COGNITIVE PSYCHOLOGY **30,** 39–78 (1996)
ARTICLE NO. 0002

Just Say No: How Are Visual Searches Terminated When There Is No Target Present?

MARVIN M. CHUN AND JEREMY M. WOLFE

Massachusetts Institute of Technology and Harvard Medical School

How should a visual search task be terminated when no target is found? Such searches could end after a serial search through *all* items, but blank trials in many tasks are terminated too quickly for that to be plausible. This paper proposes a solution based on Wolfe's (1994) Guided Search model. The probability that each item is a target is computed in parallel based on items' differences from each other and their similarity to the desired target. This probability is expressed as an *activation.* Activations are examined in decreasing order until the target is found or until an *activation threshold* is reached. This threshold is set adaptively by the observer—more conservative following misses, more liberal following successful trials. In addition, observers guess on some trials. The probability of a guess increases as trial duration increases. The model successfully explains blank trial performance. Specific predictions are tested by experiments.

Suppose that you have written an important phone number on a small piece of paper. You are searching for that piece of paper among a mess of various articles, journals, forms, and other miscellaneous paperwork on your desk. When should one stop searching? Certainly, one can stop when the phone number is found, but if it is not found, determining how long to keep looking depends on other factors. One could perform a serial, exhaustive search, checking every sheet of paper in the office until the phone number is found or no papers remain unexamined. More reasonably, one could stop searching when no *likely* papers remain unexamined. For instance, if the phone number was written on a small piece of paper, then a more efficient strategy would be to search just through items of similar size.

Marvin M. Chun is from the Department of Brain and Cognitive Sciences; Jeremy M. Wolfe is from the Center for Ophthalmic Research, Brigham & Women's Hospital and Department of Ophthalmology. Portions of this work were presented as a poster at the 1991 Annual Meeting of the Association for Research in Vision and Ophthalmology (ARVO), Sarasota, FL. This research was supported by grant EY05087, and in part by grants T32GM07484, and EY06592 from the National Institutes of Health. We thank Kyle Cave, John Higgins, Lester Krueger, Gordon Logan, and Molly Potter for their helpful comments on earlier drafts of this work. We also gratefully acknowledge Stacia Friedman-Hill and Michelle Eng for their help in various phases of this project. Correspondence concerning this article should be addressed to Marvin M. Chun, Vision Sciences Laboratory, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138. E-mail: chun@isr.harvard.edu.

0010-0285/96 $18.00

This paper examines the termination of unsuccessful searches in a laboratory task corresponding to the previous example, the visual search paradigm. Visual search has been a very useful tool in the analysis of visual processing (see Grossberg, Mingolla, & Ross, 1994; Treisman, 1988; and Wolfe, 1994, for reviews). In a visual search task, an observer looks for a designated target item among a variable number of distractor items. Typically, reaction time (RT) measures are taken to determine the time required for an observer to respond ''yes'' when a target is present or ''no'' when it is absent. The number of distractors (the ''set size'') is the independent variable. The shape and slope of RT × set size functions are the dependent measures that change with different search tasks. These changes have been used to infer two stages of visual processing. Some searches, notably those for targets defined by a single, basic feature (e.g., color, orientation, size), produce RTs that are independent of set size for target-present and for target-absent (blank) trials. This is taken as evidence for a parallel stage that can process some aspects of visual input across large parts of the visual field at one time (Treisman & Souther, 1985). Other searches (e.g., a search for a ''T'' among ''L''s) produce a linear increase in RT with set size. Characteristically, these searches yield RT × set size slopes of 20–30 ms/item on target trials and twice that, 40–60 ms/item on blank trials (Treisman, 1988). Steep slopes and this 2:1 ratio of blank to target trial slopes are indicative of a serial, self-terminating search (Kwak, Dagenbach, & Egeth, 1991; but see Townsend, 1990, for a discussion of the fact that limited-capacity parallel searches could underlie such search results).

There are visual search tasks that produce results lying between the classic ''parallel'' and ''serial'' patterns. In what we will call *guided* searches, information from the parallel stage of processing is used to guide the subsequent, serial deployment of visual attention from item to item (Cave & Wolfe, 1990; Hoffman, 1978, 1979; Wolfe, 1994; Wolfe & Cave, 1989; Wolfe, Cave, & Franzel, 1989). Thus, in a search for a red vertical target, parallel color processes can guide attention toward red items while parallel orientation processes guide attention toward vertical items producing quite shallow slopes (Cohen & Ivry, 1991; Egeth, Virzi, & Garbart, 1984; Treisman & Sato, 1990a; Wolfe et al., 1989) even though no single parallel process is sensitive to the conjunction of color and orientation (Wolfe, Chun, & Friedman-Hill, 1995). The existence of guided searches raises the theoretical problem that is at the heart of this paper. How and when should an observer terminate a search when no target is present? In principle, the answer is easy enough for strictly parallel and serial searches. In a parallel search, all items are processed in parallel, allowing for an efficient decision about target presence or absence. In a serial, self-terminating search, the observer searches until she finds the target or, on blank trials, until she has exhaustively examined the entire set of items, one by one. The endpoint of an unsuccessful guided search is less obvious. A guided search is a serial search through a subset of the items.

How does the observer know that she has exhausted that subset? Why doesn't the observer search all items on blank trials?

DECISION PROCESSES IN VISUAL SEARCH

We can imagine two different strategies for deciding when a target is not present without having to exhaustively search the entire display. People may simply search through the distractors that have a certain likelihood of being a target and ignore those items which are less similar to the target. For reasons that will be made clear below, we will refer to this as the *activation threshold* hypothesis. Another possibility is that as an observer runs in a visual search task, she may develop some internal estimate of how long it takes to find a target. With such an estimate, she may be able to make ''educated guesses,'' since the probability of a guess being correct increases as evidence is accumulated as a search trial progresses. According to this *timing hypothesis,* observers will terminate a trial when the duration of the trial exceeds some duration threshold, based on the assumption that the target should have been found by then.

In this paper, we explore these two hypotheses in the context of the Guided Search model (Cave & Wolfe, 1990; Wolfe, 1994; Wolfe & Cave, 1989; Wolfe et al., 1989). Neither of the decision mechanisms we propose are entirely dependent on the specifics of the Guided Search model. However, Guided Search is sufficiently detailed to allow for specific, quantitative predictions to be tested. The Guided Search model holds that attention is guided toward candidate target items by parallel processes that *activate* items possessing one or more target features. Activation is combined across feature types so that an item having two target features will receive more activation on average than an item having only one such feature. For example, consider a search for a red vertical item among red horizontal and green vertical distractors. In this guided search for a conjunction, the parallel feature processor for color would activate all ''red'' items while the orientation processor would activate all ''vertical'' items. Information from these two feature modules would be combined into an overall *activation map.* Attention is guided from one candidate target to the next in a serial manner in order of decreasing activation. If this combination of information from the parallel stage processes were perfect, then the target, if present, would always receive the highest activation and search for a conjunction target would be no less ''parallel'' than search for a feature target. However, activations appear to be embedded in internal noise. The result is that an average target will have a higher activation than an average distractor but, on any one trial, some distractors may have higher activation than the target. As noted, in the Guided Search model, attention is deployed from item to item in decreasing order of activation strength. Thus, ''noisy,'' high activation distractors will have to be checked and rejected by the serial stage first before the actual target is found.

We do not have direct access to the thresholds and decision rules proposed here but we can predict their impact on RTs and error rates. The paper is organized in the following manner: First we describe the proposed activation threshold mechanism in more detail and present some simulation results based on theory that the activation threshold alone accounts for blank trial performance. These simulation results deviate from the data in the literature, so we next incorporate the timing/guessing mechanism into the model. This generates better simulation results and a set of concrete predictions. Finally, these predictions are tested empirically in a series of three visual search experiments. The results are consistent with the model.

The Activation Threshold Mechanism

What happens on blank trials in a guided search? Clearly, it would be inefficient to search exhaustively through the entire display, since the average activation of the target will lie above the average activation of the distractors (otherwise there is no guidance). It should be possible to safely reject *some* distractors in parallel on the basis of their low activations. We propose that an internal *activation threshold* is used as a cutoff criterion for terminating search on blank trials. Setting the correct value for this activation threshold is essentially a signal detection problem (see Pavel, Econopouly, & Landy, 1992). This is illustrated in Fig. 1 which shows hypothetical average activation distributions for distractors and for targets compiled *across* trials for two visual search tasks. The activation of the distractors is modeled as a normal distribution with some mean and standard deviation. The distribution from which the target activation will be drawn on any given trial is simply this distractor distribution with an activation ''signal'' added to it. RTs for target and blank trials can be derived graphically from Fig. 1. On a target trial, observers will have to examine all distractors with activations above the target activation value. On average, the proportion of distractors examined will be those falling in region ''C,'' the region of the distractor distribution above the average target activation. On blank trials, the activation threshold hypothesis simply states that observers examine all distractors above an activation threshold. Thus, the proportion of distractors examined on an average blank trial will correspond to the area of the distractor distribution in regions B and C in Fig. 1. If the activation threshold is set to a level higher than the lowest *target* activation, then on some target trials, search will be terminated before the target is found. The target will be missed and a *miss* error is generated. The proportion of *target* activations that gives rise to miss errors will correspond to region A. If the activation threshold is set higher, fewer distractors will be checked on blank trials, reducing blank trial RTs. However more misses will be produced; a classic speed-accuracy trade-off. If a different task produces a greater signal (Fig. 1b), fewer distractors will lie above either the average target activation or the activation threshold and, thus, target and blank RTs will decline.

The distributions in Fig. 1 are fine theoretical constructs but it is implausible

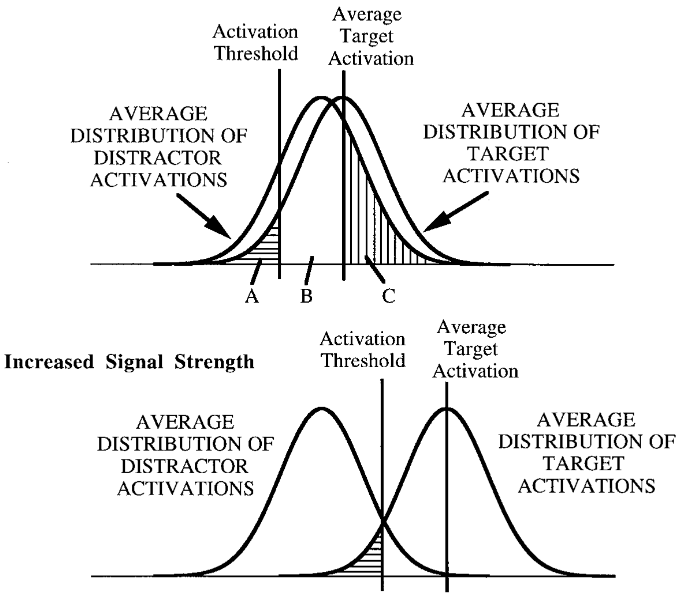

FIG. 1. Visual search can be conceived of as a signal detection problem where each distractor item has an activation drawn from some noise distribution and where the target has an activation drawn from that distribution with a signal added to it. Average target-present RTs correspond to the proportion of distractors with activations above the average target activation (area C). In the present model, blank trial RTs correspond to the proportion of distractors above the activation threshold (areas B + C) and miss rate corresponds to the proportion of targets with activations below the activation threshold (area A).

to assume that observers have any conscious or unconscious access to the precise shape of the distributions or to the magnitudes of activations. How, then, is the observer to set an activation threshold that will produce an acceptable error rate? We have modeled this setting of the threshold as an internal ''staircase'' procedure. In visual search tasks, the observer is usually instructed to respond as fast as possible while making few errors (5–10%). According to the model, the observer keeps the RTs and the error rate at the desired levels as follows: When the observer correctly terminates a blank trial search, he raises his activation threshold in order to terminate the next blank trial more quickly and improve his overall speed. When the observer commits an error, he lowers his activation threshold in order not to miss future targets with low activation, thus preserving an acceptable error rate. The specific error rate can be controlled by varying the relative size of the ''up'' and ''down'' staircase ''steps'' (Levitt, 1971). For example, if the step ''down'' following a miss is 20 times the size of the step up following a successful blank trial, then the staircase converges on an error rate of approximately 4% (see simulation below).

Blank-Target Slope Ratios: A Retrospective Analysis

In the signal detection scheme outlined above and illustrated in Fig. 1, the ratio of blank trial slopes to target trial slopes is determined by the ratio of areas B + C to area C. For standard serial, self-terminating slopes, that ratio would be 2:1. Since this paper is intended to model the full range of visual search tasks from parallel through guided to serial searches, it is important to know the normative slope ratio for a range of search tasks. To obtain this information, we retrospectively examined slope ratios for a wide range of search tasks. To do this we took 733 pairs of target and blank trial slopes from 65 different search experiments run in this laboratory over a period of 4 years. Each pair represents 300 trials run on one observer. Most observers are represented several times because they were tested on several search tasks but at least 100 individuals are represented. Search tasks include simple feature searches, complex (''serial'') feature search, easy and hard conjunction searches, and ''serial'' searches (e.g., ''T'' among ''L''s). The slowest searches are those for conjunctions of two colors or two orientations (Wolfe et al., 1990) and some complex orientation searches (Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). Most of the data has been reported in previous work from this laboratory (cf., Wolfe, 1994).

The results of this analysis are shown in Fig. 2. The main figure shows all data with target slopes between 0 and 60 ms/item and blank slopes between 0 and 150 ms/item. The inset shows all slopes. The solid line is the 2:1 ratio line and *not* the regression line. It is clear that a 2:1 ratio is a reasonable description of the main trend in the data. For the entire data set, the regression slope is 1.99 with an R^2 value of 0.77. Removing the highest slopes produces a slope of 1.70. The deviation from 2.0 seems to reflect the influence of points near 0.0 where the ratio becomes meaningless. If, for example, we examine the 168 points with target slopes between 5 and 12 ms/item (a reasonable definition of ''guided'' searches), the regression line has a slope of 2.04, though R^2 is much reduced (0.12). The purpose of this analysis is to show that a 2:1 slope ratio is the most reasonable fit to the full range of various search tasks that have been tested in this laboratory. This is *not* meant to imply that every search task will show 2:1 slope ratios. Indeed, as shown in Fig. 2, there is considerable variation of slope ratios. Breaking down the analysis by search type (e.g., feature search vs conjunction search) does not lead to any systematic variation between search task and slope ratio. In sum, for general modeling purposes, we choose to simulate the search task to approximate the normative 2:1 slope ratio that best fits the entire range of data.

The 2:1 ratio raises a problem. Returning to Fig. 1, we obtain a 2:1 slope ratio when area B equals area C. This may be the case for one signal strength, but as the distribution of target activations slides left and right with changes in signal strength, the ratio of area B to area C will change. Specifically, as

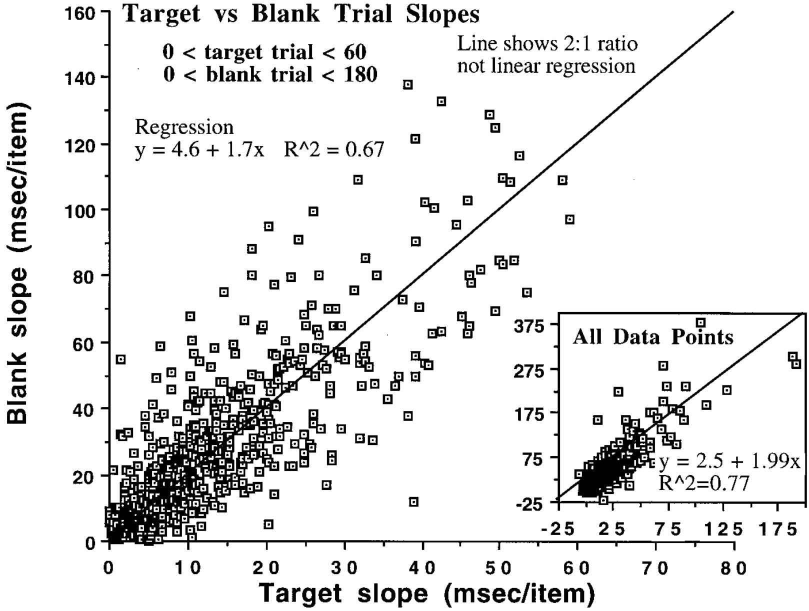

FIG. 2. Blank trial slope as a function of target trial slope. Each of the 733 points represents one observer performing 300 trials on one visual search task. Over 70 different search tasks are represented. Though there is substantial scatter, it is clear that a 2:1 slope ratio is a good approximation to the data over the full range of search tasks.

signal strength increases, area C will decrease faster than area B as can be seen in Fig. 1. If the activation distributions for targets and distractors were normal and had the same standard deviation, the slope ratio should increase systematically with increasing signal strength. As discussed above, in the actual data, there is no evidence for any such systematic deviation change in slope ratios.

There are, no doubt, a host of possible solutions to this problem. The solution used in Guided Search 2.0 (Wolfe, 1994) is illustrated in Fig. 3. If the targets are drawn from a distribution whose standard deviation *declines* as the signal strength *increases*, then the distance, in activation units, between the average target activation and the activation threshold decreases with increasing signal strength and an average 2:1 ratio can be maintained.

Though we do not know the neural correlates of target and distractor activations, the idea that target activations may become more precise as signal strength increases seems plausible. One scenario that we have simulated could be labeled ‘‘neuron recruitment.’’ Suppose that each ‘‘neuron’’ in the system has the same variability. The number of neurons activated by an item is directly related to the signal strength of that item: more signal, more neurons. The overall activation for an item is the average of the activated neurons. By

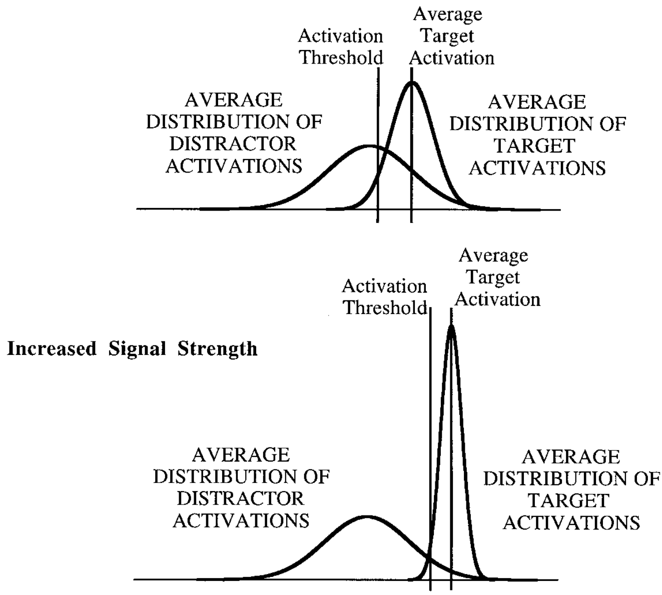

FIG. 3. Target and blank trial slopes are in a 2:1 slope ratio when areas B and C of Fig. 1 are equal. This relationship will not be maintained as signal strength changes if distractor and target distributions are normal and have the same variance. One way to preserve the relationship is shown here. Reduce the variance of the target distribution as signal strength increases.

the law of large numbers, an increased number of activated units with unit variance results in smaller variability in their averaged output. Thus, stronger signals have smaller variance (see Zohary, Shadlen, & Newsome, 1994). In the full Guided Search 2.0 simulation of Wolfe (1994), this scheme produced results that were in good qualitative agreement with human data. In the more limited simulation reported in the remainder of this paper, we manipulate the variance directly with a function that reduces the variance as signal strength increases and that yields the approximate 2:1 slope ratio seen in the human data. Details are discussed below.

Simulating the Activation Threshold

In the simulation, each trial was modeled as a list of activations. Distractor activations were generated from a normal population with a mean of 300 and a standard deviation of 100 arbitrary units. On half the trials, one distractor was replaced by a target item. For any block of trials, this target was drawn from a distribution with a mean defined by the size of the signal and a standard deviation calculated to satisfy two constraints: First, the activation threshold must be set so that the proportion of distractor activations lying between the activation threshold and the target average will be equal to the proportion of distractor activations above the target average (areas B and C in Fig. 1). This will yield a 2:1 slope ratio. Second, the activation threshold must lie about 1.7 standard deviations below the target average on the target distribution. This will yield an approximate 4% error rate (area A in Fig. 1). Different error rates can be produced by changing the distance from target average to activation threshold. Given these constraints (and a tabulated normal distribution), the required standard deviation of the target distribution can be determined for any signal strength. For computing purposes, relationship of the standard deviation of the target distribution to the signal can be well approximated by

$$SD_{\text{target}} = SD_{\text{distractor}} \times \frac{1}{1 + (2 \times Z_{\text{signal}})} \tag{1}$$

Where Z_{signal} is the Z-score of the signal in the distractor distribution. When there is no signal, SD_{target} equals $SD_{\text{distractor}}$. As the signal grows, SD_{target} decreases, going to zero in the limit. (Note that this equation has no theoretical implications of its own. It represents pure curve fitting.)

For each trial in the simulation, set size was randomly set to 4, 10, or 16. RT was determined by counting the number of distractors with activations above the *target activation* for target trials and by counting the number of distractors with activations above the *activation threshold* for blank and miss (error) trials. These counts were converted to RTs in milliseconds by multiplying by 50 ms, an estimate of the amount of time required to process a single item. Four hundred fifty milliseconds were added to each simulated trial RT as a constant to reflect the time required to make a response.

The activation threshold was initially set to the mean of the distractor distribution. It was adjusted from trial to trial in a staircase manner. If the simulated observer correctly terminated a blank trial, the threshold was increased by one step (5 arbitrary units of activation). If the simulated observer made an error, the threshold was decreased by k steps, where ‘‘k’’ is the parameter that determines the error rate. For instance, a ‘‘k’’ of 20 steps (100 units) yields an error rate of about 4%. The threshold was not changed on successful target-present trials. Thus, for any run of the simulation for the activation threshold, the relevant parameters were signal strength and the staircase parameter. We are not proposing that real observers have any specific notion about these staircase steps. The step size is merely a way to express the real observer’s automatic effort to set an appropriate activation threshold. Our implementation of this is a computationally simple approximation to the idea of incrementing the threshold in little steps with each successful ‘‘no’’ response and of decrementing the threshold in big jumps with each ‘‘miss’’ error. Successful blank trials should suggest to the observer that the search could be terminated more quickly and that the threshold could rise. Misses demonstrate that a search was terminated too quickly and that the threshold should be lowered. The size of the drop in threshold is determined by the

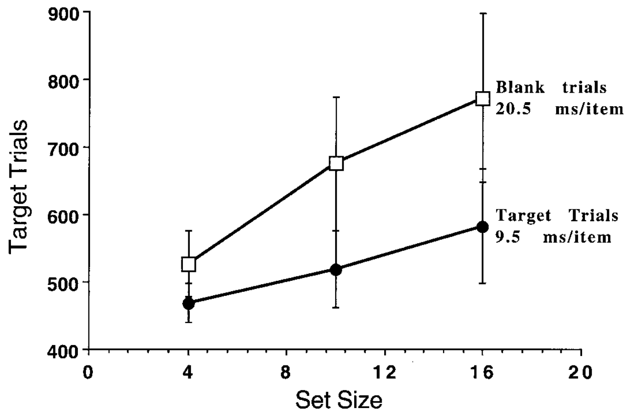

FIG. 4. Simulated data for a visual search with a signal strength of 100 and an error rate of 4%.

observer's tolerance for errors. Lower tolerance should yield bigger step sizes. Note that the observer does *not* need to keep track of percent error. He merely reacts to the feedback from each trial with a lesser or greater degree of caution.

We first present the simulation results for the activation threshold mechanism alone. Fig. 4 shows average RT as a function of set size for 600 trials on one simulated observer. The staircase parameter is 20 and the resulting miss error rate is 4.3%. These results would be typical for a ''guided'' search (e.g., conjunction of color and orientation (Treisman & Sato, 1990b; Wolfe et al., 1989). Note also that the variability increases with set size and is greater for blank trials than for target trials, an attribute of real search data (Ward & McClelland, 1989) not accounted for in the earlier versions of the Guided Search model.

The way in which the simulation achieves the results in Fig. 4 is illustrated in Fig. 5. Figure 5a shows the activation distributions for targets and distractors that underlie the performance shown in Fig. 4a. Average simulated target activation is 396 with a standard deviation of 31. Average distractor activation is 292; the standard deviation equals 99. 15% of the distractors are above the average target activation and will need to be checked on an average target trial. If a full serial examination of all items yields 50 ms/item, then a search through 15% of the items would yield a slope of 7.5 close to the obtained regression slope of 9.5 ms/item, as in Fig. 4a. Forty-two percent of distractors are above the average activation threshold of 313 yielding a blank trial slope of about 21 as in Fig. 4a. The average activation threshold for trials in which the target was missed was 366. This was close to the theoretical threshold value of 343 which would lead to a predicted error rate of 4%. Figure 5b simply illustrates the continuous adjustment of the threshold over

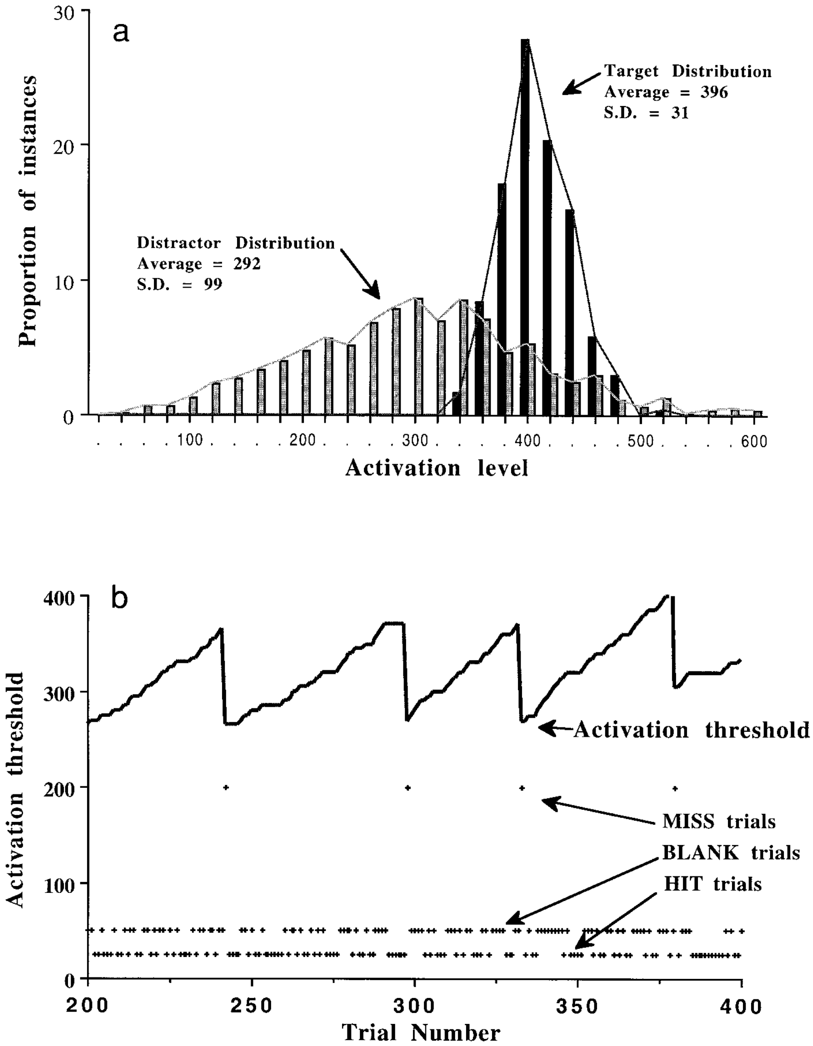

FIG. 5. (a) shows the activations for targets and distractors underlying the simulated data plotted in Fig. 4a. (b) gives a record of the changes in the activation threshold over 200 trials. Miss error trials are accompanied by sharp decreases in the activation threshold. The threshold gradually increases with each succesful blank trial.

200 of the 500 trials and its relationship to the trial type. The threshold goes up a little when blank trials are correct, down a lot when target trials are missed.

The threshold setting staircase parameter determines how blank trials are terminated. Simulated results for systematic variation of this parameter are

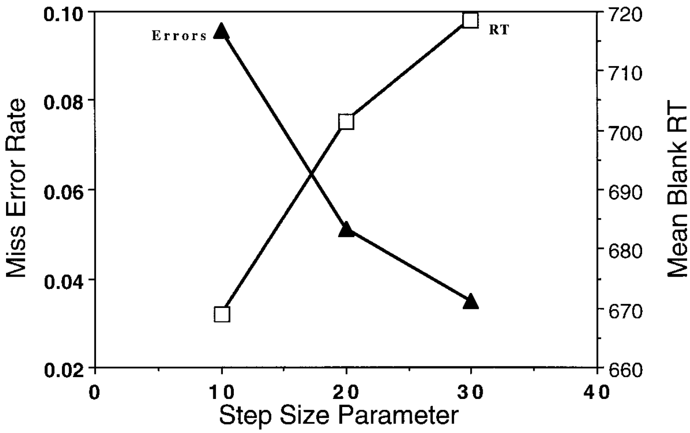

FIG. 6. Simulation results showing the effect of the step size parameter on Blank trial RT and miss error rate.

shown in Fig. 6. Increasing the staircase ''step size'' decreases miss errors and increases blank trial reaction times—a classic speed-accuracy tradeoff. The simulation shows the results for six ''observers'' averaged over runs at each of three different signal strengths (0, 100, 200). A more liberal criterion (smaller step size) means fewer distractors are checked. Blank trial searches terminate more quickly at a cost of more miss errors.

Interrupt (Timing) Mechanism: The Problem of False Alarms

The activation threshold hypothesis provides a concise explanation for how blank trials are terminated. It makes adequate predictions regarding blank trial RTs and miss error rates. However, the activation threshold can say nothing about false alarms. While false alarms typically occur less frequently than miss errors, the presence of false alarms require some explanation within the context of any model of visual search. We propose that false alarm errors are typically produced by an ''educated guessing'' strategy based on timing estimates. That is, due to miscellaneous factors such as boredom, fatigue, frustration, anticipation, etc., there may be a small proportion of trials where observers simply terminate a trial with a guess. These guess responses will be correct half of the time probabilistically, allowing the observer to terminate a trial more quickly. Moreover the probability of a guess being correct increases as time elapses within a search trial. Suppose an observer has determined through practice that he usually takes less than 1000 ms to find a target. If more than a second has elapsed on a certain trial, the observer may respond ''no'', guessing that it was likely to have been a blank trial since the target had not been found by then. Observers may also incorrectly guess ''yes'' on a proportion of trials, producing a handful of false alarms.

A version of this timing hypothesis was previously offered as the primary mechanism for terminating blank trials in an earlier version of Guided Search (Wolfe & Cave, 1989). According to this account, an observer could use an internal timing threshold to terminate a search trial. It would be based on some implicit knowledge of the average time it takes to find the target and the variance of these ‘‘yes’’ responses. This timing threshold can be estimated from the distribution of an individual’s response time distributions for target-present responses. The logic is analogous to that shown in Fig. 1 of the activation threshold. However, pilot analyses of reaction time distributions from data obtained in our lab suggested that observers do not employ a timing threshold as the primary mechanism for terminating blank trials. In our present model, we incorporate a version of the timing hypothesis into our interrupt ‘‘guessing’’ mechanism so that the probability of making a guess increases as the trial duration increases. This is a plausible assumption given that the probability of a guess being correct increases as evidence accumulates during a search trial.

We model the interrupt mechanism as a simple random guessing strategy in the simulation. The probability of the simulation for making a guessing response is controlled by a single guessing parameter, *g*. On any given trial, for any given ‘‘*g*,’’ the likelihood of making a guess response increases with reaction time. On trials where a guess response was generated, we implemented the probability of guessing ‘‘no’’ to be 80% and guessing ‘‘yes’’ to be 20%.

A schematic processing flow diagram of the full model is illustrated in Fig. 7.

Full Simulation

A complete simulation having both activation threshold and interrupt components was run to generate joint predictions for reaction time and error rate across a range of visual search tasks. Thus, we varied signal strength, the stepsize parameter, and the guessing rule parameter. We tested signal strengths 0, 100, and 200, to model the full range of visual search tasks ranging from serial to guided (conjunction) to feature search. Staircase step sizes were 10, 20, and 30. The guessing parameter was 1 or 2 (Increasing the *g* parameter increases the likelihood of guessing). Each parameter was crossed with each other in a factorial design. Four simulated ‘‘observers’’ were run at each combination of signal strength, step size, and guessing parameter.

The results of this full simulation are shown in Fig. 8. Figure 8a shows that both target and blank trial slopes increase systematically with decreasing signal strength. Target trial slopes were 22 for a signal of 0 (serial search), 8 for a signal of 100 (guided-conjunction search), and 1 for a signal of 200 (‘‘parallel’’-feature search). The corresponding error rates in Fig. 8b illustrate that the proportion of miss errors and false alarms both decrease as the signal strength gets larger. At signal strength 0, which roughly corresponds to a

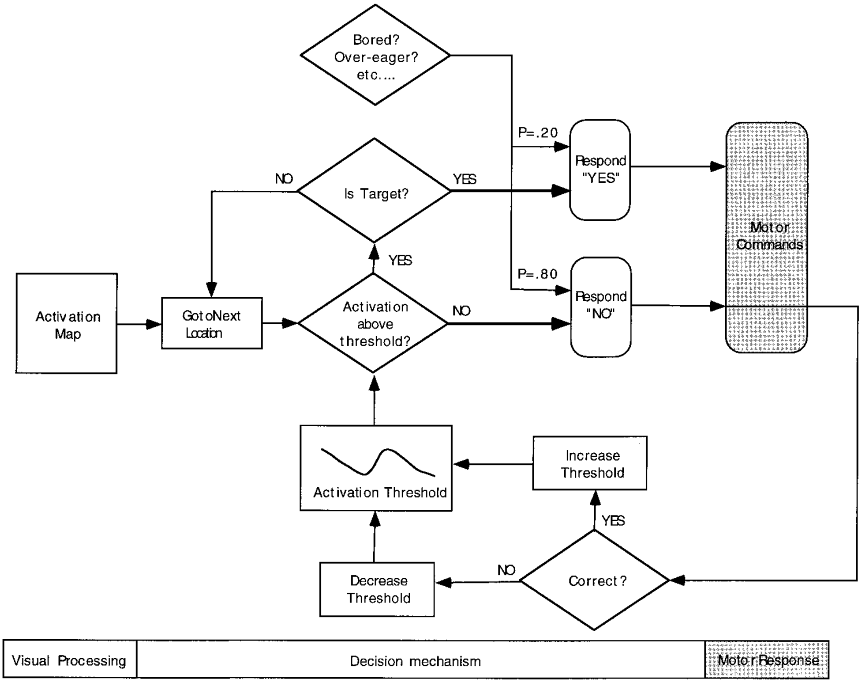

FIG. 7. A schematic architecture of the ''Just Say No'' decision model. An array of activations produced by earlier visual mechanisms serves as the input to the module which generates a ''yes'' or ''no'' response as output.

standard serial search task, miss errors and false alarm errors increase most dramatically with set size, a pattern observed in human data. Miss error rates also increase with set size in the signal strength 100 condition, as is typical in Guided Search conjunction tasks.

Predictions

The model and full simulation can be used to generate specific predictions about reaction times and error rates in visual search experiments:

1. After a miss, RTs for blank trials should increase markedly while RTs for target trials remain unchanged. Note that this is different from the usual informal assumption that observers are simply abnormally slow on the trial after an error. This assumption lies behind the custom in some labs of discarding the trial immediately following an error.
2. Manipulation of observers' error tolerance should influence blank trial slopes but have little effect on target trial slopes.
3. Manipulation of an observer's tolerance for error and/or priority on

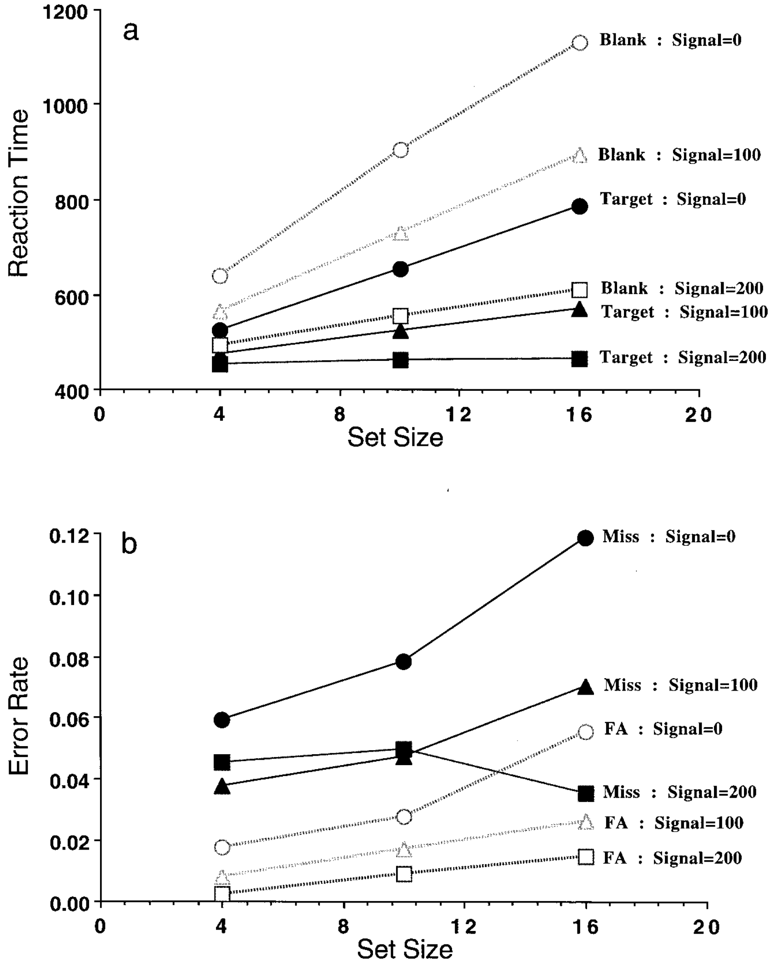

FIG. 8. RT and Error rate as a function of signal strength. The two other free parameters of the simulation (step size and guessing) were varied systematically in a factorial design. Data shown here was averaged across these parameters. The greater the signal, the shallower the slope of the RT × set size function for both target and blank trials. Miss errors occur more frequently than false alarms. Both misses and false alarms show a set size dependency as signal strength is decreased (increased task difficulty).

speed should reveal speed/accuracy tradeoffs *within* that observer's data. Across observers, it is harder to predict the relationship between slope and error rate since, for example, an observer with a low tolerance for error (slows blank RTs) might also have a high signal strength (speeds blank RTs).

4. An RT distribution for ''no'' responses can be obtained from each observer. Analysis of the placement of ''misses'' within that ''no'' distribu-

tion should show higher proportions of miss errors at fast RTs compared to slow RTs because misses will occur on those trials when the activation threshold has been set too high and search terminates too quickly.

5. If we assume that average activation declines as a function of eccentricity, then both RTs and errors should increase with eccentricity (cf., Useful Field of View: Ball, Beard, Roenker, Miller, & Griggs, 1988; Ball, Roenker, & Bruni, 1990; Sekuler & Ball, 1986).

6. If observers terminate trials when too much time has elapsed, then there should be elevated guessing responses for tasks producing longer RTs. This predicts more errors at larger set sizes for more difficult search tasks, but not for efficient feature search tasks.

In the experimental portion of this paper, we compare the results of our model of search termination to empirical data. In Experiment 1, we examine the dynamic properties of the threshold during a single block of trials. As predicted, blank trial RTs go down as long as the observer is responding correctly and rise following each error. Target trials, not dependent on the activation threshold, are relatively unaffected. RTs are slower and miss error rates are higher for peripheral targets than for targets appearing near fixation. Analysis of within observer blank trial RT distributions also reveals higher miss error rates among fast responses. In Experiment 2, we varied the observer's payoff matrix to demonstrate how the resulting speed/accuracy tradeoffs fit with the predictions of the model. To examine whether blank trial termination could be based solely on a timing strategy (interrupt mechanism), we mixed trials with different sets of distractors in a single block in Experiment 3. Blank trial RTs were different for different sets of distractors suggesting that a timing strategy alone would be insufficient for explaining how blank trials are terminated.

EXPERIMENT 1

Dynamic Thresholds

In the first experiment, observers were run in two different search tasks. By varying the number of features the target shares with the distractors in the two tasks, we could manipulate the difficulty of the search (corresponding to signal strength in the model). We computed a trend analysis of the reaction time data to look for effects of our hypothesized activation threshold. We also present an analysis of the individual reaction time distributions obtained from this experiment. The pattern of results was consistent with the predictions of a dynamic activation threshold.

Method

Participants. Sixteen observers were tested. They were recruited from Harvard Medical School, Northeastern University, Massachusetts College of Art, and the Massachusetts Institute of Technology. All observers were paid for their participation. They all had normal or corrected-to-normal acuity and were also tested for color blindness. All were naive to the purposes of this experiment.

Apparatus and stimuli. Stimuli were presented on a standard television monitor that was part of a modified ‘‘Sub-Roc 3-D’’ video game. Displays were controlled by an IBM PC-XT with IBM-YODA graphics. Stimuli were saturated red and blue X’s and O’s on a black background (CIE, International Commission on Color, x, y coordinates: red, .62, .36; blue, .14, .07). Observers viewed an 11.3 degree by 11.3 degree field bounded by a green square. Observers fixated a small green central point. Individual items fit within an 0.85 degree by 0.85 degree square. They could be placed at any of 36 locations in a slightly irregular 6 by 6 array. For the entire experiment, a single set size of 25 was used. On each trial, the items were presented on 25 randomly chosen loci within the 36 possible positions in the array. On 50% of the trials, one of these locations contained the target. Positions of targets and distractors, presence or absence of a target was random across trials.

Observers searched for targets defined by a triple conjunction of color, form, and size. Each of the observers was tested in two conditions. The Easy Condition was a triple conjunction task where each distractor shared one feature with the target. In the more difficult, Hard Condition, two thirds of the distractors shared two features with the target. In the Easy Condition, observers were instructed to look for a *small red O* among the three types of distractor items: Big *red* X; Big blue *O*; and *Small* Blue X (Big stimuli measured .80 by .80 degrees in visual angle, small stimuli measured .45 by .45 degrees). Each distractor shared one feature with the target. The Hard Condition was a double feature, triple conjunction task. The target was a *small blue O* which shared two features with two of the previous distractors big *blue O* and *small blue* X, and shared no features with the big red X. All distractor items appeared with equal probability in both conditions. Note that the same set of distractors were used in both conditions. Thus the physical stimuli for blank trials were the same for both tasks. Any differences in blank trial RTs between the two conditions must reflect a difference in the proportion of the display being searched and cannot be attributed to physical differences in the visual displays. Previous studies have shown that triple conjunction tasks with distractors that share two features with the target exhibit steeper search slopes similar to those found in standard conjunction tasks. When the target shares only one feature with the target, search slopes are shallow (Quinlan and Humphreys, 1987; Wolfe et al., 1989).

Procedure and design. Observers initiated each trial by pressing a button with the left hand, causing the search display to appear immediately. The search display remained on the screen until the observer made a response. Observers responded by pressing one of two buttons with the right hand depending on target presence or absence. Reaction time was measured from

TABLE 1
Results of Experiment 1

Condition	Reaction time (ms) Target trial mean (*SD*)	Reaction time (ms) Blank trial mean (*SD*)	Error rate (%) Target trial (MISS)	Error rate (%) Blank trial (False alarm)
Single feature conjunction	663 (180)	836 (236)	3.36	1.26
Double feature conjunction	775 (225)	1098 (292)	8.01	1.61

stimulus onset to the response of the observer. Observers were given feedback for correct and incorrect responses.

Each condition was run in a single block of 330 trials, with the first 30 trials discarded from the data analysis as practice. Each observer was run in both conditions, with the order of the two blocks counterbalanced across the 16 observers.

Results and Discussion

Average Results

Table 1 shows average RTs and error rates for both conditions. There was a main effect of search task condition, $F(1,15) = 23.75$, $p < .001$, replicating previous studies showing that a triple conjunction task is significantly faster when the distractors share only one feature with the target (Easy Condition) than when they share two (Hard Condition). The target/blank reaction time difference was highly significant, $F(1,15) = 35.26$, $p < .001$, as was the interaction between condition and target/blank trial reaction time, $F(1,15) = 19.30$, $p < .001$. Observers made significantly more errors in the Hard Condition, $F(1,15) = 50.83$, $p < .001$. More miss errors were made than false alarms across the two conditions, $F(1,15) = 41.67$, $p < .001$, and the interaction between Condition $\times$ Error type was significant, $F(1,15) = 19.56$, $p < .001$. This is consistent with our model's assumption that miss errors occur more frequently than false alarms in visual search tasks, and that the pattern of miss errors would be more sensitive to various task manipulations. Also the variance on blank trials is larger than the variance on target trials. This replicates Ward and McClelland's (1989) blank trial data and argues against any simple, exhaustive serial self-terminating search strategy through a fixed subset of the blank trial display. A simple exhaustive search model would predict lower variance because the same number of items would be examined each time.

Reaction Time and Activation Threshold

The raw trial data from this experiment provides a database of reaction times which can be examined for evidence that observers were changing an internal activation threshold in a manner analogous to the simulation described above. On-line adjustment of an activation threshold should be reflected in blank trial RTs. Faster blank trial RTs indicate that fewer items were being checked. Slower blank trial RTs suggest a more conservative threshold value resulting in more extensive checking of display items. To look for evidence of threshold adjustment, we examined the trials surrounding miss errors. Averages of blank trial RTs were computed for the five successful blank trials preceding and following a miss. If a miss was followed by another miss within five successful blank trials, only those blank trials within the interval between the two misses were included in the analysis for each miss error. Each observer's data was computed as a deviation from her average blank trial RT and these normalized RT differences were then averaged across observers. The results of this analysis for the Hard Condition are shown in Fig. 9a. As predicted by the model, observers become more liberal and, thus, faster over time until the error is made. After the error, we assume observers become more conservative, the threshold is raised, and RTs take a sharp jump up. For comparison, Fig. 9b shows the same analysis carried out on simulated data. The shape of the results is clearly similar (if less noisy) in the simulated data.

Similar results have previously been reported for errors and responses following errors in continuous-performance choice-response tasks (Rabbitt, 1966, 1967). Rabbitt (1966) showed that in these tasks RTs for errors were faster than correct responses, and that the first correct response following an error-correcting response (observers were required to immediately correct any errors they made) was slower than other equivalent correct responses. These results were confirmed in a separate study where the first two correct responses following an error were shown to be slower than correct responses not preceded by an error (Rabbitt, 1967). He discusses that errors may be made when observers try to respond faster than their maximum rate of gain of information permits. These speed/accuracy tradeoffs in a choice response to a signal are comparable to the pattern of fast errors and subsequent RT slowing we showed for responses to target absent displays. In departure from Rabbitt's general findings however, our activation threshold model predicts that RT slowing would be restricted to responses for blank trials. This was confirmed in the following analysis.

A central prediction of the model is that errors should have their effect on blank trial RTs but not on target trial RTs. In order to test this hypothesis, 6 conditional RTs were computed. Three of these were conditionalized RTs for blank trials. These were the average RTs for: (1) a blank trial immediately preceded by a blank trial (neg_NEG); (2) a blank trial immediately preceded by a correct hit (hit_NEG); and (3) a blank trial immediately preceded by a miss error (miss_NEG). The remaining three were the analogous conditionalized RTs for target trials: (1) neg_HIT; (2) hit_HIT; and (3) miss_HIT. Each of these was expressed as a deviation from the unconditionalized average RT

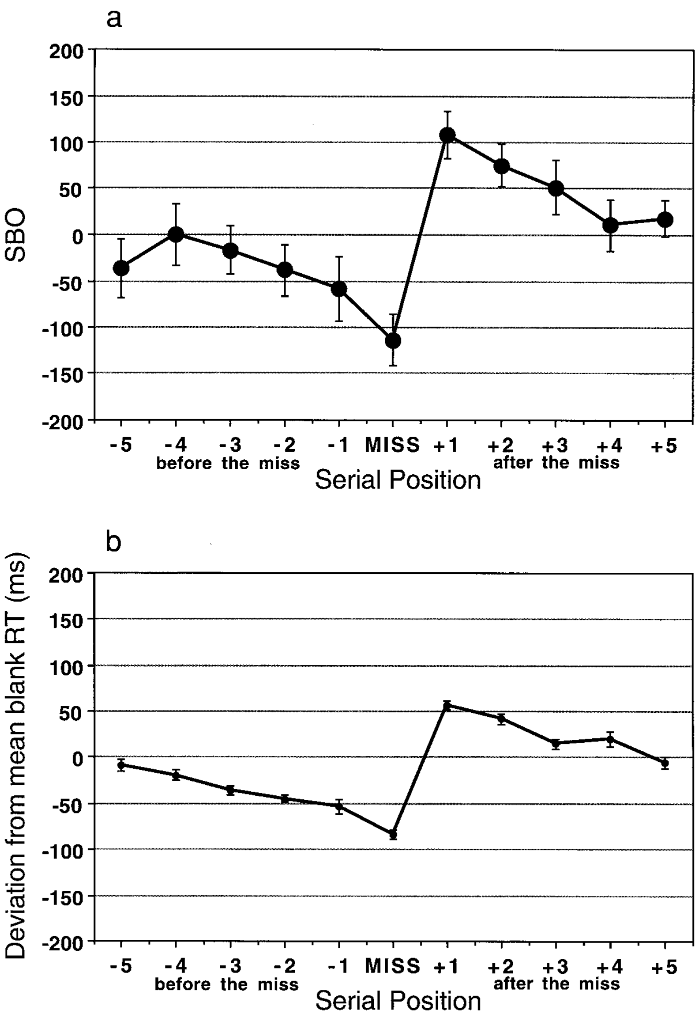

FIG. 9. Blank trial RTs as a function of their serial position relative to a miss. RTs increase sharply following a miss and decline slowly at other times. (a) shows real data. (b) shows that the simulation produces very similar if much less noisy results.

for either blank or target trials as appropriate. The normalized averages from Experiment 1 are shown in Fig. 10a.

Some observers produced only a few miss errors resulting in empty cells for the conditionalized data (miss_NEG or miss_HIT). The data from these observers (three in the Easy condition and one in the Hard condition) were excluded from the analysis shown in Fig. 10a and from the corresponding statistical tests. For both conditions, the average RTs for successful blank

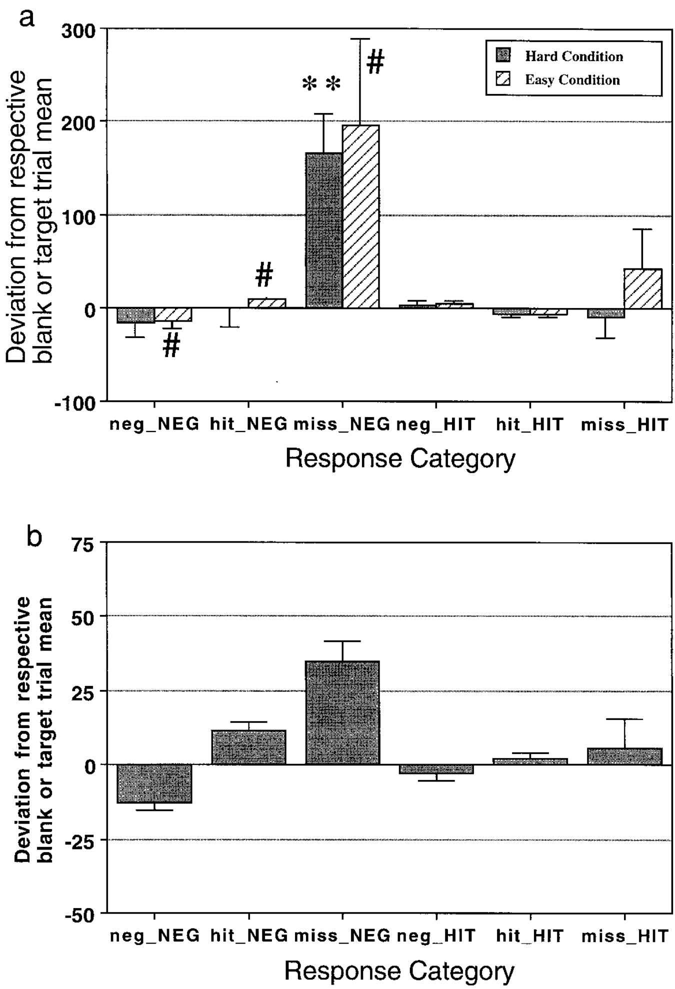

FIG. 10. Change in RT relative to baseline for blank and target trials conditionalized on the nature of the preceding trial. (a) shows that for real data the RTs increase markedly for blank trials that follow a miss. The two sets of bars represent two tasks as described in the text. The * and # are shown to indicate statistical significance from baseline (two-tailed t-test, ** = $p < .01$ level, * = $p < .05$, # = $p < .10$) The model (b) produces a similar pattern of results. Error bars represent standard error of the mean.

trials appearing immediately after a miss were longer than the zero baseline (Easy Condition: $t(12) = 2.11$, $p = .06$; Hard Condition: $t(14) = 3.81$, $p < .01$). This indicates that the activation threshold jumps towards a more conservative value to reduce the probability of making another miss, resulting in a longer reaction time for the blank trial. As predicted, however, the increase in RT after a miss is seen *only* for the blank trials: the average of target RTs directly following a miss trial were not significantly different from the target trial RT baseline (for both Conditions, $p > .34$). This result supports the idea that a miss error causes a dynamic shift in an activation threshold used for terminating a search trial, and that this threshold does not affect the time it takes to locate a target. Thus, we can reject the common notion that observers are globally slower after errors. For purposes of comparison, Fig. 10b shows the same analysis for the simulated data shown in Fig. 9b. Simulated and human observers produce comparable results.

Error Rates and Location

The activation threshold hypothesis also predicts that missed targets will generally be items with low activation values. According to the model, attention is serially deployed to items in order of decreasing activation. Target events with low activations will thus result in slower target RTs and have a higher likelihood of being missed. Although we do not have direct access to the internal activation values of any particular target item, due to decline in spatial vision with retinal eccentricity and to other attentional factors (Useful Field of View: Ball et al., 1988, 1990; Sekuler & Ball, 1986), one can hypothesize that the average activation of target items decreases with increasing eccentricity. If so, slower RTs and higher miss error rates should be found for targets appearing in peripheral locations. In our search task, targets could appear in any one of 36 locations within a square 6×6 grid. We computed the mean RTs for target present trials as a function of target location. The results of this analysis for both conditions are shown in Fig. 11. It is clear from this figure that RTs are fastest for items appearing near foveal fixation and are slower for targets that appeared in peripheral locations. The probability of missing a target appearing at any given location is shown in Fig. 12. Figure 12 illustrates that targets which appeared in the periphery were more likely to be missed. This is the opposite of the speed-accuracy tradeoff. RT and error rate were significantly correlated (Hard condition: $R^2 = .52$, $t(34) = 6.1$, $p < .001$; Easy condition: $R^2 = .32$, $t(34) = 4.0$, $p < .001$). As predicted by the activation threshold model, targets with hypothetically low average activations yielded both (1) slower RTs and (2) higher probabilities of being missed. These results are comparable to those of Carrasco and colleagues (Carrasco & Evert, 1991; Carrasco, Evert, Katz, & Chang, 1995).

Speed-Accuracy Tradeoff

If the data are examined in another way, the activation threshold hypothesis does predict a speed-accuracy trade-off *within* observers. The probability of missing a target should increase the earlier a search is terminated. This can be examined through an analysis of the RT distributions for ''target absent'' responses. Using latency data from a serial, self-paced two-choice task, Rabbitt and Vyas (1970) analyzed the probability that a response will be an error

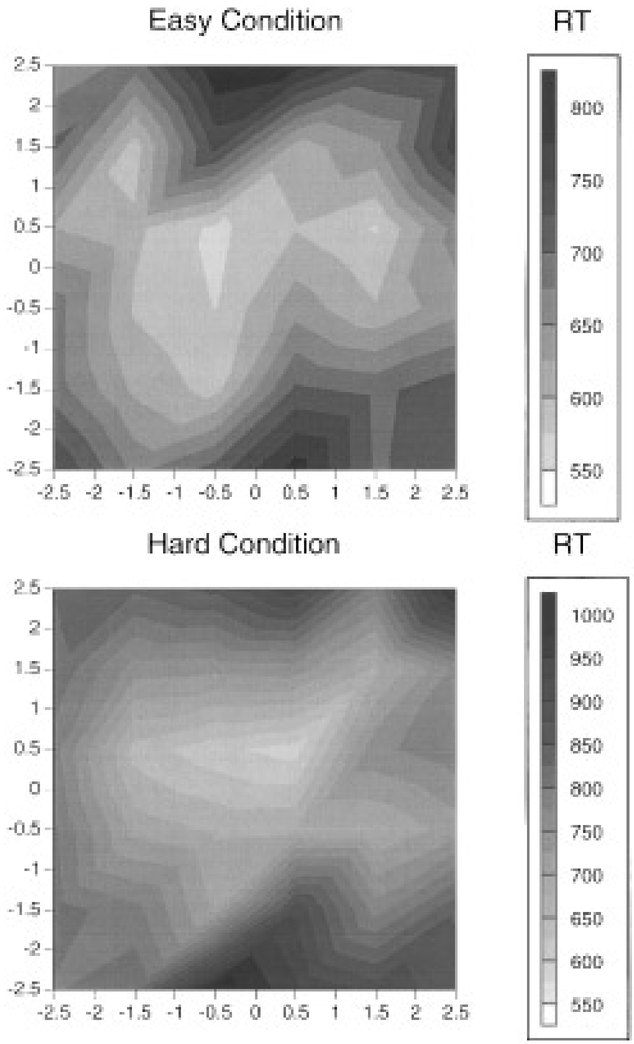

FIG. 11. Contour plot showing mean RT for correct hits as a function of target position. RTs are higher in peripheral locations than in central locations.

as a function of RT interval (which they termed as T functions). As would be expected, a higher proportion of errors were found for the faster RTs within a latency distribution. Analogous results were found for the target absent response data in our task. For each observer, pentiles for ‘‘absent’’ response trial RTs were computed, and within each pentile, the probability of making a miss error was computed. The corresponding analysis for false alarms plotted against the ‘‘yes’’ response time distribution as a function of

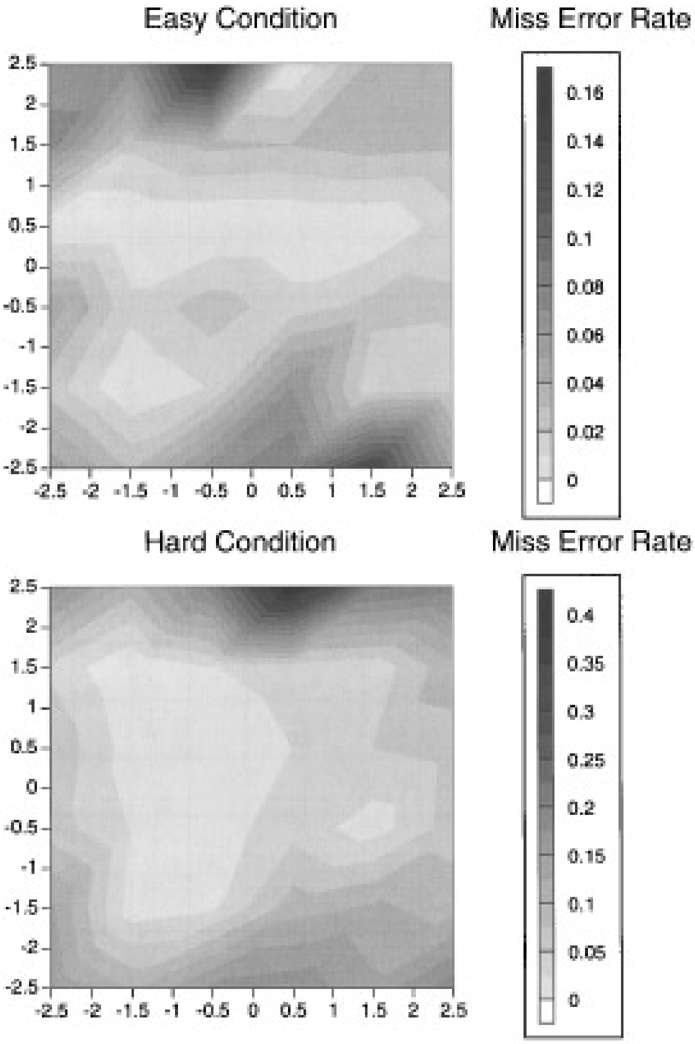

FIG. 12. Contour plot showing mean Miss error rate as a function of target position. Observers were more likely to miss a target appearing in the periphery than at central regions.

pentile was also analyzed. Pentiles, rather than deciles, were used because false alarms are quite rare in this data set. The results of these analyses for the Hard and Easy condition are shown in Figs. 13a and 13b, respectively. In both conditions, the probability of making a miss error was higher within the cohort of fast ''target-absent'' responses than within slower responses. This classic speed-accuracy tradeoff, shown *within* a distribution of an observer's responses, is consistent with predictions from the activation threshold

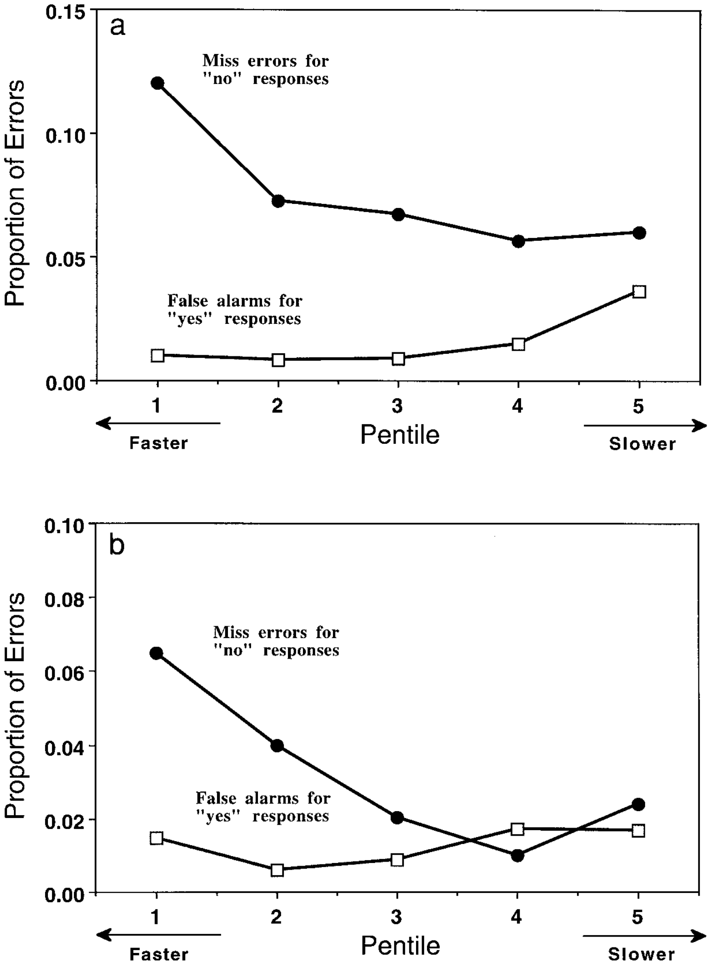

FIG. 13. Proportion of miss or false alarm errors made as a function of ''no'' and ''yes'' RT distribution pentile, respectively. Observers were more likely to make a miss for fast responses than for slower responses.

hypothesis. It is worth noting that in contrast to the miss error data, the pattern of false alarms did not show any strong speed-accuracy tradeoffs with reaction time pentile.

Serial Search

The Guided Search model assumes serial checking through the subset of display items that have activation values above the hypothesized threshold. For ''parallel'' searches, that subset may be a single item on target trials

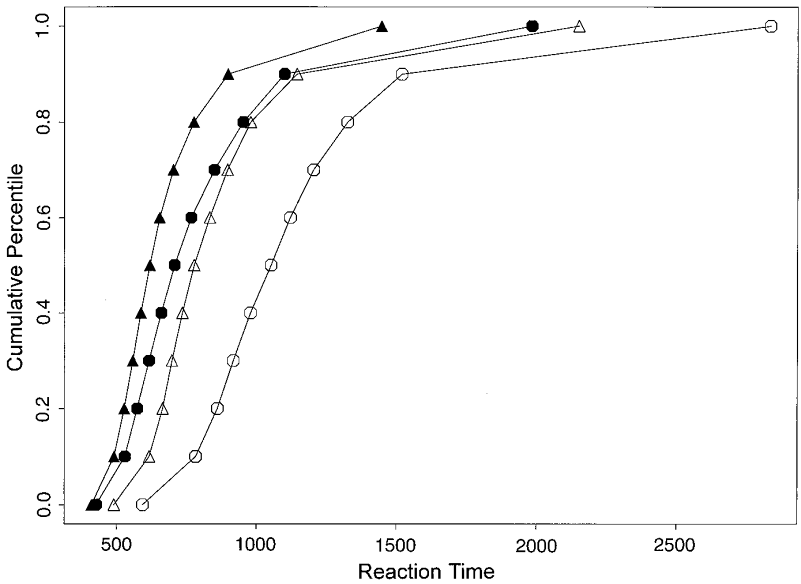

FIG. 14. Cumulative distributions for correct hit and correct blank trials in Experiment 1. Solid circles: hard condition hit distribution; solid triangle: easy condition hit distribution; hollow circle: hard condition blank distribution; hollow triangle: easy condition blank distribution.

while for truly serial searches, the subset may be the set of all items. Models of this sort (see also Treisman's Feature Integration Theory: Treisman & Gelade, 1980; Treisman & Sato, 1990) require that more items are searched on an average blank trial than on an average target trial. Closer examination of the RT distributions for hit and blank trials does suggest support for this idea. Average RTs are, of course, longer for blank trials than for target trials. They also tend to have higher variance. However, since these effects could be produced by positive skew on the blank trial RT distribution, increased mean and variance are not strong evidence for an increase in the number of items examined. If target and blank trials actually differ in the number of serial comparison processes, then the *minimum* of the RT distributions should be longer for blank as compared to target trials (Ratcliff, 1978, 1979; Ratcliff & Murdock, 1976; Townsend, 1990). To examine whether the RT differences between target and blank trials were due to increased positive skew for blank trial distributions, we computed separate target and blank trial cumulative distributions for each observer. The vincentized average compiled across observers (Ratcliff, 1979) is shown in Fig. 14. It is clear from the RT distribution analysis that the differences between target and blank trial RT means reflect a shift in the overall distribution of responses and not just increased positive skew, supporting the Guided Search assumption that more items are being searched through on blank trials than on target trials. We do not present this analysis as evidence for a distinction between serial or a limited-capacity parallel mechanisms for search (Townsend, 1990), nor does the Guided Search model make strong commitments to the serial checking assumption (Wolfe, 1994; Wolfe & Cave, 1989; Wolfe et al., 1989). However, the RT distribution analysis is at least consistent with the serial checking assumption currently implemented by the model and simulation.

We have presented a wide range of analyses in support of an activation threshold account for termination of blank trials. In sum, the most important aspect of these data for the present model of search termination is the finding that RT changes after a miss have the largest impact on blank trials. Target trial RTs, theoretically independent of the setting of the activation threshold, are empirically less influenced by the effects of a miss. We have also shown support for the notion that observers dynamically adjust the speed of their target-absent responses from trial-to-trial, terminating as quickly as possible while minimizing the number of miss errors made. We do not presume that observers have direct access to their activation thresholds. The trend analysis shown in Fig. 9a does suggest that a heuristic for controlling performance would be to follow a staircase procedure as implemented in our simulation (described in the Introduction). A correct response suggests that a faster response may be made. An error indicates that the current activation threshold is too high. Such feedback allows the observer to estimate the limits of his/her performance for terminating blank trials with a certain degree of accuracy.

In these respects, the dynamic adjustment of an activation threshold proposed here is similar to the tracking model elaborated by Rabbitt and his collaborators (1969; Rabbitt & Vyas, 1970) to describe performance in serial self-paced choice-response tasks. This model assumes that observers have an internal T function which limits their efficiency in any given task and additionally assumes that observers must discover the nature of this T function to shape their RT distribution to conform to its characteristics. This can be done by progressively making faster responses until errors are made. An error would indicate that the speed of responding should be modified to go more slowly. In other words, observers buy information about their internal T function by making errors.

The activation threshold model hypothesizes that the ‘‘cost’’ associated with making an error is what drives adjustment of the threshold. In the next experiment, we attempt to more directly control the activation threshold by manipulating observers’ tolerance for errors.

EXPERIMENT 2

Manipulating Observers’ Tolerance for Errors

In the model, the staircase step-size and the resulting activation threshold are determined by the observer’s tolerance for errors. If an observer is unwilling to commit errors, the downward step in activation threshold following a miss will be large. This will result in a lower average threshold and more distractors will be examined on blank trials. Unfortunately, we do not have direct access to internal activation thresholds in humans even if those thresholds exist. However, the model does make concrete predictions about the effects of manipulating observers' tolerance for errors. First, the increase in blank trial RT after a miss should be greater when the error tolerance is lower. Second, regardless of the tolerance for errors, there should be little or no impact of a miss on target trial RTs. Third, overall blank trial RTs should increase as error tolerance decreases but overall target trial RTs should remain largely unchanged. To test these hypotheses, we repeated the Hard Condition (two-shared feature triple conjunction search) of Experiment 1 with an added manipulation of error tolerance.

Method

Procedure and design. The apparatus and methods used in this experiment are the same as in Experiment One except for the error tolerance manipulation. As is routine in signal detection tasks, observers' criteria were manipulated by changing the payoff matrix. In this experiment, the ''payoff'' was a score in the manner of a video game. After each trial, a score was flashed in the middle of the screen. This score was a positive number for correct responses and a negative number for errors. The number of points awarded for each correct response was computed by subtracting the reaction time from a constant (i.e., faster RTs yield more points). Constants for target and blank trials were individually adjusted for each observer based on a practice block of trials. For example, if an observer's constant was 700 ms for hit trials and she made a correct response in 550 ms, then she would have been awarded 150 points for that target trial. If the reaction time was over the criterion value, then zero points were awarded.

The scoring and feedback on error trials was the crucial manipulation for this experiment. Two payoff matrices were used. In the Liberal Condition, observers were informed that they would suffer a 3000 point *penalty* for false alarm errors (wiping out 10 to 20 trials worth of correct trial points), and *no* penalty for miss errors. They were specifically instructed that false alarms involved responding ''yes'' to trials where the target was not present, and that misses involved responding ''no'' to trials when the target was present. The penalty for false alarms was included in the liberal condition to prevent observers from simply making random keypress responses. In the Conservative Condition block, a 5000 point penalty for misses was added, while the penalty for false alarms was kept constant at 3000 points.

Observers were run in three blocks of 30 warm-up trials and 300 data trials each. The first block was a practice block to familiarize observers with the task and scoring procedures. The average reaction times for target and blank trials in the practice block were used to adjust the constant for each individual

TABLE 2
Results of Experiment 2

Condition	Reaction time (ms) Target trial mean (*SD*)	Reaction time (ms) Blank trial mean (*SD*)	Error rate (%) Target trial (MISS)	Error rate (%) Blank trial (False alarm)
Liberal	667 (190)	814 (196)	15.3	1.9
Conservative	669 (179)	876 (186)	6.2	0.9

observer. The Liberal and Conservative conditions were run in the two latter blocks, with the order of these conditions counterbalanced across observers.

Even though correct responses with RTs slower than the scoring cutoff criterion received a score of zero, they were included in the data analysis.

Participants. Sixteen observers from the previously described observer pool were tested.

Results and Discussion

The RT and error data for both conservative and liberal payoff conditions are presented in Table 2. The basic manipulation worked. Miss error rates were 2.5 times greater in the liberal condition than in the conservative condition (15.3% vs 6.2%; $F(1,15) = 10.24$, $p < .01$). As predicted by the model, there was no effect of payoff condition on the average RTs for correct target trials ($p > .84$) but there was a significant effect on correct blank trials, $F(1,15) = 5.21$, $p < .05$). A lower tolerance for errors does not make observers globally more conservative in their responses. It makes them more specifically reluctant to terminate unsuccessful searches. Although there was a significant difference in the proportion of false alarm errors made in the two conditions, $F(1,15) = 7.38$, $p < .05$, the significant interaction between payoff condition and error type, $F(1,15) = 9.01$, $p < .01$, suggests that blank trials were more strongly affected by the payoff manipulation.

The results of the conditional RT analysis are shown in Fig. 15. The data from two observers were excluded from the current analysis because they made so few misses that critical cells in their data were empty (e.g., no miss errors followed by a hit). As shown in Fig. 15, the average deviation of blank trial RTs directly following a miss error were significantly longer than baseline (Liberal condition: $t(13) = 2.35$, $p < .05$; Conservative condition, $t(13) = 4.43$, $p < .001$). The change in blank trial RT following a miss was greater in the conservative condition than in the liberal condition, $t(13) = 2.14$, $p = .052$. Following the predictions of the model, this result suggests that observers made larger decreases in activation thresholds following misses in the conservative condition.

As in Experiment 1, the mean RTs for targets that appeared in peripheral

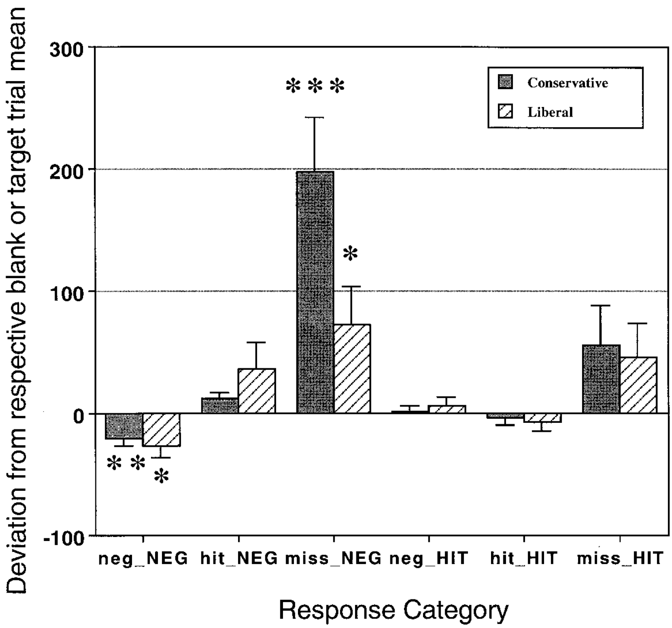

FIG. 15. Same as Fig. 8a but showing the effect of manipulating observers' tolerance for errors. When errors are heavily penalized, blank RT following an error increases more steeply than when errors are lightly penalized.

locations were longer than for targets appearing near foveal fixation. The probability of making a miss error was higher for these peripheral target locations. An error analysis as a function of response time distribution pentile indicates that within observers and within each error-tolerance condition, the probability of missing a target was higher for fast ''no'' decisions compared to slower ''no'' responses. The probability of making false alarm errors did not show any such dependency on the relative speed of corresponding ''yes'' responses.

In general, Experiments 1 and 2 provide gratifying support for the activation threshold model of blank trial search termination. Observers are significantly slower on blank trials but not on target trials after a miss. Error tolerance alters average blank trial RTs but not average target trial RTs. Miss errors occur mainly for targets with low activation (as inferred from their mean RT), and the probability of missing a target is higher for fast responses than for slower responses.

Up to now, we have ignored predictions from the interrupt ''guessing'' mechanism. In the introduction we asserted that a guessing mechanism alone, even a sophisticated one based on timing estimates, would be insufficient for predicting performance on blank trials. The next experiment tests predictions of such a timing strategy.

EXPERIMENT 3

Testing a Timing Hypothesis

We propose that blank trial termination involves a combination of activation thresholds and guessing based on the duration of the trial. In Experiment 3, we test the possibility that a timing strategy *alone* could explain how searches are terminated. Since the speed-accuracy tradeoffs in the previous experiments could be produced by a timing threshold as well as by an activation threshold, we designed a different test to directly compare the two termination rules. The results of Experiment 3 suggest that the timing mechanism cannot be the sole determinant of search time on blank trials.

Consider a color × orientation conjunction search with one target type (red vertical) and three distractor types (red horizontal, green vertical, and green horizontal). In Guided Search, these different distractors will have different characteristic activations. At the very least, green horizontal distractors that share no features with the target should have lower activations than the other two distractor types. Varying the proportions of different distractors is known to alter search times (Egeth et al., 1984; Pashler, 1987; Poisson & Wilkinson, 1992; Zohary & Hochstein, 1989). In this experiment we vary the proportions *within* a block of trials. Timing and activation thresholds make different predictions about responses to different types of trials in a mixed block of trials. Unless observers could adjust a timing threshold based on the perceived proportion of one type of distractor in a display, a single timing threshold should govern all blank-trial terminations in a mixed block of trials. Blank trial RTs should be similar for different types of trials and error rates should be higher for the harder task. If observers use an activation threshold as a termination rule, however, the harder tasks will have more items with high activation and will take longer. Error rates will be relatively stable for different tasks within the block.

Method

Two versions of the experiment were run. In the Easy version, observers looked for a red vertical target among three distractor types: green vertical, red horizontal, and green horizontal. In the Hard version, the target was a red, left oblique among red, right oblique, green left oblique, and green right oblique. Set size was kept constant at 25 items for all trials in the experiment. Each version of the experiment consisted of 30 practice and 300 experimental trials. These trials were comprised of a random mixture of three types of displays. In *color dominant* displays, two thirds of the 24 distractor items shared the same color with the target and one third of the items shared the same orientation with the target. In *orientation dominant* displays, two thirds of the items shared the same orientation with the target and one third of the items shared the same color with the target. In *mixed* displays, one third of the items shared the same color, one third of the items shared the same

TABLE 3
Results of Experiment 3

	Target trials mean	Blank trials mean
Easy task		
Red vertical target		
Color dominant	675	866
Orientation dominant	706	754
Mixed	646	799
Hard task		
Red left oblique target		
Color dominant	786	1030
Orientation dominant	826	855
Mixed	753	857

orientation, and the remaining third of the distractors shared no features with the target (green horizontal in the easy condition; green, right oblique in the hard condition). The line stimuli used were .85 degrees in visual angle (CIE x, y coordinates: red −.62, .36; green −.34, .57).

Ten observers from the previously described paid volunteer pool were run on both blocks in a single one-hour session. The order of the two versions of the experiment were counterbalanced across observers. All procedures and apparatus used were identical to Experiment One. Feedback was given for correct and incorrect responses.

Results and Discussion

Table 3 shows the average RTs for target and blank trials for each of the three display types in each task. For purposes of this paper, it is the blank trials that are of interest. There was a significant effect of distractor ratio (display type) on blank trial RT in both versions of the experiment. Specifically, RTs were longer when there were more red items (Easy Version: Color dominant-Orientation dominant condition, $t(9) = 2.2$, $p < .05$, Color dominant and Mixed condition, $t(9) = 1.9$, $p < .05$; Hard Version: Color-Orientation, $t(9) = 4.5$, $p < .001$, Color-Mixed, $t(9) = 3.6$, $p < .01$). This suggests that in this experiment, the distractors sharing color with the target had higher average activations than distractors sharing orientation. Thus, they were more likely to lie above the activation threshold and more likely to be checked on blank trials. A timing hypothesis would predict this result *only* if separate timing thresholds were maintained for each distractor ratio. It seems implausible that observers could keep track of the requisite information.

The results of this experiment make it unlikely that blank trials are terminated primarily by a timing strategy. Our earlier proposals to this effect were incorrect (Wolfe & Cave, 1989; Wolfe et al., 1989). The activation threshold hypothesis provides more concise, intuitive predictions for the data. The presence of some false alarm errors makes it reasonable to propose that guessing does occur on a limited number of trials.

GENERAL DISCUSSION

This paper has presented a model for termination of unsuccessful visual searches in the context of the larger Guided Search 2.0 model (Wolfe, 1994). Preattentive processes evaluate the probability that each item in the display could be the target. Items are examined in decreasing order of that probability. Unsuccessful searches are terminated when no remaining items have probabilities above a termination threshold. The threshold is set adaptively and automatically by the observer in order to place performance at some desired combination of speed and accuracy. If the error rate is too high, the threshold is lowered so that searches are terminated later, slowing the RT. If the error rate is low, the threshold is raised so that fewer items are examined, speeding the RT but increasing the error rate. Given many trials of the same search, observers come to (implicitly) know something about the time required to complete a successful search. They terminate some trials with a guess, influenced by the duration of the trial. That is, they are more likely to guess as time passes within a trial. They guess ‘‘no’’ most of the time but ‘‘yes’’ enough of the time to produce the small number of false alarms seen in these tasks.

Experiments 1 and 2 provide support for specific predictions of the activation threshold portion of the model. (1) RTs for target-absent response decline a little after correct target-absent responses and rise markedly after each miss while RTs for target-present responses were unaffected. (2) Miss errors were generally faster than correct target-absent responses. (3) Speed-accuracy tradeoffs were found within each individual’s distribution of blank trial RTs. (4) Both RTs and error rates increase with eccentricity as predicted if internal activation values of items decrease as a function of eccentricity (see Carrasco & Evert, 1991; Carrasco et al., 1995). (5) Manipulation of observers’ tolerance for miss errors in Experiment Two had an impact on blank trial RTs but little effect on target trial RTs. Finally, Experiment 3 falsifies a pure timed guessing version of the model (as was proposed in an earlier version of Guided Search, Wolfe and Cave, 1989).

The model has relatively few free parameters. By just varying threshold stepsize, k, and guessing rate, g, we were able to produce reaction time and error rate patterns which are consistent with the empirical data. These simulated parameters may vary unpredictably between real human observers, yet our simulation results were robust even when averaged across a variation of these parameters.

If the model is run n times with the same parameters, there is less variability in the model data than in data collected from n real observers. This is not surprising since multiple runs of the simulation with fixed parameters are the equivalent of multiple runs of a single, very stable observer—not single runs of *n* variable observers. The model has two parameters that could be adjusted to account for variability between observers. First, the staircase rule (e.g., 20-down, 1-up). This corresponds to the signal detection theory parameter, b, reflecting the observer's tolerance for errors. This parameter is under some voluntary control. The variability in the data cannot be accounted for purely in terms of variation in b. Criterion shifts alone would predict strong speed-accuracy tradeoffs *across* observers. As the activation threshold rose, error rates would rise and slopes would fall. We do not see a correlation between error rates and slopes *across* observers suggesting that other factors contribute to variation in the data. The second available parameter is the signal strength. In the simulation, the signal strength is simply a number in some arbitrary units. In real search tasks, a red target among green distractors generates a signal of some strength and that signal may vary from observer to observer. Presumably, this is under less voluntary control though a number of investigators have suggested that perceptual learning effects may reflect an ability to improve the signal to noise ratio. Similar effects could underlie improvements in visual search with practice.

The model has a few underlying assumptions that are worth examining. First, it is assumed that the underlying distribution of internal noise is normal. This is a matter of convenience though it seems plausible enough in a biological system. If future physiological research found a convincing correlate of item activations and if those distractor activations turned out to have some non-normal distribution, the model would have to be revised.

The second assumption is that the standard deviation of the activation distribution decreases as the signal strength increases. This, it must be admitted, is a somewhat arbitrary solution to a specific problem. Any model must deal with the fact that empirical blank to target trial slope ratios are roughly 2:1 over the full range of search tasks. Our solution is to have the signal distribution change with signal strength (Zohary et al., 1994). There may be other solutions to this problem (e.g., non-normal distributions) but any model must be able to handle a wide variety of search tasks without producing systematic changes in slope ratios.

A third assumption is that, during the course of a single trial, observers must somehow internally ''mark'' items that have been rejected. In this model, observers are proceeding down an activation list. It is important that they not revisit already rejected loci. There is evidence for what has been called ''inhibition of return'' in attentional tasks (Klein, 1988; Mackeben & Nakayama, 1988; Posner & Cohen, 1984; Tipper, Driver, & Weaver, 1991). Unfortunately, Klein's (1988) evidence for inhibitory marking of locations in visual search has proven hard to replicate (Wolfe & Pokorny, 1990). Nevertheless, given the results from other attentional paradigms, it seems very likely that some sort of inhibitory tagging occurs during visual search. In our simulation, this inhibitory tagging is perfect. It seems likely that this is an oversimplification and that real observers may recheck some items during a trial. Although not currently implemented, it would be trivial to simulate rechecking of items within the activation threshold model. Tagging would be achieved by suppressing the activations of checked items. Noisy inhibitory tagging could be simulated by allowing these suppressed activations to regain strength as the duration of search lengthens within a trial. Thus there would be an increased probability of rechecking for items that were checked early, or for items which had initially high activations, or on trials for which the current activation threshold is set at a low value.

Search Termination and Other Visual Search Models

How does this model of search termination fit with visual search models other than Guided Search? The diverse array of visual search models available in the literature concentrate on elucidating the mechanisms involved in finding a target. Existing models can be broadly categorized into two major camps: one which assumes that conjunction of features occurs within localized regions of space in an item-by-item serial manner (Treisman & Gelade, 1980; Treisman & Sato, 1990; Wolfe, 1994; Wolfe et al., 1989); the other proposing that conjunctions of features for every item are processed in parallel across the visual field (Duncan & Humphreys, 1989; Humphreys & Müeller, 1993; Ward & McClelland, 1989). Hybrid models have also been proposed which postulate that conjunction search is serial over clumps of items but parallel within these clumps (Pashler, 1987). A more recent model developed along these lines proposes that visual representations are organized into multiple boundary and surface representations, and that search progresses recursively through the candidate groups formed by these processes (Grossberg et al., 1994). The Guided Search 2.0 model currently implements serial checking of each item, but as noted in the introduction, it does not make strict commitments to the serial processing assumption. The issue of whether conjunction search is serial or parallel is still extensively debated in the literature and is beyond the scope of this paper.

We wish to argue that the threshold mechanism described in this paper should be a part of any search model regardless of whether the underlying visual processing is serial or parallel in nature. In particular, even parallel models of visual search incorporate ‘‘thresholds’’ to modulate search accuracy. Indeed, a number of the alternative search models have properties reminiscent of those that we propose here.

Humphreys and Müeller (1993) developed a computationally explicit connectionist model for visual search which assumes that conjunctions of line segments (T’s and L’s) are computed in parallel across the visual array. Based on principles of similarity and spatial proximity, parallel processing of form allows for grouping of items within a display. As the name of their model, SEarch via Recursive Rejection (SERR), implies, search operates via recursive rejection of areas of the field where grouping of similar items has been achieved. When distractor–distractor similarity is high, all distractors may be grouped and rejected in a single step, allowing for search to be completed independently of display size. For heterogeneous displays, grouping will only succeed over a smaller subset of the items. A recursive process of grouping and rejection of target-absent sub-groups continues until all of the items in the display have been examined. Therefore, search time increases with distractor heterogeneity. There is also an increased likelihood that a target will be missed in heterogeneous displays, and this likelihood increases as a function of set size. To account for observers' tendencies to minimize the number of errors made, they proposed that observers rely on a *response criterion* to maintain an acceptable error rate across various conditions within a task. Optimizing speed/accuracy tradeoffs results in increased search time as a function of the number of distractors. In SERR, this response criterion was implemented in the form of rechecking (rerunning the simulation on a number of trials to produce a more acceptable error rate). The response criterion implemented in SERR to successfully simulate human search performance corresponds to our ''activation threshold.'' In the context of Guided Search, we propose that an activation threshold is used to achieve an optimal balance between speed and accuracy in search and that this activation threshold is under the dynamic control of the search process. Likewise, the response criterion of SERR is used to balance miss rates as a function of the efficiency of recursive rejection of distractor groups. In other words, the thresholds used in both models can be conceptualized to be independent of the specific mechanics of earlier visual processes (i.e., whether visual features are computed in a serial or a parallel manner).

A resource-limited parallel model, based on the diffusion model of Ratcliff (1978), was proposed by Ward and McClelland (1989) as an alternative to models assuming serial, self-terminating search. In this model, processing is viewed as a random-walk, evidence-accumulating process that proceeds simultaneously over all display locations. Detectors accumulate evidence for target presence in parallel across the visual field. Each detector begins at a resting activation level which drifts towards a positive or negative threshold value as evidence is accumulated. A positive decision is made when a detector crosses a positive threshold and search terminates when all detectors have crossed their negative threshold. In addition, their model assumes that the resting activation level and thresholds are variable in response to task variables. In particular, dynamic adjustment of these resting activation levels and thresholds was implemented in their simulation to approximate actual error rates shown in the empirical data of human observers. As in SERR, a threshold component, which is functionally equivalent to the activation threshold we propose in the context of Guided Search, was critical to allow their model to realistically simulate human performance.

Pashler (1987) proposed a cluster model arguing against an item-by-item serial search process. RT slope data obtained from search tasks using set sizes smaller than 8 suggest that search within this lower range of display size may be parallel and self-terminating. Rather than search proceeding in an item-by-item basis, he proposed that search is serial over clumps of items, with search processes being parallel within each clump. Although a computational model was not developed to differentiate predictions of processing clusters versus processing of single items, it is clear that at display set sizes which exceed the processing capacities of a clump processor, a process is involved where search is serial over several clumps. In other words, within his model, observers may utilize thresholds that determine the number of clumps to process, or the thoroughness with which processing is done within each individual cluster. This is also applicable to the search model by Grossberg et al. (1994) which proposes that search occurs through recursive selection of candidate groupings of items until a target object is identified.

Therefore, although the actual implementation of thresholds used to achieve realistic accuracy in various computational simulations varies from model to model, it seems that some sort of adjustable threshold component is required in any model of visual search because human observers adjust their performance to reduce the proportion of miss errors, resulting in speed/accuracy tradeoffs shown in the empirical data (see also Krueger & Shapiro, 1980).

As we stressed throughout the paper, one need not be an adherent of the Guided Search model in order to see the need for a new account for blank trials. Consider the variance of target and blank trials in a simple serial self-terminating search. Targets are found after a random search through the items; sometimes on the first attentional fixation, sometimes on the last. On average, half of the items must be examined. On blank trials, search is always exhaustive. Ignoring other factors, this argues that the variability of blank trial RTs should be lower than that for target trial RTs.[1] As Ward and McClelland (1989) pointed out and as we and others have confirmed in a number of experiments, this is rarely the case. Variance of blank trial RTs is almost always higher than that for targets. This raises a problem for a simple serial search model. There are a variety of solutions to this problem (Ward and McClelland suggest one). However, it is a useful feature of the present model that it produces variances that have the pattern seen in the data.

[1] If you do not ignore other factors, it has been brought to our attention that it is not always true that self-terminating search for targets should produce greater RT variance than an exhaustive search on target-absent trials. Townsend and Ashby (1983, Propositions 7.7 and 7.8, pp. 194–196) have shown that exhaustive processing time variance can be larger than self-terminating processing time variance when the variance of individual element processing times, var(T), is greater than $((n+1)/6)[E(\mathrm{T})]^2$, where n is the number of elements, and E(T) is the mean comparison time for an element. On statistical grounds, for a set size of 25 as used in the present experiments, the variance of exhaustive processing will be larger than the variance of self-terminating processing if the *SD* of individual element processing times exceeds 2.082 times the mean comparison time of each element, E(T). We thank Gordon Logan and Lester Krueger for pointing this out.

Another consistent pattern found in a variety of visual search tasks is an increase in errors with larger set sizes (Egeth et al., 1983; Pashler, 1987; Treisman & Gormican, 1988; Wolfe et al., 1989; Ward & McClelland, 1989). In our simulation, set size effects on error rates are generated when observers guess on a small proportion of trials to produce faster reaction times. While a guess has only a 50% probability of being correct at the beginning of a trial, as evidence accumulates, more ‘‘educated guesses’’ can be made increasing the probability of a guess being correct. Observers guess with increasing probability as a search progresses. This lowers RTs somewhat and increases error rates at the higher set sizes.

In summary, the Guided Search model and, indeed, any model of visual search needs to be able to explain how observers terminate unsuccessful searches. We suggest that observers mainly use a staircase-like procedure to set an activation threshold. They then search through those items whose activations are high enough to make them plausible candidate targets. When that list has been exhausted, search is terminated. If search is terminated incorrectly, observers become more conservative on subsequent trials. When the speed of a search becomes imperative to a task, a number of guessing responses may also be found, producing both misses and false alarms.

In real world searches, we may learn and remember a variety of activation thresholds for a variety of tasks. In a search to determine if the cat has been put out or a search for a late night snack in the refrigerator, unsuccessful searches are not exhaustive searches. They are terminated when all likely locations have been examined. ‘‘Blank trial’’ searches become more exhaustive when the costs of failure increase. Consider the search for a missing but important piece of paper among the piles on your desk or the task of a photoanalyst, searching satellite images for evidence of hostile missile deployment. In such cases, blank trials are likely to be terminated only with the greatest reluctance.

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,** 2210–2219.

Ball, K. K., Roenker, D. L., & Bruni, J. R. (1990). Developmental changes in attention and visual search throughout adulthood. In J. T. Enns (Ed.), *The development of attention: Research and theory* (pp. 489–508). Amsterdam: Elsevier North-Holland.

Carrasco, M., & Evert, D. L. (1991, November). *The effect of target position in a conjunction visual search task.* Paper presented at the Annual Meeting of the Psychonomics Society, San Francisco, CA.

Carrasco, M., Evert, D. L., Katz, S. M., & Chang, I. (1995). *Eye movements, covert attention, and spatial resolution in the processing of conjunction visual searches.* Manuscript submitted for publication.

Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology,* **22,** 225–271.

Cohen, A., & Ivry, R. B. (1991). Density effects in conjunction search: Evidence for coarse

location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance,* **17,** 891–901.
Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review,* **96,** 433–458.
Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance,* **10,** 32–39.
Grossberg, S., Mingolla, E., & Ross, W. D. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review,* **101,** 470–489.
Hoffman, J. E. (1978). Search through a sequentially presented visual display. *Perception & Psychophysics,* **23,** 1–11.
Hoffman, J. E. (1979). A two-stage model of visual search. *Perception & Psychophysics,* **25,** 319–327.
Humphreys, G. W., & Müeller, H. J. (1993). SEarch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology,* **25,** 43–110.
Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature,* **334,** 430–431.
Krueger, L. E., & Shapiro, R. G. (1980). Why search for target absence is so slow (and careful!): The more targets there are, the more likely you are to miss one. *Journal of Experimental Psychology: Human Perception and Performance,* **6,** 662–685.
Kwak, H., Dagenbach, D., & Egeth, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception and Psychophysics,* **49,** 473–480.
Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of Acoustic Society of America,* **49,** 467–477.
Mackeben, M., & Nakayama, K. (1988). Fixation release facilitates rapid attentional shifts. *Investigative Ophthalmology and Visual Science (suppl.),* **29,** 22.
Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics,* **41,** 191–201.
Pashler, H. (1987). Target-distractor discriminability in visual search. *Perception & Psychophysics,* **41,** 285–292.
Pavel, M., Econopouly, J., & Landy, M. S. (1992). The psychophysics of rapid visual search. *Investigative Ophthalmology and Visual Science,* **33,** 1355 (abstract 3313).
Poisson, M. E., & Wilkinson, F. (1992). Distractor ratio and grouping processes in visual conjunction search. *Perception,* **21,** 21–38.
Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bowhuis (Eds.), *Attention and Performance X* (pp. 55–66). Hillside, NJ: Erlbaum.
Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception & Psychophysics,* **41,** 455–472.
Rabbitt, P. M. A. (1966). Errors and error correction in choice response tasks. *Journal of Experimental Psychology,* **71,** 264–272.
Rabbitt, P. M. A. (1967). Time to detect errors as a function of factors affecting choice response time. *Acta Psychologica,* **27,** 131–142.
Rabbitt, P. M. A. (1969). Psychological refractory delay and response-stimulus interval in serial, choice response tasks. In W. G. Koster (Ed.), *Attention and Performance II* (pp. 195–219). Amsterdam: North-Holland.
Rabbitt, P. M. A., & Vyas, S. M. (1970). An elementary preliminary taxonomy for some errors in laboratory choice RT tasks. *Acta Psychologia,* **33,** 56–76.
Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review,* **85,** 59–108.
Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin,* **86,** 446–461.
Ratcliff, R., & Murdock, B. B., Jr. (1976). Retrieval processes in recognition memory. *Psychological Review,* **83,** 190–214.

Sekuler, R., & Ball, K. K. (1986). Visual localization: Age and practice. *Journal of the Optical Society of America A,* **3,** 864–868.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object centered inhibition of return of visual attention. *The Quarterly Journal of Experimental Psychology,* **43A,** 289–298.

Townsend, J. T. (1990). Serial and parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished. *Psychological Science,* **1,** 46–54.

Townsend, J. T., & Ashby, F. G. (1983). *The stochastic modeling of elementary psychological processes.* Cambridge: Cambridge Univ. Press.

Treisman, A. (1988). Features and objects: The fourteenth Bartlett memorial lecture. *The Quarterly Journal of Experimental Psychology,* **40A,** 201–237.

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review,* **95,** 15–48.

Treisman, A., & Sato, S. (1990). Conjunction search revisited. *Journal of Experimental Psychology: Human Perception and Performance,* **16,** 459–478.

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

Ward, R., & McClelland, J. L. (1989). Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception and Performance,* **15,** 664–672.

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

Wolfe, J. M., & Cave, K. R. (1989). Deploying visual attention: The guided search model. In T. Troscianko & A. Blake (Eds.), *AI and the eye.* Chichester, UK: Wiley and Sons.

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

Wolfe, J. M., Chun, M. M., & Friedman-Hill, S. R. (1995). Making use of texton gradients: Visual search and perceptual grouping exploit the same parallel processes in different ways. In T. Papathomas (Ed.), *Early vision and beyond* (pp. 189–198). Cambridge, MA: MIT Press.

Wolfe, J. M., Friedman-Hill, S. R., Stewart, M. I., & O'Connell, K. M. (1992). The role of categorization in visual search for orientation. *Journal of Experimental Psychology: Human Perception and Performance,* **18,** 34–49.

Wolfe, J. M., & Pokorny, C. W. (1990). Inhibitory Tagging in Visual Search: A failure to replicate. *Perception and Psychophysics,* **48,** 357–362.

Wolfe, J. M., Yu, K. P., Stewart, M. I., Shorter, A. D., Friedman-Hill, S. R., & Cave, K. R. (1990). Limitations on the parallel guidance of visual search: Color $\times$ color and orientation $\times$ orientation conjunctions. *Journal of Experimental Psychology: Human Perception and Performance,* **16,** 879–892.

Zohary, E., & Hochstein, S. (1989). How serial is serial processing in vision? *Perception,* **18,** 191–200.

Zohary, E., Shadlen, M. N., & Newsome, W. T. (1994). Correlated neuronal discharge rate and its implication for psychophysical performance. *Nature,* **370,** 140–143.

(Received August 23, 1994)
(Revision received February 18, 1995)