Spatial Vision, Vol. 17, No. 4-5, pp. 327–352 (2004)

Also available online - www.vsppub.com

Termination of a visual search with large display size effects

DENIS COUSINEAU [1,*] and RICHARD M. SHIFFRIN [2]

[1] *Département de Psychologie, Université de Montréal, C. P. 6128, succ. Centre-ville, Montréal, Quebec, H3C 3J7, Canada*
[2] *Psychology bldg., Indiana University, Bloomington, Indiana 47405, USA*

Received 3 June 2003; revised 4 November 2003; accepted 29 November 2003

Abstract—The ability to locate an object in the visual field is a collaboration of at least three intermingled processes: scanning multiple locations, recognizing the object sought (the target), and ending the search in cases when the target is not found. In this paper, we focus on the termination rule. Using distribution analyses, it is possible to assess the probability of termination conditional on the number of locations examined. The results show that on some trials without target, the participants carried out more comparisons than there are objects in the display; in other conditions, they carried out fewer comparisons than objects. Because there were very few errors, the premature stops were not pure guesses. We present models to account for these findings. The distributions of terminations help determine the slopes of the functions relating response time to set size.

Keywords: Visual search; distributions; termination rule.

INTRODUCTION

In typical visual search experiments, a to-be-sought target is presented and followed by a test display containing D objects. Half the trials contain one target, half do not, and the participant decides which has occurred. Accuracy is usually high and response times (RT) are graphed as a function of D.

This area of research is made both interesting and challenging by the fact that the process of search is generally an amalgam of many lower-level abilities. For example, Townsend and Nozawa (1995) identified the following skills used to carry out a visual search:

- The ability to recognize a target and select the corresponding response.
- The process by which multiple locations are examined (the scanning process).
- The rule determining when to stop a search (the termination rule).

*To whom correspondence should be addressed. E-mail: Denis.Cousineau@Umontreal.CA

The recognition process is of course a critical and essential component (Biederman, 1987). One can try to isolate recognition by presenting just one object in the test display. In this case the visual search task reduces to a same-different judgment task (Bamber, 1969), a situation very similar to an identity priming task (Huber *et al.*, 2002). Even in this seemingly simple task, a number of factors affect both accuracy and response time, and a full understanding of performance is yet to be achieved (Sternberg, 1998).

Response selection is in principle an important component of search, but because most search studies use just two responses ('Target present' and 'Target absent'), this process is not usually an object of modeling or empirical study (but see Logan, 1994). Because our study uses the typical two responses, we do not try to assess separately the contributions of object recognition and response selection to overall performance.

Commonly held views suggest that scanning the stimuli can be done either in serial or parallel, particularly depending on stimulus difficulty. At least two types of results support this dichotomy. The results rely critically on the slope of the function relating mean RT to D (the display size defined as the total number of stimuli displayed). In some studies the slope is near zero (the function is close to flat), suggesting a parallel search (e.g. Treisman and Gormican, 1988). In others, the RT functions are roughly linear with a slope clearly above zero, a finding for which many researchers have proposed a serial comparison mechanism (e.g. Shiffrin and Schneider, 1977). This distinction is not truly clear-cut because it has long been known that parallel models can produce such linear functions if target rejections are not independent and/or if the scanning rates vary with D (e.g. see Townsend and Colonius, 1997).

The termination rule indicates how a search is terminated. One simple rule would involve scanning all the locations (exhaustive search) before making a response (Sternberg, 1966, proposed such a rule for memory search). However, if there is a cost in scanning unneeded locations once a target has been identified, a more efficient rule might involve termination of the search as soon as a target is found (self-terminating search, as in Shiffrin and Schneider, 1977). An efficient termination rule on target-absent trials would be exhaustive, proceeding one location at a time and avoiding scanning any locations more than once (Horowitz and Wolfe, 1998).

Although models of search with restrictive assumptions have been oft tested, truly general conclusions have proved difficult to reach, partly because the data collected and examined are usually too simple to support strong inference. Saying this in another way, less restrictive assumptions allow many different model types to predict the same data patterns. For example, the conjunction of a serial and self-terminating search with constant scanning rate and constant rejection rate predicts that the present targets are found on average halfway through the sequence of scans. As a consequence, the slope of the target absent mean RT is twice that of the target present mean RT function (Wolfe, 1998). However, other models make the same prediction (Lu and Dosher, 1998). Townsend and Ashby (1983) exhibited a plethora of models that do mimicking at the level of the mean RT.

In what follows, we present an approach that goes a long way toward allowing identification of the termination rule. To preview the results, the termination rule found in our experiment could not be termed exhaustive or self-terminating, partly because it involved many premature stops. Our examination of the termination rules in this study also allowed us to draw a number of inferences concerning the nature of scanning. The approach we used is not restricted to mean RT but instead looks at the whole distribution of RT, reducing the risk of model mimicking considerably.

Distribution analyses

The following discussion is predicated upon the assumption that scanning is serial, an assumption that will be tested on the data presented in the Experimental section.

A typical visual search paradigm uses uniformly random target placement. Therefore, the target, if present, is located in any of the D display locations with a probability $1/D$. Thus a serial self-terminating search (SSTS) will, regardless of scan order, encounter a target in scan position i with probability $1/D$. If the time to carry out a single scan is unimodal, then in the $D = 1$ condition, both 'target-present' and 'target-absent' responses will be made according to a unimodal distribution. In the $D = 2$ target-present condition, 50% of the RT should occur at the same time as in the $D = 1$ condition, and the other 50% should occur later, after the time needed to reject the first stimulus and scan the second location. Similar logic applies to larger values of D. Of course, the components of the response time distributions will not necessarily be visible, particularly if the variance of one scan is high relative to the mean time. On the other hand, if the mean time to scan and switch to a new location is low relative to the variance of a scan, the target present responses should form a multimodal distribution, as illustrated in the left part of Fig. 1. Superimposing two or more distributions is called a mixture; mixtures of distribution will be defined more precisely in the Results section. Finally, assuming independence of the processing times for different locations and different scanning orders, and assuming that the variance in the time to reject a distractor is comparable to the variance of accepting a target, the variance of the ith scan should be i times the variance of the first scan.

In the case of a target absent trial, all the responses should occur following the scan of the last location (exhaustive search), as seen in Fig. 1, right part. Again, the variance of the (unique) component should be D times larger than in the $D = 1$ condition. In principle, if the $D = 1$ target present distribution can be characterized in some way, and if these various auxiliary assumptions hold true, the only remaining unknown is the scanning rate.

A more complex situation would occur if the participants could terminate the search prematurely, for example with a guess. Indeed, because the participants are generally allowed a small number of errors (often 5%), they might skip a few

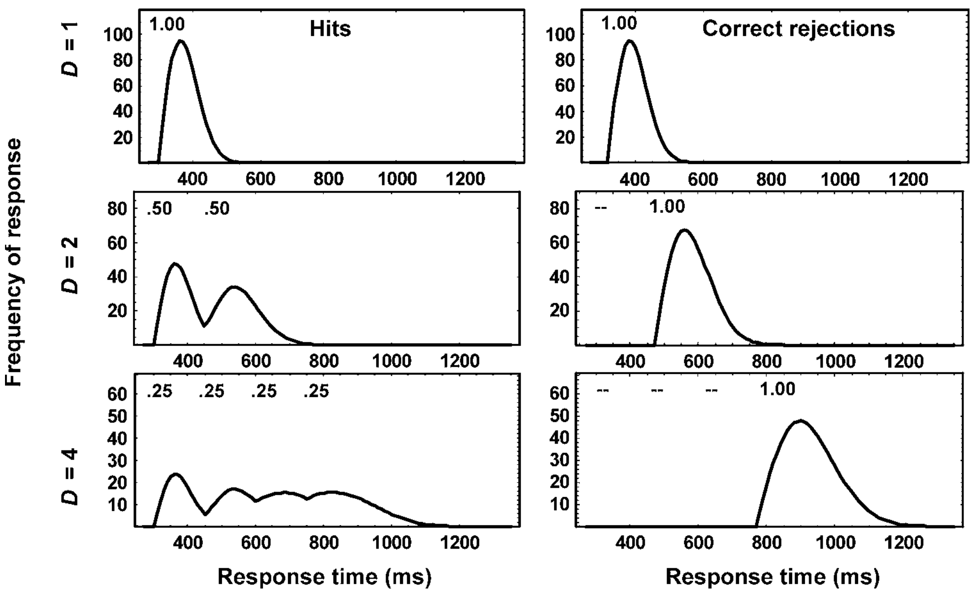

Figure 1. Response frequencies predicted by a serial self-terminating search (SSTS) for display sizes 1, 2 and 4 (rows), assuming that each scan results in a unimodal distribution and that the scanning rate is large (i.e. the separation between component is large). The quantities within each panel indicate the probability that the search stops at the corresponding scan. Each panel assumes a total of 360 trials (for comparison with the Experiment). Left panel: target-present trials, showing a self-terminating search; right panel: target-absent trials, showing an exhaustive search.

locations to speed their mean RT. If a participant makes pure guesses prematurely on 10% of the trials, and has no errors otherwise, the resultant error rate would match the allowed 5% (because the other 5% would be correct guesses).

It must be pointed out that errors are not distributed uniformly. Their number generally increases as display size increases. Furthermore, it is often found that only the percentage of missed targets ('misses'), and not the percentage of falsely located targets ('false alarms') increases with display size. To explain such results, Cousineau and Larochelle (in press) assumed that all guesses are 'target absent' responses, and that the probability of making such premature guesses increases as the duration of search increases. As premature guessing increases over search duration, the percentage of hits will decrease over search duration (e.g. Zenger and Fahle, 1997); i.e. fewer hits will occur late, a scenario seen in the left part of Fig. 2. Within a given display size, the misses will be more numerous if the premature stops are made early rather than late, a situation illustrated in the central part of Fig. 2. Finally, the correct rejection distributions will be contaminated by the correct early guesses, a situation shown in the right part of Fig. 2.

In the remainder of this article, we use SSTS(=) to refer to the standard serial self-terminating search without premature termination (self-terminating on target present, exhaustive on target absent). This model predicts that the hits are distributed equally across the D locations. We use SSTS(*) to refer to a serial search model

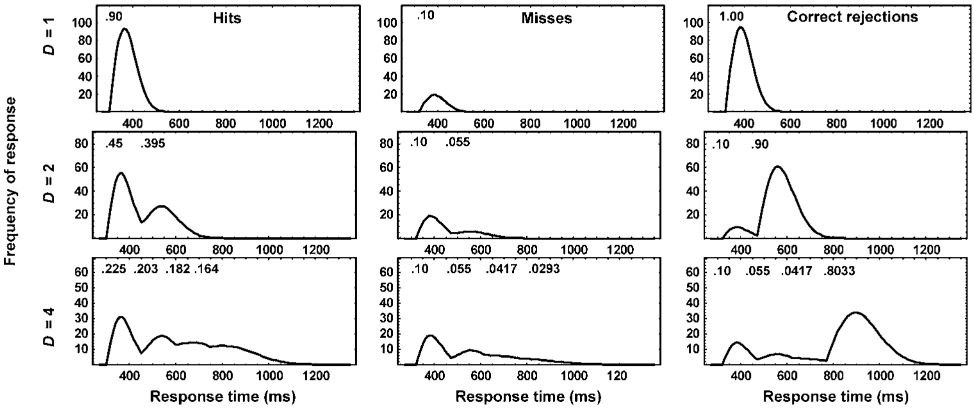

Figure 2. Distributions predicted by a serial self-terminating search with some early guesses for display sizes 1, 2 and 4 (rows), assuming that each scan results in a unimodal distribution and that the separation between modes is large. In this scenario, the base rate of a premature stop before each location is 10%. Left panel: target-present trials, showing a self-terminating search contaminated by some early guess so that the last modes are under-represented; central panel: missed target on target-present trials, which shows the distributions of the incorrect guesses; right panel: target-absent trials, showing an exhaustive search partly contaminated by early correct guess.

that is mostly self-terminating but where a certain number of trials can end with premature termination. The sole difference between SSTS(=) and SSTS(*) is the process of termination.

The Experiment section that follows presents highly trained participants (45 sessions) who performed a search task that resulted in a large set size effect. We used difficult stimuli to induce serial comparisons and to slow scanning to the point where different modes in the response time distributions for target-present responses would become visible.

EXPERIMENT

To produce a search task with large display size effect, we created novel stimuli that would be difficult to distinguish. The use of such stimuli typically decreases the likelihood of a parallel search and increases the magnitude of the effects. At least two operational definitions of stimulus difficulty have been offered in the past. The first one is related to the learning schedule of the participants. Stimuli become ‘easy’ to detect when they are consistently mapped to the response (Shiffrin and Schneider, 1977). The second one is based on the constituents of the stimuli, the features. If only one feature (e.g. the color) identifies uniquely the to-be-found target (disjunctive search), the participants locate the target rapidly and accurately, which suggests a parallel scanning process. On the other hand, if a conjunction of features is required (e.g. blue S among red S and blue T), search is often slower and seems to induce a serial scanning process. Stimulus consistency and stimulus composition can be pitted one against the other to evaluate their relative importance, but it is not the purpose of this article (see Larochelle, Lefebvre and Cousineau, in preparation). One might think it optimal to combine varied mapping with feature conjunctions to achieve our purposes, but varied mapping also tends to increase variability, use of different strategies, and general task stress, undesirable characteristics for a study in which participants must provide data for 45 or more sessions. We therefore used a conjunction of features to create the targets but used consistent mapping.

Method

The results reported here are from the standard visual search conditions of a larger study carried out at Indiana University (see Note 1).

Subjects. Four participants (3 women), all right-handed and with vision corrected to normal, completed a total of 74 sessions; sessions 45 and up were devoted to transfers that are not discussed here. They were paid an average of 8 dollars per session. Participant C was the first author.

Stimuli. The stimuli are shown in Fig. 3, top panel. These stimuli were intended to be very difficult to discriminate; the distractors used on each trial made it impossible to use the presence or absence of just one spoke to determine whether a given display item was or was not a target (a conjunctive search). Two of the four targets are defined by the presence of two spokes indicating a clock time

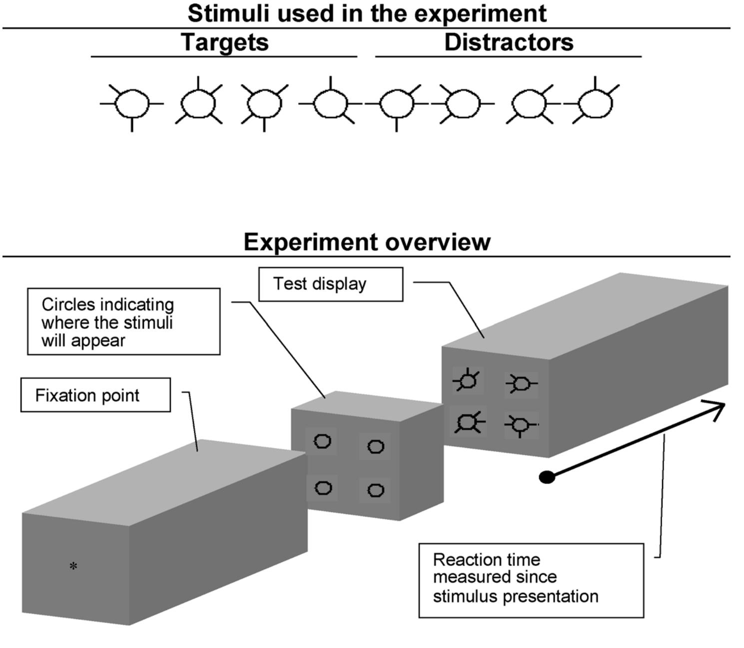

Figure 3. (Top) Stimuli used in the experiment. (Bottom) A typical target present trial, where the display size D is 4.

of four o'clock (no distractor contains this pattern). The other two targets are defined by the presence of two spokes indicating a clock time of ten-thirty (no distractor contains this pattern). The three naive participants were not informed of this structure, and post-experiment questioning suggested they never noticed it. The features, excluding the central circle that is common to all targets and distractors, do not touch, preventing the possibility that new features emerge at the junctions. In addition, we designed the set so that each target is more similar to the set of distractors than to the other targets (a claim that can be verified using various physical similarity measures). The participants had to learn the targets; the set of target stimuli was shown on a sheet of paper during the first session only; they remained fixed for the 74 sessions of 648 trials each (consistent mapping), one third of which are from the standard visual search conditions. Only the last ten sessions of training (sessions 35 to 44) are reported in this article.

Procedure. The procedure is outlined in the bottom part of Fig. 3. Each trial started with a fixation star lasting 1000 ms in the middle of the display. For 500 ms, circles showed where the stimuli were to appear (an atypical procedure that was adopted in consideration of the sequential conditions making up the bulk of the experiment); then the test display appeared and remained until a response was given. The size of the test display, seen at 50 cm was within 2° vertically and 3° horizontally. The number of stimuli on the test display (D) was 1, 2 or 4, varying at random. The positions used were the corners of an imaginary square. On 50% of the trials, a target was present in a random position. No distractors were repeated on a given trial. Responses were given with the right hand using the '1' and '2' keys on the numeric keypad of a computer keyboard, followed by feedback on speed and accuracy for 1 s. Participants were informed to respond as fast as possible but not to exceed 5% errors. No particular instructions related to eye movements were given. Recapitulative feedback was provided every 108 trials.

Results

Figure 4 presents the individual mean RT and the percent of errors for 'target present' and 'target absent' responses as a function of display size. Each point is based on 360 observations, less the few trials on which errors occurred (and less twelve trials for which RT exceeded 1.8 seconds).

Error rates were relatively low: misses of targets averaged 4.8%; false alarms on negative trials were virtually absent for three of the four participants, averaging 1.4% (see lower graphs in Fig. 4; Chun and Wolfe, 1996; Zenger and Fahle, 1997). The four participants produced roughly linear RT functions. The largest negative to positive slope ratio (for A) was about 2 : 1, and the lowest (for B) was about 1.5 : 1. If one drew inferences only from these ratios (a fairly common approach in the literature), many interpretations would be possible (e.g. A could have used serial self-terminating search, B could have alternated between serial self-terminating

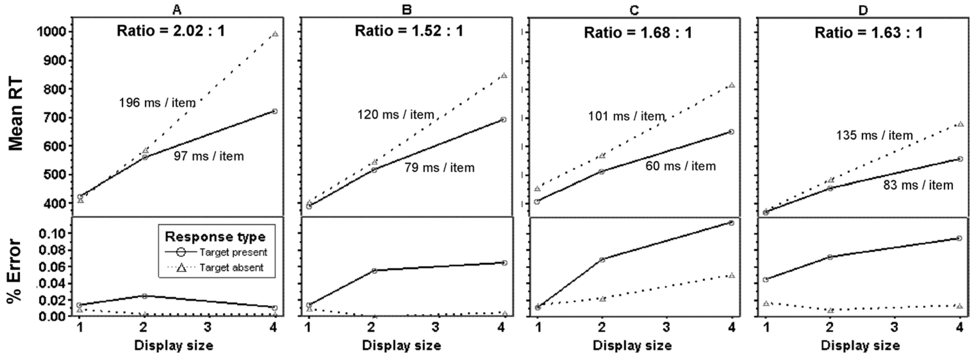

Figure 4. Mean reaction time and percent of errors as a function of participants and display size. Top panels present mean reaction time for both target present and target absent trials as a function of display size. Standard error bars are smaller than the symbols. Bottom row shows percent of errors. Circles are for missed target errors whereas triangles are for false alarms.

and serial exhaustive search, and C and D could have used mixed strategies, etc.). Unfortunately, as shown by Townsend and many others, different forms of both parallel and serial processing can produce identical mean response time predictions (e.g. Townsend and Ashby, 1983), reducing the strength of any conclusion drawn from mean RT slopes or slope ratios. Thus we must look at the data in more detail. Although various theorems demonstrate mimicry between parallel and serial processing at the level of mean RT, stronger conclusions can be drawn from the entire RT distributions, especially if one limits consideration to psychologically plausible and physically realizable models. The RT distribution results reviewed next provide some evidence that all the participants used the same search process and that this search process was serial.

Target-present RT distributions. Figure 5 shows the individual RT distributions for target-present responses (Ashby *et al.*, 1993; Luce, 1986). They are strikingly similar for the participants A and B, and show clearly distinct modes. This suggests that similar processes are at work, despite the fact that the slope ratios for A and B were the most dissimilar among the four participants. This result tends to highlight the inadequacy of using slope ratios to infer processing mechanisms. The distributions for participants C and D did not reveal clearly separate modes, so more detailed analyses are required to reveal the processes that give rise to them. In the next section, we carry out some numerical estimation that suggests all four participants may have used basically the same serial scanning process. For now we draw some qualitative conclusions based on the patterns for participants A and B. Overall, their data support a serial, self terminating processing for the following reasons:

(1) Distinct modes are visible (clearest for display size two) and the probability of responding between two modes is quite small. Plausibly, the first mode

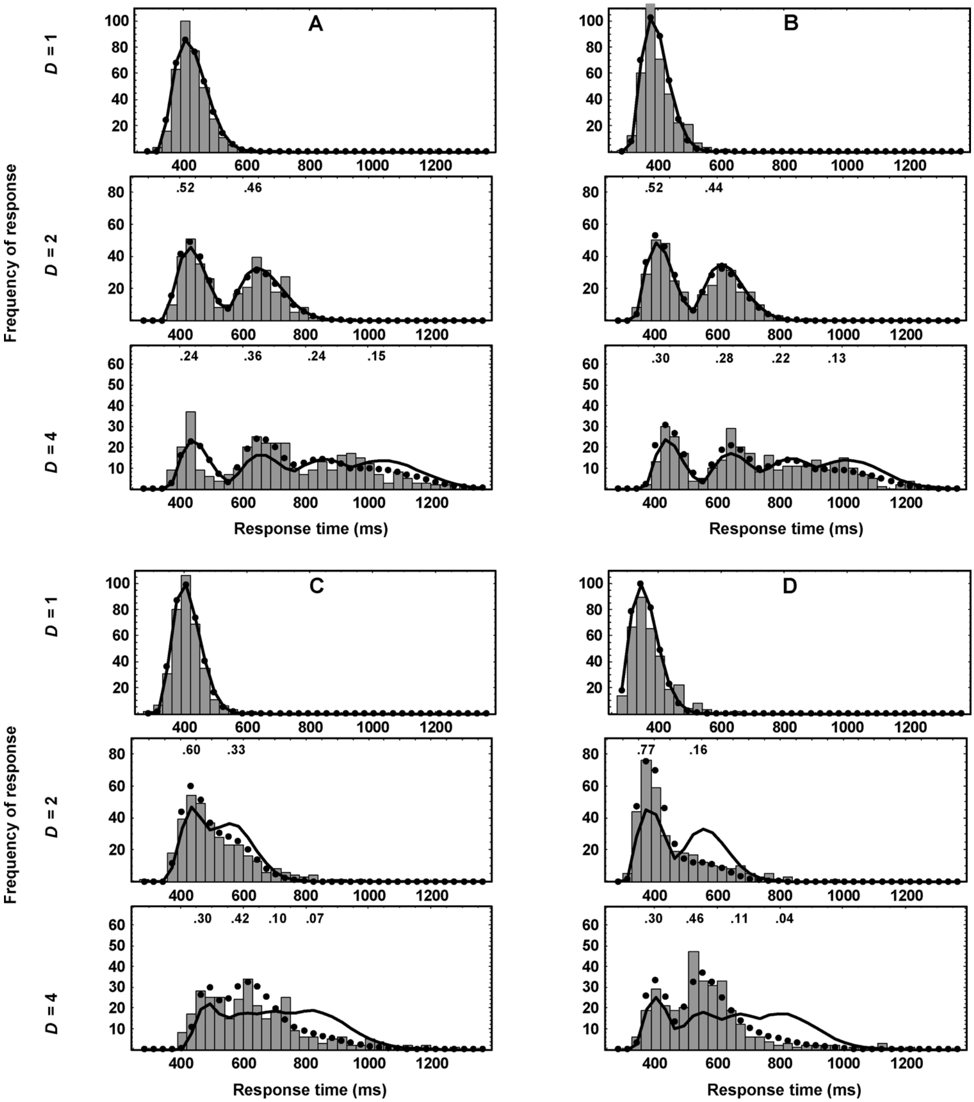

Figure 5. Response time distributions for hits (correct target present trials) for each participant (columns) and display sizes 1, 2 and 4 (rows). Histograms show the observed frequencies of response (bin width is 30 ms). Solid lines show the predicted density for a serial self-terminating search. The points show predicted frequencies calculated using unequal probability of locating the target on the ith location (see text). All predictions include a residual response time component (see text).

corresponds to the first comparison, with a response made if a target is found, and subsequent modes correspond to targets found on subsequent comparisons.

(2) The mode visible for display size one is present for the other display sizes, with about the same shape. This means that the first comparison process is independent of whether or not other locations remain to be processed. Similarly,

the second mode for display size two appears for display size four, again with about the same shape.

(3) The temporal distance between the first and second mode is about the same for display sizes two and four. Therefore, the time to scan is also independent of how many items remain to be processed. This temporal distance (called x_s in the model next) is rather large (estimated to be 190, 180, 110 and 140 ms for participants A, B, C and D respectively). Assuming a serial search, this distance corresponds to the scanning rate.

Of course, there might exist a parallel search model of some sort that could mimic these results. As shown by Townsend and Ashby (1983), the possibilities of mimicking seriality with a parallel model are endless. However, if the resources to scan items are not shared (as suggested by point 3) and if the resources to process locations are not shared either (as suggested by point 2), the resulting parallel model might not be very plausible. As we will argue later, there might exist a parallel process running along with the serial process, but it is not responsible for the observed modes.

However, there is more to the data than the evident modes in the distributions for participants A and B, so we turn next to a quantitative model for all four participants. As we shall see, a key to understanding the data is an appropriate characterization of the termination rule.

Modeling the termination rule for the correct target present data. The modes reported in the previous section for participants A and B are interesting and diagnostic; to our knowledge, it is the first time that modes have been reported. Finding such modes may have been facilitated by the difficulty of our search task and the extensive training given under consistent mapping. The first factor may have increased the temporal distance between successive scans whereas the second may have reduced variability. However, modes for participants A and B notwithstanding, the simplest serial search model cannot account for all the results. For example, according to the simplest model, each component should cover an equal proportion of the data ($1/D$), but as will be seen from the models that follow, this is not the case even for participants A and B.

The SSTS(=) model assumes that (i) there is a time to scan a location and decide whether it is a target or not. If more than one location is shown, these processes are repeated until a target is found. (ii) It also assumes that there is a residual time, including response selection and motor response times. This part of the latency is done only once in a trial and the sum of the residual process and the scanning process(es) determines the observable response time. To begin, we establish a baseline by fitting a quantitative version of the SSTS(=) model to the RT distributions of hits.

Scanning and decision process. Let f_D^+ be the probability density of correct target present decision times in display size condition D, and let P_D^+ be the percent correct in that same condition (so that $1 - P_D^+$ is the target present error rate, also termed the miss rate). Let f_{iD}^{+} represent the distribution of times in condition D when the target is found on the ith serial position. The distributions f_{D}^{+} for the display conditions $D > 1$ results from a mixture of the distributions f_{1D}^{+} to f_{DD}^{+}.

A mixture of distributions occurs when two or more data sets resulting from different processes are pooled together (Yantis *et al.*, 1991). Assuming that the distributions of each process are known and that the proportions of data taken from each are also known, the probability density of the mixture is a weighted sum of the individual densities. In the visual search context, let p_{iD} be the probability that a target-present response is made correctly at the ith comparison for display size D. According to SSTS(=) the target is equally likely to be found at any location and the components should have equal weights. This is expressed in the model by having equal proportions $p_{iD}^{+} = P_{D}^{+}/D$. Based on these, the global distribution for correct target present responses (containing all the components at once) for display size D is a mixture of distributions (Cramér, 1946; Dolan *et al.*, 2002; Feller, 1966) given by:

$$f_{D}^{+} = \frac{1}{P_{D}^{+}} \sum_{i=1}^{D} p_{iD}^{+} f_{iD}^{+},$$

where the denominator P_{D}^{+} gives the percent correct obtained in that condition. This model provides a baseline to test the termination rule. In essence, it assumes that the search is equally likely to stop at any of the D locations.

If the component distributions f_{1D} to f_{DD} are known, the mixture can be computed and fitted to the data. Because the results indicated that the first component was comparable, independent of the display size condition, let $f_{1D}^{+} = f_{1}^{+}$ for all D. Further, if a target is not found on a first scan, the same process is repeated on a second location. The probability density resulting from the addition of two random decision times with known densities is given by the convolution of the two densities (denoted *). In the present case, the same process is repeated i times, therefore $f_{iD}^{+} = f_{1}^{+} * \cdots * f_{1}^{+}$.

In sum, the distribution of decision times at the ith scan is given by a convolution and the overall distribution across all locations is a mixture of D convolutions. If the SSTS(=) model is correct, the only unknown in all this is f_{1}^{+}, the probability density in the $D = 1$ condition. Under this assumption, we will extract information from f_{1}^{+} to model the modes in the other conditions.

To characterize the f_{1}^{+} distribution, we decided to find the best-fitting Weibull distribution for each subject. Note that the choice of a Weibull distribution is not critical since we are not interested in a specific distributional shape. In a similar model, Bricolo, Gianesini, Fanini, Bundesen and Chellazi (2002) used exponential instead of Weibull distributions. The Weibull distribution (Cousineau and Larochelle, 1997) is defined by three parameters: x_s, the position of the distribution, the point where it starts to rise; β_s, the scale of the distribution, a quantity proportional to standard deviation; finally γ, the asymmetry of the distribution. In the present model, x_s represents the time to perform a scan and decide, and β_s is the variability in this process. Finally, to reduce to a minimum the number of parameters, γ was fixed at 2.0 for all subjects in all conditions (Cousineau *et al.*, 2002).

Residual processes. As said earlier, the standard SSTS model assumes that there is a residual time which adds to the latency of the decision process described above and which have its own variability. To model this aspect, we convolved the above densities with the density of the residual process. This process is characterized by two parameters, the base residual time and the variability. As previously, we chose a Weibull distribution with the shape parameter γ held constant at 2. However, a close inspection of Fig. 5 shows that the point where the distributions start to rise differ slightly with display size. The distributions are shifted by about 30 ms to the right in the $D = 4$ condition relative to the $D = 1$ condition. The same result was found by Hockley (1984) and by Bricolo *et al.* (2002). To account for this fact, we allowed the base response time to vary with D. The parameters are noted x_D^+ and β_r in the following.

Fitting the model. Fitting was done in two steps. First, the f_1^+ distributions were fitted to find the best-fitting β_s and β_r. Then, all the distributions were fitted, holding β_s and β_r constant, to find the best-fitting x_s and x_D^+. The last two parameters being additive, the f_1^+ distribution alone is not sufficient to separate them. The distributions were fitted by minimizing minus the log likelihood index of fit (Cousineau and Larochelle, 1997; Cousineau, Brown and Heathcote, in press; Van Zandt, 2000).

The best-fitting parameter values are shown in the top and middle parts of Table 1. As seen, the scanning process is slow, taking from 110 in the best case (participant C) to 180 ms per item in the worst case (participant B).

Because we did not measure eye movements, we cannot determine to what degree they might have been occurring. We think it likely they took place at least some of the time, given the long scan times; for participants A and B, scanning rates were slow enough to be compatible with eye movements for each display item. However, Bricolo *et al.* (2002) found scan rates that were much larger even though trials with eye movements were discarded. Further, participants C and D had scanning rates that were probably too fast (120 ms per item) to allow eye movements to keep up with scans. In any event, eye movement is an issue independent of the termination rule. Also note that the scanning rates estimated from the model do not match the scanning rates estimated from the slopes seen in Fig. 4. Whereas both estimates are reasonably close for participant A (190 *vs.* 196 ms/item), they are very different for participant B (180 *vs.* 120 ms/item). These discrepancies will be explained when the SSTS(*) will be examined.

Finally, the base residual times increase with D. This could result from the general difficulty of the decision to be made or from the motor response. In either case, this result is not compatible with a strict serial search model. However, Sternberg, Knoll and Turok (1990) found that the time to initiate a motor response of taps depended on the number of taps. Thus, the changes in the residual base time could be a result

Table 1.
Parameters common to both the SSTS(=) and the SSTS(*) model

	Participants				
	A	B	C	D	Average
Scanning process					
x_s	190	180	110	140	155
β_s	107	88.8	85.6	92.5	93.4
Base response times for hit trials					
x_1^+	125	120	193	118	139
x_2^+	155	150	233	158	174
x_4^+	164	180	283	182	202
β_r	19.2	17.8	36.1	11.8	21.2
Base response times for correct rejection trials					
x_1^-	145	172	275	157	187
x_2^-	45.0	60.1	193	48.2	86.6
x_4^-	39.8	69.7	177	41.9	82.1

All parameters are expressed in ms.

of post decision components. Moreover, modeling of the ‘target-absent’ RT will suggest that it is not a perceptual effect since the base residual times are constant for some of the conditions when the target is absent.

The minimized minus log likelihood statistics are shown in the top part of Table 2. To assess the quality of fit of the mixture, a χ^2 measure of fit was computed. This test computes the differences between the observed frequencies and the expected frequencies from the SSTS(=) model. To do so, the data were divided into bins of 20 ms (except where the observed frequencies were smaller than 5 counts, in which case the bin and its successive bin were merged — Hays, 1973). The statistics is given by

$$G^2 = \sum_i \frac{(n_i - o_i)^2}{n_i},$$

where o_i is the observed count in the ith bin and n_i is the expected count in that same bin, computed from the model using the mixture. The test uses a number of degree of freedom based on the total number of bins minus the number of parameters minus 1. The number of bins varies between subjects, between display size conditions and between responses; it ranged from 11 to 12, 17 to 20 and 19 to 30 in the $D = 1$, $D = 2$ and $D = 4$ conditions respectively. The results are shown between parentheses in Table 2. This χ^2 test is an approximate test for two reasons. First, the bin sizes and locations are arbitrarily chosen and different results occur when they are changed. Second, the test assumes that the underlying distributions

Table 2.
The minus log likelihood (and the χ^2 index of fit in parenthesis) for the SSTS(*) model and the SSTS(=) model for participants A through D

	Model	Participants			
		A	B	C	D
Hits					
$D = 1$	SSTS(=,*)	1881.1 (9.98)	1902.6 (12.1)	1836.2 (3.65)	1885.2 (9.86)
$D = 2$	SSTS(=)	2101.1 (23.6)	2149.5 (14.7)	2165.6 (35.6)**	1983.2 (84.4)**
	SSTS(*)	2099.4 (23.6)	2121.5 (10.1)	2086.9 (10.3)	1851.5 (17.8)
	Improvement	1.70	28.0**	78.7**	131.7**
$D = 4$	SSTS(=)	2395.4 (46.9)*	2266.3 (36.5)	2189.0 (59.6)**	2144.9 (159)**
	SSTS(*)	2372.2 (30.8)	2253.7 (17.9)	2165.2 (31.0)*	2053.7 (25.6)
	Improvement	23.20**	12.6**	23.8**	91.2**
Correct rejections					
$D = 1$	SSTS(=,*)	1840.5 (16.9)	1784.2 (13.9)	1758.7 (15.4)	1904.6 (11.3)
$D = 2$	SSTS(=)	2065.0 (16.3)	2110.2 (12.1)	2111.8 (31.8)**	2125.4 (21.8)*
	SSTS(*)	2065.0 (16.3)	2110.2 (12.1)	2026.5 (19.7)	2055.5 (18.8)
	Improvement	0.00	0.00	85.3**	69.9**
$D = 4$	SSTS(=)	2267.3 (19.2)	2587.5 (28.4)*	2968.5 (49.6)**	3027.6 (132)**
	SSTS(*)	2267.3 (19.2)	2336.7 (17.9)	2312.0 (10.4)	2293.0 (23.8)
	Improvement	0.0	250.8**	656.5**	734.6**

* $p < 0.05$.
** $p < 0.01$.
The fit of the models are assessed using an approximate χ^2 index of fit. The improvement is assessed using an exact test, the Likelihood ratio test.

are truly Weibull distributions. However, we are not really interested in exploring that hypothesis.

As seen in Table 2, the SSTS(=) model fits the $D = 1$ condition well, and the $D = 2$ condition well for two subjects, but does not fit the $D = 4$ condition for most subjects.

The solid lines in Fig. 5 show the distributions predicted by the SSTS(=) model. The figure explains why the quality of the fits decreased with display sizes. It is not a problem with the estimation of the scan rate (the distance between modes) but a problem with the height of the modes. First, as seen in the $D = 4$ conditions, the last predicted modes should still produce a visible bump for all participants whereas the data show no fourth mode. Second, in the $D = 2$ condition, the height of the first mode is always underestimated. The same is true for modes 1 and 2 in the $D = 4$ conditions.

One way to deal with these deviations involves relaxing the assumption that termination occurs with equal probability at all scan positions (possible reasons for this inequality will be taken up in the General Discussion). To generalize the model, therefore, we allowed the p_{iD}^{+} to vary freely, except that their sum must equal P_D^+. Ideally we should refit all the other parameters under this assumption, but we thought it more informative (and conservative) to fix all the other values to those given in Table 1. The resultant model, termed SSTS(*), was discussed in the Introduction, and the predictions are given as the dots on Fig. 5. The best-fitting values of p_{iD}^+ are given above each mode in each panel of Fig. 5. As seen, the predictions of the SSTS(*) match closely the observed frequency of response times for all participants, so that little could be gained by refitting the other parameter values.

To test the fit, a χ^2 test of fit was performed in the same manner as previously. The results are shown in Table 2. However, because the SSTS(*) model is a less constrained version of the SSTS(=) model, there exists an exact test, the Likelihood Ratio Test (Bozdogan, 1987), to check whether the improvements are significant. Seen the other way around, SSTS(=) is nested within the more general SSTS(*) model. According to the test, let the index of fit LRT be twice the improvement in the minimized minus log likelihood index of fit. This test is said to be an exact test because LRT does not depend on the bin sizes (contrary to the above tests) and it is exactly distributed as a χ^2 distribution with the degree of freedom given by the number of fixed parameters (that is, 2 and 4 in the $D = 2$ and 4 conditions respectively).

This test shows that the p_{iD}^+ are significantly not equal across modes ($p < 0.01$), except in one case (participant A, $D = 2$, $\chi^2(2) = 3.40$, $p > 0.05$). The top part of Table 2 shows the improvements for all the participants in the $D = 2$ and $D = 4$ conditions.

Consider now the p_{iD}^+ estimates given in Fig. 5. All participants exhibited a tendency to terminate early, but this trend was quite extreme for participants C and D. For example, for participant D when $D = 2$, only 15.7% of the hits occur when processing the second location, as compared to about 50% predicted by the standard serial search model ($1/D$ less half of the 6% of misses observed in that condition).

A different way of summarizing these results for hits is given in the top part of Table 3, giving the probabilities of finding the target in the ith scan location, according to the serial self-terminating search and the unconstrained search model, averaged across participants. As seen in the numbers for SSTS(*), there is a pronounced tendency for search to terminate early.

It is useful to estimate what may be termed the 'efficiency' of the termination rule, based on E, the expected number of scans carried out until a target is found on a positive trial. For $D > 1$, let efficiency be defined as

$$\frac{E(D) - E(1)}{D - 1}.$$

For the SSTS(=) model,

$$E(D) = \frac{D + 1}{2},$$

Table 3.
Proportions p_{iD} of finding the target on the ith scan as a function of the display size D, as predicted by the SSTS(=) model and as estimated by the SSTS(*) model averaged across subjects

	ith scan	SSTS(=)	SSTS(*)	Deviation
Hits				
$D = 2$	1	0.475	0.60	+13%
	2	0.475	0.35	−13%
	$\overline{P}_2^+$	$\overline{0.95}$	$\overline{0.95}$	
$D = 4$	1	0.23	0.29	+6%
	2	0.23	0.38	+15%
	3	0.23	0.17	−6%
	4	0.23	0.10	−14%
	$\overline{P}_4^+$	$\overline{0.92}$	$\overline{0.92}$	
Correct rejections				
$D = 2$	1	—	0.00	0%
	2	0.99	0.95	−4%
	3	n/a	0.04	+4%
	$\overline{P}_2^-$	$\overline{0.99}$	$\overline{0.99}$	
$D = 4$	1	—	0.00	0%
	2	—	0.00	0%
	3	—	0.41	−41%
	4	0.98	0.57	+41%
	$\overline{P}_4^-$	$\overline{0.98}$	$\overline{0.98}$	

$\overline{P}_D^+$ is the percent of hits averaged across subjects in the D size condition.
$\overline{P}_D^-$ is the percent of correct rejections averaged across subjects in the D size condition.

and efficiency equals 0.5. In other words, half of the items are checked on a typical trial. For SSTS(*), $E(D)$ is given by $\sum_{i=1}^{D} i \times p_{iD}$. Based on the p_{iD} values given in Fig. 5 for SSTS(*) the resultant efficiency values for hits are given in the top part of Table 4. There is a large difference between the participants, participant A being closest to 0.5 and participant D the most distant. On average, the participants scan about one third (33%) of the locations for $D = 2$ and $D = 4$.

Summary of the hit distributions. The unconstrained SSTS(*) model provides a quite adequate model for the hit distributions, suggesting that all participants used a form of serial search, despite the wide variation in the slopes and the negative to positive slope ratios.

The major result of this section is that the search is definitely not simply self-terminating. With the possible exception of participant A, participants B, C, and D are more likely to end the search early relative to a randomly ordered search. One explanation involves premature termination with a guess, but that model would predict large numbers of errors for large D, and we shall see that this prediction is inconsistent with the observed data.

Table 4.
Efficiency of the search for each participant and for display sizes 2 and 4

	Participants				
	A	B	C	D	Average
Hits					
$D = 2$	0.44	0.40	0.26	0.09	0.30
$D = 4$	0.43	0.35	0.24	0.25	0.32
Correct rejections					
$D = 2$	1.00	1.00	1.04	1.06	1.03
$D = 4$	1.00	0.93	0.77	0.71	0.85

SSTS(=) predicts an efficiency of 0.5 for target present and an efficiency of 1.0 for target absent.

Correct 'target absent' RT distributions. Figure 6 presents the correct rejection RT distributions in the same format as in Fig. 5. As was the case for hits, the results for participants C and D differed from those for A and B. It is interesting that participant A seems to wait to finish scanning before responding negatively, and has minimum negative RTs of almost 750 ms in the $D = 4$ condition. The other participants tend to respond negatively at early times, a tendency most pronounced for participant D, who begins to respond negatively as early as 400 ms in the $D = 4$ condition.

If negative trials are based on an exhaustive scan, then the distributions in Fig. 6 should match those predicted by SSTS(=) for the last component. The SSTS(=) predictions are shown in Fig. 6 with a solid line. The exhaustive serial search model fits essentially perfectly for participant A, but poorly the other participants. This result is supported by χ^2 measures of fit presented in the lower part of Table 2.

An approach more consistent with the SSTS(*) model for the hits would allow premature terminations on negative trials. Let these termination probabilities be termed p_{iD}^{-}. There is no particular reason why the p_{iD}^{-} should match the p_{iD}^{+}, so these were re-fitted freely, as well as the base response times parameters, denoted x_D^{-}; the other parameter values were carried over from those used to fit the positive trials.

The resultant predictions for SSTS(*) are given in Fig. 6 as the dotted lines. These predictions are quite good. The results of the Likelihood Ratio tests for improvement from SSTS(=) to SSTS(*) are shown in the bottom part of Table 2. Further, the average across participants of the termination probabilities by scan number by D value are given in the bottom part of Table 3.

The x_D^{-} estimates are given at the bottom of Table 1. As seen, the x_D^{-} values were close to identical in the $D = 2$, and $D = 4$ conditions. Therefore, the increase in the base time seen on hits is not the result of a pre-decisional stage since by definition, the presence or absence of a target is not known at that moment.

More interesting are the estimated p_{iD}^{-} values shown in the upper part of each panel in Fig. 6. The estimated termination probabilities show that all participants tend to use exhaustive or close to exhaustive search. However, only A tends to search exhaustively on all display sizes. In the $D = 2$ condition, participants C

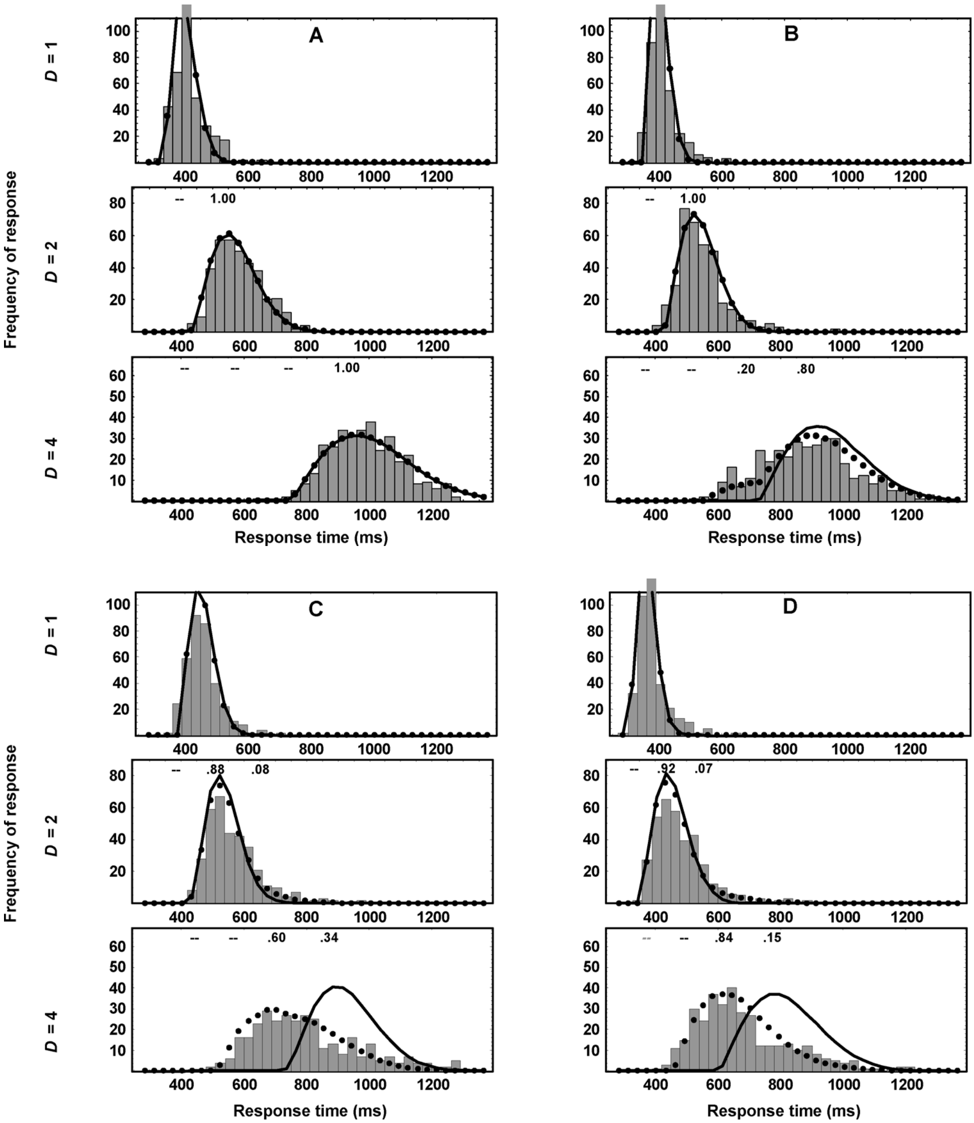

Figure 6. Response time distributions for correct rejections (correct target absent trials) for each participant (columns) and display sizes 1, 2 and 4 (rows) in the same format as in Fig. 5.

and D seem to make a small percentage of 'extra' scans. This is significant for both participant C ($p^+_{32} = 0.08$, $\chi^2(1) = 258.8$, $p < 0.01$) and participant D ($p^+_{32} = 0.07$, $\chi^2(1) = 294.8$, $p < 0.01$) (see Note 2). In the $D = 4$ condition, participants C, and D tend to stop most often after three scans.

A different way to display these results is in terms of 'efficiency', as given in the lower half of Table 4. A result of 1.0 indicates a perfectly exhaustive search, as is the case for participant A in both $D = 2$ and $D = 4$ conditions, and for participant B in the $D = 2$ condition. Participants C and D have an efficiency measure above 1 in the $D = 2$ condition, indicating that they occasionally performed more scans than needed. The efficiencies for participants B, C, and D when $D = 4$ are all well below 1.0, showing in another way that search terminates early. Once again these estimates of premature termination have implications for error rates, and we turn to this point now.

RT distributions of missed trials. According to one simple hypothesis, the premature stops are guesses. These guesses are perhaps made as a response to the instructions to respond 'as fast as possible without making many errors'. Guessing would produce both errors and accidental correct responses in proportion to their numbers at each scan number. The failure to find false alarms in large numbers seems to imply that guesses are usually 'target absent'. This guessing hypothesis has implications for RTs that are explored in Fig. 7, which gives the RT distribution for the erroneous target absent responses. Participant A made almost no misses, and her error data will not be discussed further. Participant B also made relatively few errors, making interpretation difficult. However, participants C and D made just enough errors (a little less than 50) to allow separate modes to be observed.

To assess whether there is a relation between the miss and possible guesses, the unconstrained model was fit to the missed data with the total proportion equal to the percent of miss in that condition. The remaining parameters were taken without modification from the correct rejection fits.

The estimated proportions (seen above the histograms in Fig. 7) show that for participants C and D in the $D = 2$ condition, as was the case for the correct rejections, three scans are sometimes performed when the target is missed. This occurred on about 2% of the target-present trials. This is significantly different from zero in both cases (participant C: $p_{32} = 0.03$, $\chi^2(1) = 94.3$, $p < 0.01$; participant D: $p_{32} = 0.02$, $\chi^2(1) = 77.2$, $p < 0.01$).

In the $D = 4$ condition, the misses mostly occurred during the third scan, although participant D also does misses on the second scan (significantly different from zero; $\chi^2(1) = 11.76$, $p < 0.05$). To estimate the amount of premature stops, the estimated p_{iD} must be weighted according to the number of trials not ended earlier with a target found response. For example, participant C does 5% of misses on the third scan. However, he finds the target on the first or second scan 72% of the times (see the percentages in Fig. 5). Thus, on the remaining 28% of the trials, he does a premature stop 18% of the times (0.05/0.28). Similar computations show that premature stops resulting in a missed target occurred on average on 20% of the trials. The exception is participant D, $D = 4$, second mode, which correspond to a 5% premature stop rate.

Summary of the correct rejection and the miss distributions. One interesting point to note is the similarity of the base response times in the $D = 2$ and $D = 4$ conditions. It suggests that the extra time found on the hit distributions

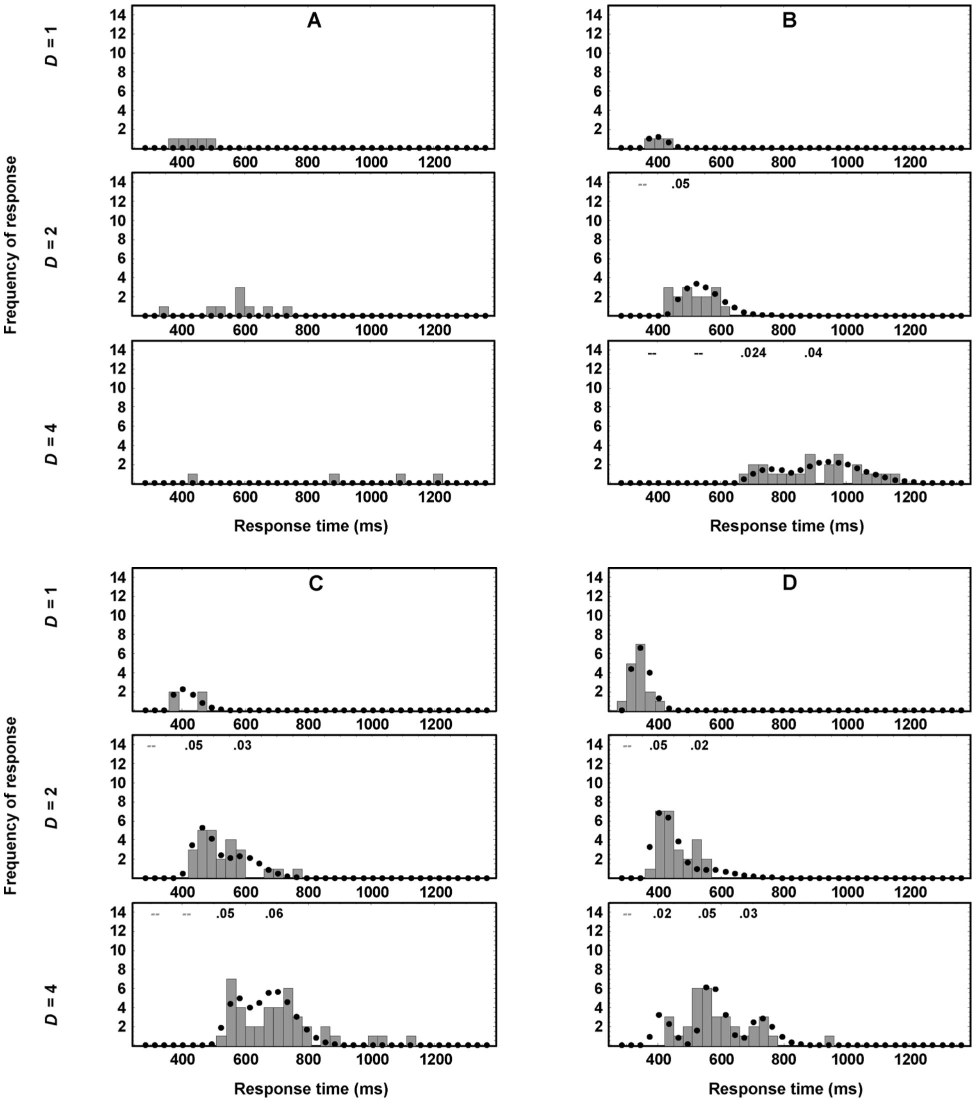

Figure 7. Response time distributions for misses (erroneous target present trials) for each participant (columns) and display sizes 1, 2 and 4 (rows) in the same format as in Fig. 5.

are latencies following the perception of the target. A similar assumption also improved the fits in Cousineau and Larochelle (in press), even though they only fit the mean and standard deviation of RT, not the entire RT distribution. It is tempting to draw a relation with attentional blink studies where the perception of the first target suspends processing of following targets for about 100 to 300 ms (Raymond *et al.*, 1992). Here, the extra processing time is smaller (the average difference between positive and negative residual time is 53 ms) but the participants are highly trained. Jolicoeur, Lefebvre and Cousineau (in preparation) found that the attentional blink effect was smaller after only four hours of practice. Recall that the present participants had from 35 to 44 hours of practice.

The fact that some subjects tended to perform three scans in the $D = 2$ condition and tended to finish with three scans as well in the $D = 4$ condition suggests that the termination process may have its own ballisticity. Once initiated ($D > 1$), it would tend to always perform the same number of scans.

The most important results of this section are that (in the $D = 4$ condition) i) when the target is present, the participants C and D guessed on 20% of the trials; ii) when the target is absent, they performed premature stops on nearly 70% of the trials. These percentages should match if the participants were performing pure guesses. This suggests that premature stops are above-chance guesses. This issue is discussed next.

GENERAL DISCUSSION

Let us recapitulate the termination rules adopted by the participants. Participant A is almost the perfect illustration of the self-terminating/exhaustive termination rule. Her hits exhibited D modes and her correct rejection, only one. Furthermore, her negative to positive ratio is exactly 2 : 1. Participant B, despite the difference in ratio (Fig. 4), differed only superficially from participant A. Her termination rule might best be described as a quasi-self-terminating process The hit distributions are similar to those of participant A, but she used premature stops, mostly before initiating a fourth scan in the $D = 4$ condition. The 20% of premature but correct 'target absent' responses found in this condition have a large effect on performance, reducing the mean 'target-absent' RT and consequently, the ratio. As seen, the incorporation of a few premature stops increases considerably the variance on target-absent trials. Some studies found that the 'target-present' responses have sometimes far larger variance than 'target-absent', as would be expected from a serial self-terminating search (Shiffrin and Schneider, 1977) but in other studies, the variances can be surprisingly similar (Treisman and Gelade, 1980; Ward and McClelland, 1989). If the amount of premature stops varies across studies, as it seems to vary across these two participants, the differences among studies would be understandable.

Participants C and D have a different distribution of hits. In the $D = 4$ condition for example, they are equally likely to find the target on the second scan than on any other scans. They also say 'target absent' on the third scan more often than on the fourth scans. Misses are also more frequent on the third scans. Likewise, these subjects performs a third scan in the $D = 2$ condition.

From these results, we may infer a few tentative conclusions: i) the search order may not be random, but informed with some efficiency; ii) the termination rule has important impact on the 'scanning rates' measured by RT slopes and on the negative to positive slope ratios.

Informed search

Deviations from a self-terminating search could be the result of a strategic speed-accuracy trade-off (Meyer *et al.*, 1988). If such was the case, the premature stops would be guesses, resulting in many errors. However, participant A performed a fourth scan in the $D = 4$ condition 15% of the times (compared to the expected one quarter of the times) and at the same time maintained a 1% error rate. Participant B did make more errors, but if the proportions of hits and of misses are added together, participant B has the same distribution of hits as participant A, still deviating significantly from the expected $1/D$. The deviations are even more pronounced for participants C and D, where more than 70% of the hits are performed on the first or second scans.

An alternative explanation more compatible with the data involves some form of parallel processing that guide the search to begin with those positions more likely to contain a target. This idea is consistent with the Guided Search model (Wolfe, 1994) where an activation map is built at the beginning of a trial. Likewise, it is possible that a parallel search operates concurrently with the serial search, the serial and parallel processes racing to locate a target (although this race model might have trouble with the modes lining up for different D values). In either case, the notion of a random search is inadequate. We rather need a notion of 'informed serial search'.

This conclusion is a bit challenging for classical models of visual search since the present stimuli were defined by a conjunction of features. Treisman and Gelade (1980; also Treisman and Sato, 1990) proposed the Feature Integration Theory (FIT, also see Wolfe, 1994, for a similar model). FIT assumes distinct maps that are feature detectors over the display. Each map individually can be accessed in parallel, so that searching for a unique feature is fast. However, binding features from more than one map requires attention, a slow process. Because attention is limited-capacity, multiple locations must be bound serially.

An alternative view to FIT is to be found under the generic term of noise models. The general idea is that accuracy decreases with stimulus difficulty not because capacity is limited but rather because noise is present on the processing pathways. For example, McElree and Carrasco (1999) showed that, with limited viewing time and varying a signal-to-response delay, accuracy reached a different asymptotic level depending on stimulus difficulty. This result cannot be explained by simply postulating limited capacity processing. Another example taken from cued detection task (Dosher and Lu, 2000) also suggests that noise is an integral part of the visual system and that there are mechanisms aimed specifically at reducing it (also see Eckstein, 1998; Palmer, 1998). Because noise is modeled as a random variable, it is always possible that some trials are perceived as 'clearer' than others. In such cases, a parallel, preattentive search would be possible, even for conjunctively defined stimuli, thereby preventing the use of a serial scan of the display. This preattentive mechanism could drive the search studied here, informing it of the most likely locations.

The impact of the termination rule

The termination rule adopted by the subjects threatens in profound ways the calculations that used to be performed on the data. For example, participants A and B were similar with respect to scanning rates (x_s were 190 and 180 ms per item respectively) and the distribution of hits. However, a look at the computed slopes for negative trials (shown in Fig. 4) shows that the values are the most different (196 *vs.* 120 ms/item respectively). This is in large part the result of the quasi-exhaustive termination rule chosen by participant B (see Note 3). Incidentally, this termination rule also reduces the observed 'target present' slope, although by a smaller value (about 10 ms per item for subject B). As a consequence, the ratios are not very diagnostic: Similar distributions of hits can produce very different ratios (compare participants A and B) whereas very different distributions of hits can produce similar ratios (compare participants B and D).

Many of the tests of serial *vs.* parallel processing assume a very specific termination rule (self-terminating *vs.* exhaustive). However, as was shown here, there are other termination rules. Further, this dimension may be independent of the scanning processing issue. The point is that the conclusions one would draw looking only at the slope ratios in Fig. 4 can be quite different from those drawn from analyses of the distributions in Figs 5 and 6.

It is not common to observe separate modes, and our findings probably result from the convergence of several factors, such as a very difficult search task that slows comparison time, a great deal of training, advance indication of the positions of the display items, and very reliable participants. In addition, the comparison times of participants A and B are slow enough that they could conceivably have made eye-movements from one display item to the next. However, the same unconstrained model (SSTS*) applied as well to participants C and D whose scanning rates are closer to what is found in the literature with difficult stimuli.

The deviations from the standard serial self-terminating search proved easy to fit with a mixture of distributions. Therefore the distribution of hits (summarized in Table 3) or the efficiency index (Table 4) could be used as a theoretical-free measure of search efficiency. A search where all the locations have equal probability would correspond to a zero-level of efficiency. Whether this measure is practical in all situations remains to be seen, but it would be a less controversial measure than the negative-to-positive slope ratio.

Acknowledgements

We would like to thank Lori Arnold, Judy Blackburn, Amy Criss, Alicia Justus and Peter Murray for their comments on an earlier version of this text.

This research was supported in part by funding from the Fonds de Recherche sur la Nature et les Technologie du Québec and the Conseil de Recherches en Sciences Naturelles et en Génie du Canada.

NOTES

1. The larger study involved additional conditions in which the stimuli were presented sequentially, at SOAs up to 50 ms per object or per feature. The research involving those conditions is being prepared for a much longer submission at a later time. All the conditions were mixed across trials, but the difference between the standard conditions we analyze in this article and the sequential conditions was not very evident to the participants. The reported that some trials appeared a bit 'fuzzy' or seemed to 'flicker' slightly.
2. Individual parameter comparisons reported in the text are performed using the Likelihood Ratio Text where the SSTS(*) model is compared with a nested model in which one parameter is set at a target value. The significance of the decrement in fit, penalized by 2, is tested on a χ^2 table with one degree of freedom.
3. The different residual base times also play a role. By subtracting the $D = 4$ mean RT from the $D = 1$ mean RT to compute the slope, they do not cancel out.

REFERENCES

Ashby, F. G., Tein, J.-Y. and Balakrishnan, J. D. (1993). Response time distributions in memory scanning, *J. Math. Psychol.* **37**, 526–555.

Bamber, D. (1969). Reaction times and error rates for "same"-"different" judgments of multidimensional stimuli, *Perception and Psychophysics* **6**, 169–174.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding, *Psychol. Rev.* **94**, 115–147.

Bozdogan, H. (1987). Model selection and Akaike's Information Criterion (AIC): The general theory and its analytical extensions, *Psychometrika* **52**, 345–370.

Bricolo, E., Gianesini, T., Fanini, A., Bundesen, C. and Chelazzi, L. (2002). Serial attention mechanisms in visual search: A direct behavioral demonstration, *J. Cognitive Neuroscience* **14**, 980–993.

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

Cousineau, D. and Larochelle, S. (1997). PASTIS: A program for curve and distribution analyses, *Behavior Research Methods, Instruments and Computers* **29**, 542–548.

Cousineau, D. and Larochelle, S. (2004). Visual-memory search: An integrative perspective, *Psych. Res.* (in press).

Cousineau, D., Goodman, V. and Shiffrin, R. M. (2002). Extending statistics of extremes to distributions varying on position and scale, and implication for race models, *J. Math. Psychol.* **46**, 431–454.

Cousineau, D., Brown, S. and Heathcote, A. (2004). Fitting distributions using maximum likelihood: Methods and packages, *Beh. Res. Meth. Instr. Comp.* (in press).

Cramér, H. (1946). *Mathematical Methods of Statistics*. Princeton University Press, Princeton, NJ, USA.

Dolan, C. V., van der Maas, H. L. J. and Molenaar, P. C. M. (2002). A framework for ML estimation of parameters (mixtures of) common reaction time distributions given optional truncation or censoring, *Behavior Research Methods, Instruments, and Computers* **34**, 304–323.

Dosher, B. A. and Lu, Z.-L. (2000). Noise exclusion in spatial attention, *Psychol. Sci.* **11**, 139–146.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing, *Psychol. Sci.* **9**, 111–118.

Feller, W. (1966). *An Introduction to Probability Theory and Its Applications, Volume II*. John Wiley, New York.

Hays, W. L. (1973). *Statistics for the Social Sciences*. Holt, Rinehart and Winston, New York.

Hockley, W. E. (1984). Analysis of response time distributions in the study of cognitive processes, *J. Exp. Psych. Learn. Mem. Cognit.* **10**, 598–615.

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

Huber, D. E., Shiffrin, R. M., Lyle, K. B. and Quach, R. (2002). Mechanisms of source confusion and discounting in short-term priming 2: Effects of prime similarity and target duration, *J. Exper. Psychol.: Learning, Memory and Cognition* **28**, 1120–1136.

Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations, *J. Exper. Psychol.: Human Perception and Performance* **20**, 1015–1036.

Lu, Z.-L. and Dosher, B. A. (1998). External noise distinguises attention mechanisms, *Vision Research* **38**, 1183–1198.

Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization.* Oxford University Press, New York.

McElree, B. and Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches, *J. Exper. Psychol.: Human Perception and Performance* **25**, 1517–1539.

Meyer, D. E., Irwin, D. E., Osman, A. M. and Kounios, J. (1988). The dynamics of cognition and action: mental processes inferrde from speed-accuracy decomposition, *Psychol. Rev.* **95**, 183–237.

Palmer, J. (1998). Attentional effects in visual search: relating search accuracy and search time, in: *Visual Attention*, Wright, R. D. (Ed.), pp. 348–388. Oxford University Press, New York.

Raymond, J. E., Shapiro, K. L. and Arnell, K. M. (1992). Temporary suppression of visual processing in a RSUP task: An attentional blink? *J. Exp. Psych. Hum. Per. Perf.* **18**, 849–860.

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

Sternberg, S. (1966). High-speed scanning in human memory, *Science* **153**, 652–654.

Sternberg, S. (1998). Inferring mental operations from reacton-time data: How we compare objects? in: *An Invitation to Cognitive Science: Methods, Models, and Conceptual Issues*, Scarborough, D. and Sternberg, S. (Eds), 2nd edn, pp. 365–445. Bradford Books, Cambridge, MA, USA.

Sternberg, S., Knoll, R. L. and Turok, D. L. (1990). Hierarchical control in the execution of action sequences: Tests of two invariance properties, in: *Attention and Performance XIII*, Jeannerod, M. (Ed.), pp. 1–56. Lawrence Erlbaum Associates, Hillsdale, NJ, USA.

Townsend, J. T. and Ashby, F. G. (1983). *Stochastic Modeling of Elementary Psychological Processes*. Cambridge University Press, Cambridge, UK.

Townsend, J. T. and Colonius, H. (1997). Parallel processing response times and experimental determination of the termination rule, *J. Math. Psychol.* **41**, 392–397.

Townsend, J. T. and Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories, *J. Math. Psychol.* **39**, 321–359.

Treisman, A. and Gelade, G. (1980). A feature-integration theory of attention, *Cognitive Psychology* **12**, 97–136.

Treisman, A. and Gormican, S. (1988). Feature analysis in early vision: evidence from search asymmetries, *Psychol. Rev.* **95**, 15–48.

Treisman, A. and Sato, S. (1990). Conjunction search revisited, *J. Exper. Psychol.: Human Perception and Performance* **16**, 459–478.

Van Zandt, T. (2000). How to fit a response time distribution, *Psychonomic Bull. Rev.* **7**, 424–465.

Ward, R. and McClelland, J. L. (1989). Conjunctive search for one and two identical targets, *J. Exper. Psychol.: Human Perception and Performance* **15**, 664–672.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search, *Psychonomic Bull. Rev.* **1**, 202–238.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychol. Sci.* **9**, 33–39.

Yantis, S., Meyer, D. E. and Smith, J. K. (1991). Analyses of multinomial mixture distributions: New tests for stochastic models of cognition and action, *Psychol. Bull.* **110**, 350–374.

Zenger, B. and Fahle, M. (1997). Missed targets are more frequent than false alarms: a model for error rates in visual search, *J. Exper. Psychol.: Human Perception and Performance* **23**, 1783–1791.