Courtney, A. J. (1998). Revising and validating the random search model for competitive search. *Perceptual & Motor Skills* **87**, 251–260.

Gibson, B. S., Li, L., Skow, E., Salvagni, K., and Cooke, L. (2000). Memory-based tagging of targets during visual search for one versus two identical targets. *Psychological Science* **11**, 324–328.

Gilchrist, I. D., and Harvey, M. (2000). Refixation frequency and memory mechanisms in visual search. *Current Biology* **10**, 1209–1212.

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

Horowitz, T. S., and Wolfe, J. M. (2001). Search for multiple targets: remember the targets, forget the search. *Perception & Psychophysics* **63**, 272–285.

Horowitz, T. S., and Wolfe, J. M. (2003). Memory for rejected distractors in visual search? *Visual Cognition* **10**, 257–298.

Krendel, S. E., and Wodinsky, J. (1960). Search in an unstructured visual field. *Journal of the Optical Society of America* **50**, 562–568.

Kristjánsson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science* **11**, 328–332.

McCarley, J. S., Wang, R. F., Kramer, A., Irwin, D. E., and Peterson, M. S. (2003). How much memory does oculomotor search have? *Psychological Science* **14**, 422–426.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., and McCarley, J. S. (2001). Visual search has memory. *Psychological Science* **12**, 287–292.

Scinto, L. F. M., Pillalamarri, R., and Karsh, R. (1986). Cognitive strategies for visual search. *Acta Psychologica* **62**, 263–292.

Takeda, Y. (in press). Search for multiple targets: Evidence for memory-based control of attention. *Psychonomic Bulletin & Review.*

von Muhlenen, A., Muller, H. J., and Muller, D. (2003). Sit-and-wait strategies in dynamic visual search. *Psychological Science* **14**, 309–314.

Wolfe, J. M., Alvarez, G. A., and Horowitz, T. S. (2000). Attention is fast but volition is slow. *Nature* **406**, 691.