

Guided Search 2.0 A revised model of visual search

JEREMY M. WOLFE

*Harvard Medical School, Boston, Massachusetts
and Brigham & Women's Hospital, Boston, Massachusetts*

An important component of routine visual behavior is the ability to find one item in a visual world filled with other, distracting items. This ability to perform *visual search* has been the subject of a large body of research in the past 15 years. This paper reviews the visual search literature and presents a model of human search behavior. Built upon the work of Neisser, Treisman, Julesz, and others, the model distinguishes between a preattentive, massively parallel stage that processes information about basic visual features (color, motion, various depth cues, etc.) across large portions of the visual field and a subsequent limited-capacity stage that performs other, more complex operations (e.g., face recognition, reading, object identification) over a limited portion of the visual field. The spatial deployment of the limited-capacity process is under attentional control. The heart of the guided search model is the idea that attentional deployment of limited resources is *guided* by the output of the earlier parallel processes. Guided Search 2.0 (GS2) is a revision of the model in which virtually all aspects of the model have been made more explicit and/or revised in light of new data. The paper is organized into four parts: Part 1 presents the model and the details of its computer simulation. Part 2 reviews the visual search literature on preattentive processing of basic features and shows how the GS2 simulation reproduces those results. Part 3 reviews the literature on the attentional deployment of limited-capacity processes in conjunction and serial searches and shows how the simulation handles those conditions. Finally, Part 4 deals with shortcomings of the model and unresolved issues.

Imagine the following scenario: you are in the midst of reading this article in the latest edition of the *Psychonomic Bulletin & Review* when you are summoned to a meeting. You place the journal on top of that two-foot stack of things to be read on your desk and depart. Unfortunately, while you are out, a shelf collapses spreading the journals, books, and an assortment of other detritus across the floor. When you return and survey the scene, your first thought is to return to this article. Obviously, you would need to undertake a visual search for the journal. If we assume that the journal is not hidden under other objects, you could efficiently locate all the purple rectangles of a certain size among other items. Those items might include *distractor* items having the *target* shape, color, or size. The target journal might be partly occluded. Still, this would not strike you as a demanding task. Although you would need to stop to read

the date on this issue of the journal in order to distinguish it from other, visually similar issues of the journal, you would not need to read the date on copies of the *Journal of Experimental Psychology: Human Perception & Performance* that share shape but not color with the target. Nor would you need to attend to a coffee mug even if its color was a perfect match to the journal's color.

The relative ease with which humans can search for targets in a crowded visual world is the subject of this paper. The ability is not trivial. The visual system cannot fully process all of its input. There is not enough room in the skull for all of the neural hardware that would be required to perform all visual functions at all locations in the visual field at the same time (see, e.g., Tsotsos, 1990). The visual system has two basic approaches to this problem. The first is to discard input. Thus, the retinal image is processed in its full detail only at the fovea. In the periphery, information is much more coarsely sampled. Receptors are spaced more widely, ganglion cell receptive fields are larger, and the cortical representation of the periphery is smaller (White, Levi, & Aitsebaomo, 1992; Wilson, Levi, Maffei, Rovamo, & DeValois, 1990). The second approach is to process information selectively. A large set of visual functions can be performed only in a restricted part of the visual field at any one moment. The identification of psychology journals is one case. Reading is a less strained example. Given

I thank Stacia Friedman-Hill and Alexander Bilsky for work on all aspects of this project. Ideas presented here were improved by conversations with Marvin Chun, Anne Treisman, Kyle Cave, and Nancy Kanwisher. I thank Asher Cohen, Patricia O'Neill, Alexander Bilsky, Greg Gancarz, Christoph Nothdurft, Jim Pomerantz, and two anonymous reviewers for comments on various drafts of this paper. This research is supported by NIH-NEI Grant RO1-EY05087 and by AFOSR Grant F49620-93-1-0407. Correspondence should be addressed to J. M. Wolfe, Center for Ophthalmic Research, 221 Longwood Ave., Boston, MA 02115 (e-mail: wolfe@psyche.mit.edu).

large enough print, it is possible to read text in any portion of the visual field, but it is not possible to read two messages in two parts of the field at one time.

Not all visual processes are limited to one portion of the visual field at a time. As Neisser (1967) suggested, there are *parallel processes* that operate over large portions of the visual field at one time, and there is a second set of *limited-capacity* processes that are restricted in their operation to a smaller portion of the visual field at one time. In order to cover all of the visual field, these limited-capacity processes must be deployed *serially* from location to location.¹

It is obvious that the visual system's limited resources should not be deployed and are not deployed in a random fashion. They are under attentional control. By "paying attention" to a specific locus in the visual field, we bring to bear at that locus some of these limited-capacity visual processes. The control of this deployment can be exogenous, based on the properties of the visual stimuli, or endogenous, based on the demands of the "user" of the visual system. These are not mutually exclusive. Deployment can be based on the subject's wish to look for a specific visual stimulus. In this paper, we will consider the deployment of attention in the exogenous and mixed cases.

This paper has three objectives.

1. It provides a fairly extensive review of the visual search literature. There has been an explosion of research in visual search in the past few years. Though it cannot claim to be comprehensive, this paper will survey that literature and discuss some of the theoretical issues raised by the data.

2. It describes a revision of our Guided Search model of visual search. The original version (Cave & Wolfe, 1990; Wolfe & Cave, 1989; Wolfe, Cave, & Franzel, 1989) requires modification in light of new data and further thought—hence, Guided Search 2.0.

3. It presents a computer simulation of substantial parts of the model. The simulation presented here can, with a single setting of its parameters, produce data that closely resemble human data for a wide variety of standard search tasks.

The review of the literature is presented in the context of the Guided Search model. For the reader with an interest in specific issues in the visual search literature, I have tried to provide section headings that will permit a selective reading of the paper. Part 1 describes the scope of the problem and the details of the guided search model and its simulation. Readers uninterested in those details might skip to the summary of Part 1. Part 2 reviews visual search data on searches for targets defined by a single feature and shows that the simulation can reproduce the results of a variety of feature search results. In Part 3, we turn to conjunction searches and serial searches. Here, too, the simulation reproduces the results of experiments with human subjects. Finally, Part 4 discusses some of the omissions and errors of the current version of Guided Search.

1. GUIDED SEARCH 2.0 Model and Simulation

Some Recent History

Briefly, Guided Search holds that the early stages of the visual system process all locations in parallel but are capable of extracting only a limited amount of information from the visual input. Subsequent processes can perform other, more complex tasks but are limited to one or, perhaps, a few spatial locations at a time. Information gathered by the parallel front end is used to restrict the deployment of the limited-capacity processes to those parts of the visual field most likely to contain items of interest. This guidance is not perfect, but it is far more efficient than random deployment of attention. Guided Search began as a modification of Treisman's feature integration model (Treisman, 1988; Treisman & Gelade, 1980). It owes intellectual debts to Neisser's division of visual processing into preattentive and attentive stages (Neisser, 1967), to Hoffman's "two-stage" model (Hoffman, 1978, 1979), and to Egeth's work on conjunctions (Egeth, Virzi, & Garbart, 1984).

Models that live for any time past their birth change as new data appear. This raises certain problems in the literature as people build upon or criticize outmoded versions of a model. For example, the current literature is full of experiments, arguments, and computational theories based on, supporting, or attacking Treisman and Gelade's (1980) original feature integration theory. In the past 10+ years, however, Treisman's group has been very active, and the current version of feature integration is different from the original in important ways. Since the name of the model remains the same, it is not immediately obvious that the earlier version has been superseded. It might help if we borrowed a convention from the software industry and numbered the versions of our models. Accordingly, I have called the present model *Guided Search 2.0* (GS2) to denote what I hope is a substantial upgrade of the original Guided Search model (GS1).

In the first part of the paper, the GS2 model will be described. The details of the simulation and choices that must be made about parameters are also described. Part 1 ends with a discussion of the differences between GS2 and GS1 as well as other models of spatial visual attention.

The Visual Search Paradigm

One of the most profitable paradigms for studying the deployment of attention has been the visual search paradigm in which the subject looks for a *target* item among a set of *distractor* items. Two methods are commonly used. In what can be called the *percent correct* method, the experimenter briefly presents an array of items followed after some interstimulus interval (ISI) by a mask. The subject gives a forced-choice, target-present, or target-absent response and percent correct is measured as a function of ISI. The total number of items (set size)

is varied. If a task can be performed with high accuracy even with short ISIs, and if the $ISI \times$ percent correct function does not change with set size, this is taken to show that all items have been processed in parallel to a degree sufficient to do the task. If larger set sizes require longer ISIs to achieve the same percent correct as smaller set sizes, this is taken to show that a process with limited capacity is required for the search task (Bergen & Julesz, 1983).

The second method measures reaction time (RT) as a function of set size. Sample data are shown in Figure 1. The top panel shows data from a visual search for a vertical line among horizontal distractors and for a red box among green boxes. The amount of time to respond "yes" (a target is present) or "no" (no target is present) is independent of set size, suggesting that all items can be processed for orientation in parallel. By contrast, in the bottom panel, the target is a "T" and the distractors are "L"s. ("T"s and "L"s need to be presented in various orientations. Otherwise grouping and shape information make the task quite easy—see Humphreys, Quinlan, & Riddoch, 1989.) Here response times (RTs) increase roughly linearly with set size, and the slope of the $RT \times$ set size function is about twice as great for blank trials as for target trials. This pattern of results is consistent with a serial self-terminating search in which the subject deploys attention at random from item to item until the target is found or until all items have been checked (Sternberg, 1969; Treisman & Gelade, 1980).

As Townsend (1971, 1976, 1990) has pointed out, $RT \times$ set size functions are not, by themselves, definitive diagnostics for parallel versus serial processes. Some recent work with modifications of the basic search paradigm suggests that serial processing is involved in at least some of the tasks that produce linear $RT \times$ set size slopes with 2:1 ratios between blank and target trial slopes (Kwak, Dagenbach, & Egeth, 1991). However, there are limited-capacity *parallel* processes that could produce results described above as serial. At the very least, increasing slopes of $RT \times$ set size functions can be taken as a measure of the cost of each additional distractor.²

The Guided Search Model

The basic goal of the Guided Search model is to explain and predict the results of visual search experiments. More generally, the model seeks to explain our ability to find a desired visual stimulus in a normal, continuous visual scene. The architecture of the model is shown in Figure 2. This section will describe the model and our simulation of part of it. The subsequent sections will review the relevant human search data and show how simulated results compare to actual results.

Feature Maps

Input and initial processing of stimuli are assumed to be carried out in parallel across the entire visual field.³ At some point, independent parallel representations are generated for a limited set of basic visual features. These

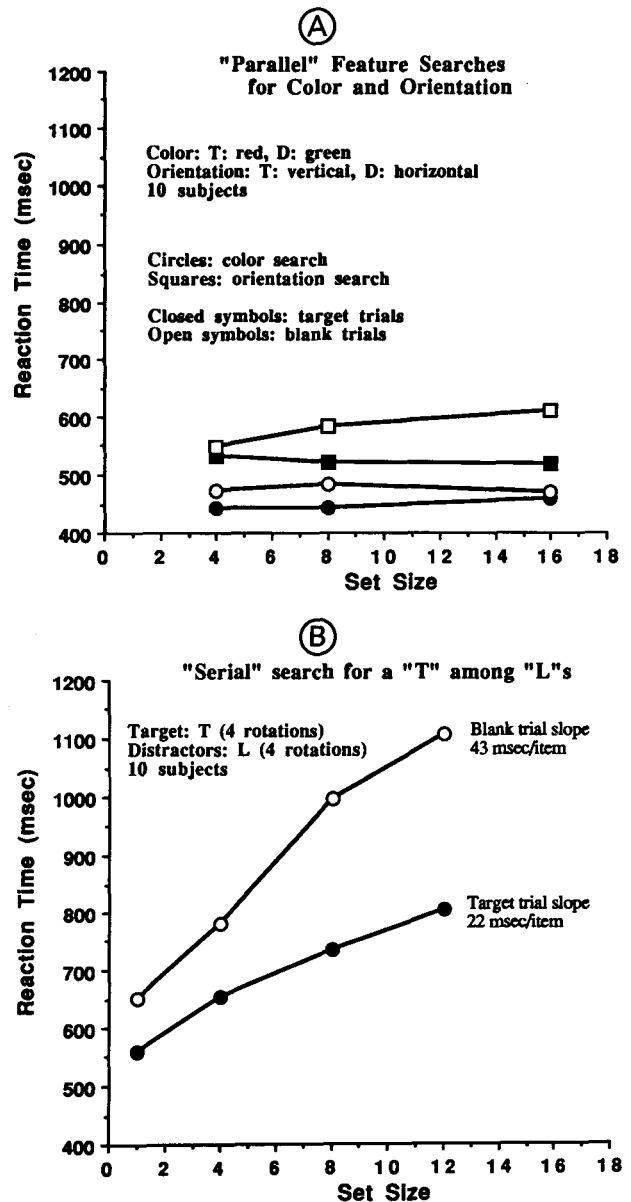


Figure 1. Data from standard "parallel" and "serial" visual search tasks. (A) Subjects search for a red target among green distractors or a vertical target among horizontal distractors (data replotted from Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). (B) Subjects search for a T in one of four possible orientations among similarly rotated Ls. All data points represent the average of 10 subjects (unpublished data from Wolfe & Friedman-Hill, 1990).

will be called *feature maps*. There are, at least, three ways to envision these feature maps.

1. There could be independent maps for each basic color, orientation, size, and so on (i.e., a map for "red," "green," "vertical," etc.). In its most extreme form, there would be no difference in the relationship between the maps for "red" and "green" and the relationship between the maps for "red" and "vertical." Something approaching this position can be found in Treisman (1985, Figure 9).

Basic Components of Guided Search

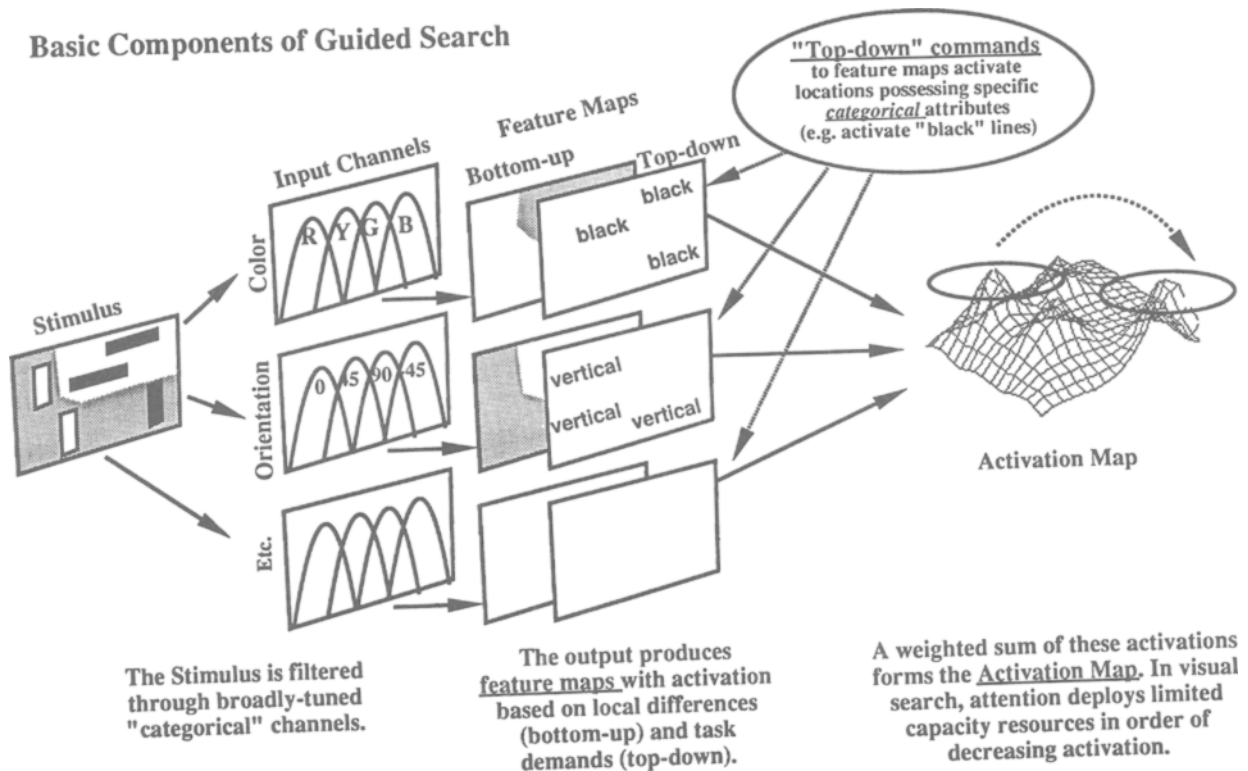


Figure 2. The architecture of GS2.

2. There might be separate maps for each feature type (color, orientation, etc.). Within the map for color, there might be some independence between the representations "red," "green," and so forth, but there would be a difference between the relationship between "red" and "green" and that between "red" and "vertical."

3. Finally, all features could be represented in a single, multidimensional map like that suggested by Duncan and Humphreys (1989), although these researchers appear to have moved their views recently toward something like Option 2 or Option 1, above (Duncan & Humphreys, 1992).

GS2 follows the second alternative of one map per feature type (Wolfe et al., 1990), although recent evidence suggests that there may be differences between features. There may be, for example, multiple color maps but only one orientation map (Nothdurft, 1993b; Wolfe, Chun, & Friedman-Hill, in press).

The set of feature maps supported by current empirical studies will be enumerated in Part 2. In the GS2 *simulation*, only color and orientation feature maps are implemented. Each is treated as an 8×8 array with numerical values designating the orientation or color at each location. For orientation, the values are given in degrees from vertical (tilted right > 0 , left < 0). For color, the values are an arbitrary mapping of spectral colors from red through green onto numerical values (red = 10, yellow = 20, etc.). These are not intended as any serious approximation of the psychophysical representation of color.

Bottom-Up and Top-Down Activation

We assume that the feature maps contain the information required for one to identify the specific colors, orientations, depths and so forth that are present in the visual input. It is to the color feature map, for instance, that one would look for the representation of the precise shade of a glass of Burgundy (though the actual act of identification may require attention). For purposes of visual search, however, the role of parallel processing of basic features is to identify locations that are worthy of further attention. In Guided Search, this is modeled as differential *activation* of locations in each feature map. The greater the activation at a location, the more likely it is that attention will be directed to that location. There are two components of activation: stimulus-driven, *bottom-up* activation, and user-driven, *top-down* activation.

Bottom-Up Activation

Bottom-up activation is a measure of how unusual an item is in its present context. The strength of bottom-up activation for one location is based on the differences between it and items at neighboring loci. It does not depend on the subject's knowledge of the specific search task. Differences refer to differences in the relevant feature space. Thus, the 40° difference between a line tilted 30° from vertical and one tilted 70° is greater than the 20° difference between 30° and 10° . A full model would require an understanding of the metrics of each feature space. This is not trivial. For example, comparisons of

orientations are not based on angular difference alone (see Rock, 1974, for a classic set of illustrations dating back, at least, to Mach, 1886/1959). In orientation search tasks, while 30° and 10° are more similar than 30° and 70° on the basis of their angular separation, we have found that 30° and -30° (mirror symmetric orientations) are more similar than angular separation alone would lead one to suspect (Wolfe & Friedman-Hill, 1992a).

There is evidence, particularly in the orientation domain, that assignments of similarity are based on the outputs of broadly tuned, perhaps *categorical*, channels (Foster & Ward, 1991a, 1991b). In the present simulation, orientations are first "filtered" by channels that respond to steep, shallow, left, or right (Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). The outputs of the channels are defined as follows:

$$\begin{aligned} \text{steep} &= (\cos(2x))^{0.25} && \text{for } -45 < x < 45 \\ \text{shallow} &= \text{abs}(\cos(2x))^{0.25} && \text{for } -90 < x < -45 \\ &&& \text{and } 45 < x < 90 \\ \text{right} &= (\sin(2x))^{0.25} && \text{for } 0 < x < 90 \\ \text{steep} &= \text{abs}(\cos(2x))^{0.25} && \text{for } -90 < x < 0 \end{aligned}$$

where 0° is vertical, and positive values are tilted to the right of vertical. These response profiles are shown in Figure 3. The specific formula is arbitrary. For the present purposes, we require only that the channels be broadly tuned to the categories found in visual search experiments.

The lower panel of Figure 3 shows the channels for color. These are quite arbitrary. They are the third root of triangular functions that have peaks at "red," "yellow," "green," and "blue." The colors are defined in arbitrary units that preserve only their ordinal position

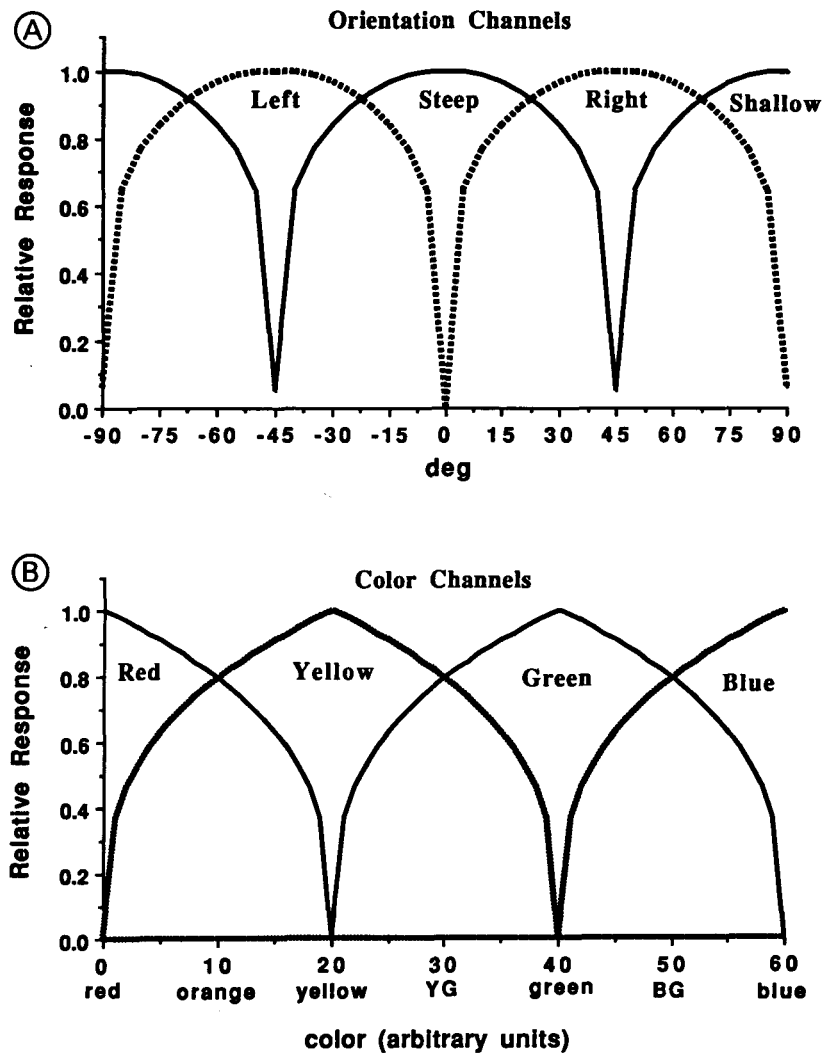


Figure 3. In the simulation, input is filtered through broadly tuned channels shown here for (A) orientation and (B) color. These channels can be considered to be "categorical," with each channel responding to a category such as "red" or "steep."

on the spectrum. However, the underlying principle is the same as in orientation. Color is coarsely categorized before being used for the guidance of attention. Experiments that would adequately define the categories have not been done.

In this simulation, stimuli are either present or absent at a location. This ignores the role of stimulus salience. In reality, a high-contrast vertical line should attract more attention than a low-contrast line. In the simulation, this distinction is not made.

Bottom-up activation is calculated separately for each feature. In the current simulation, the activation for an item is computed as follows.

1. The bottom-up activation for an item is based on comparison with all of the item's neighbors in a 5×5 array centered on the item. For each neighbor, we compute the difference between the output of each broadly tuned channel for the item and that neighbor. Items outside the 5×5 array do not influence the bottom-up activation of the item.

2. The differences are thresholded by a *preattentive just noticeable difference* (Pjnd). The idea of a preattentive jnd comes from work showing that differences that are clearly discriminable in foveal, attentive vision cannot be used by the preattentive feature processes to support parallel visual search. Thus, while subjects can perceive orientation differences of less than 1° in foveal vision, Pjnds for parallel search in orientation are on the order of 10° – 15° (Foster & Ward, 1991a). Preattentive jnds are much greater than standard jnds in color as well (Nagy & Sanchez, 1990). The result is that small differences in color or orientation do not contribute to bottom-up activation.

3. Each difference is multiplied by the strength of the response of the broadly tuned channels to the item itself. Thus, a "strong" stimulus will have a larger bottom-up activation than will a "weak" stimulus.

4. This quantity is divided by the distance from the item to the neighbor. Thus, near neighbors have a stronger effect on the activation than do far neighbors.

5. The contributions from all neighbors are averaged, and this average becomes the bottom-up activation for that item for that feature.

6. Finally, bottom-up activation of any item has a ceiling of 200 arbitrary units of activation. This is based on the intuition that there is a ceiling on texton gradients. (Does a 45° line among horizontal lines "pop out" any less than a 45° line among -45° lines?) Furthermore, if activations were allowed to grow without limit, there would be situations (notably conjunction searches) in which the simulation would make the unfortunate prediction that increasing the difference between target and distractors would make performance worse because the added bottom-up activation would act as noise. (E.g., big red vs. green differences would be more misleading than, say, white vs. pink. In fact, as discussed in Part 2, conjunction search is easier when feature differences are greater.)

Bottom-up calculations could be done in a more physiologically plausible manner. For example, the neighbor-

hood weighting function could be Gaussian rather than linear. However, the important attributes of bottom-up activation are captured by this approximation: (1) Activation of an item increases as the difference between that item and other items increases. (2) Activation decreases as the distance between an item and neighboring items decreases. This can be translated into Julesz's terminology. The patterns of bottom-up activation across an array of items are similar to "texton gradients" (Julesz, 1986). Moreover, it should be clear that this bottom-up activation embodies the similarity principles given in Duncan and Humphreys (1989). As target-distractor similarity increases, target activation will decrease and search will become harder. As distractor-distractor similarity increases, distractor activation will decrease and search will become easier. Bottom-up activation allows search for an "odd man out" even in cases in which the stimuli change from trial to trial.

Top-Down Activation

Bottom-up activation guides attention toward distinctive items in the field. It will not guide attention to a desired item if the featural properties of that item are not unusual. For this, top-down, user-driven activation is needed. Thus, in a search for green crayons in a box of many different hues, bottom-up activation will be noise. A top-down request for "green" will activate locations that might contain the correct crayon. How does the user of the visual system "talk" to the parallel feature processes? Put less colloquially, how is the desired value of the target specified for a given feature? In a series of experiments on orientation (Wolfe, Friedman-Hill, et al., 1992), we found that subjects were able to specify only one of the categories of "steep," "shallow," "left," and "right" and not the actual angle (e.g., 20°) or a combination of categories (e.g., steep and right). Accordingly, in GS2, top-down activation is accomplished by selecting the output of the one broadly tuned channel per feature (e.g., "red" for color and "shallow" for orientation in a search for red horizontal lines).

How is the correct channel selected? The goal of selection is to pick the channel that best differentiates the target from the distractors. This is not necessarily the channel that gives the largest response. For example, consider a search for a 20° tilted line among vertical distractors. The 20° line will produce a larger response from the "steep" channel than from the "right" channel. However, the 0° distractors will produce an even bigger "steep" response than will the 20° items. However, the 0° items will produce no "right" response. Thus, to select the 20° line, the correct choice is to monitor the output of the "right" filter. In the next generation of the model, channel selection will be done by the simulation as it monitors its performance with feedback. In the 2.0 version, selection is done by a series of rules.

1. The targets and distractors are examined to determine what categories are present in the target and whether they are unique. More weight is given to unique categories.

2. For each channel, the weighted response of a channel to the target is compared with the average response of the channel to the distractors. The channel with the greatest positive difference is selected.

As another example, take a 30° target among 50° and -50° distractors. The 30° target is steep and right and produces a larger response to "right." However, one of the distractors is also tilted right while neither is "steep." This makes the unique "steep" channel the correct choice.

The output of the selected channel is rescaled into arbitrary activation units such that the maximum top-down activation is 200 units. In a future version, one might set color activation higher than orientation, on the basis of evidence that color information guides attention more effectively than orientation (Egeth et al., 1984; Poisson & Wilkinson, 1992; Zohary & Hochstein, 1989).

The Activation Map

For the purposes of visual search, each feature module can be thought of as a pair of topographic maps with hills of higher activation marking locations receiving substantial bottom-up or top-down activation. Attention is attracted to the hills. The overall effect of activation in all maps is seen by summing the activations to create an *activation map* (see Figure 2).⁴ For the purposes of guiding attention, activation at a locus does not contain any information about its source. High activation from the color map "looks" the same as high activation from the orientation map in this representation. Thus the activation map contains no information about the burgundy hue of that glass of wine. It should contain a hill of activation at the appropriate location if one were looking for that glass. The purpose of the activation map is to direct attention. In the absence of any endogenous commands to the contrary, attention will be placed at the locus of highest activation. The processes under attentional control then make the decision about the actual identity of the item. In a visual search task, if the target is not found at that locus, attention will be redeployed to the locus with the next highest activation, and so forth.⁵

In a "parallel" visual search, the target item produces the highest level of activation in the activation map, regardless of the number of other items. It is, therefore, the first item to attract attention, regardless of set size. The resulting $RT \times$ set size functions have slopes near 0.0 msec/item. Parallel search is often but not always accompanied by the subjective impression of *pop-out*, the feeling that the target item leaps out of the display to grab attention (Treisman, 1985; Treisman & Gormican, 1988). Pop-out appears to occur when the activation of the target would be particularly high relative to that of the distractors (e.g., a red item among homogeneous green distractors). It may be that only strong bottom-up activation is adequate to yield subjective pop-out. Certainly, it is possible to have parallel search without pop-out (e.g., a red target among heterogeneous distractors; see Duncan, 1989; Wolfe et al., in press; Wolfe et al., 1990). See Wolfe (1992) for further discussion of this point.

The activation map makes it possible to guide attention based on information from more than one feature. This is important in the search for targets not defined by a single unique feature. Suppose, for example, that the target is red and vertical among distractors that are red horizontal and green vertical. If the subject top-down activates all red items and all vertical items, the activation map will tend to show the greatest activation for red-vertical items, even though the target is not uniquely marked in either the color or the orientation feature map. The combination of feature information permits efficient, *guided* search for these conjunctions of features, something not allowed in the original feature integration model (Treisman & Gelade, 1980) but required by more recent data (Dehaene, 1989; Egeth et al., 1984; McLeod et al., 1988; Nakayama & Silverman, 1986; Sagi, 1988; Treisman & Sato, 1990; Wolfe, Cave, & Franzel, 1989; Zohary & Hochstein, 1989). Conjunction searches are discussed more extensively later in the paper.

Cognitive Control of the Activation Map

Top-down and bottom-up activation are of different value in different tasks. Consider the conjunction search for a red vertical target among red horizontal and green vertical items. Since half the items are red, half green and half are vertical, half horizontal, the bottom-up information based on local comparisons (texton gradients) is useless. Top-down specification of the correct color and orientation is of use. By contrast, consider a task where the target is a line of orientation x among distractors of orientation y where x and y vary from trial to trial. Here, only bottom-up activation is of use. It would be helpful if the system were capable of using the informative signals while suppressing the uninformative.

Human experimental data indicates that the contributions of different features and of top-down versus bottom-up activation can be modulated by knowledge of the specific task demands (Francolini & Egeth, 1979). Furthermore, while Jonides and Yantis have shown that, all else being equal, irrelevant transients will capture attention (Jonides & Yantis, 1988; Yantis & Johnson, 1990; Yantis & Jonides, 1990), Yantis has also reported that this capture is not found if the transient is known *not* to mark the target. We have shown that a "snowstorm" of abrupt onset dots, appearing one every 40 msec, need not interfere with serial search tasks (Wolfe & Friedman-Hill, 1990). Pashler (1988) demonstrated that singletons in an irrelevant dimension need not disrupt search. Folk, Remington, and Johnston (1992) have shown that the ability of a stimulus to produce an involuntary shift of attention is modulated by the nature of the primary task (i.e., color cues will force a shift of attention in color tasks but not in abrupt onset tasks and vice versa). Finally, there is growing evidence that subjects can learn specific search tasks (Ahissar & Hochstein, 1992; Karni & Sagi, 1990, 1992; Maljkovic & Nakayama, 1992; Vieira & Treisman, 1988). Subjects may be learning what aspects of the parallel information are of use in each specific task. Accordingly, in GS2,

knowledge of the task requirements is allowed to influence the relative contributions of different feature maps to the activation map.

As with the selection of the optimal channel for top-down activation, the next generation of guided search will set these weights by monitoring its own performance. In the current simulation, weight setting is done by a series of rules. These rules are held constant over all runs of the simulation. Just as we would not expect a human subject to perform well without instruction, these rules can be thought of as the "instructions" to the computer "subject."

For the simulation results reported here, two rules were used.

1. If there is no relevant top-down information from a specific feature, set the weight of top-down activation for that feature to zero. (E.g., if all items are red, do not use top-down color information, or, if orientation changes from trial to trial, do not use top-down orientation information.)

2. If distractors share an attribute with the target, reduce the bottom-up contribution for that feature. (E.g., if the target is red vertical and there are red horizontal distractors, reduce bottom-up color input to the activation map.) The amount of the reduction varies with the ratio of distractor types. Thus, if the target is red and half the distractors are red with the other half green, bottom-up activation for color is reduced to 25% of normal. However, if there are only a relatively few red items in a sea of green, bottom-up information is useful and its contribution increases (Poisson & Wilkinson, 1992; Zohary & Hochstein, 1989). Bottom-up information is never reduced to zero because of the evidence that irrelevant singletons can capture attention and because, in some cases, irrelevant variation in another dimension seems to interfere with search (e.g., Ivry & Cohen, 1990; see also Theeuwes, 1991, 1992, though he has a rather different interpretation of his results).

For computational convenience, activations are rescaled so that distractor activations have a mean of zero. The resulting negative activations have no physiological significance. It is the rank order of activations that is important.

The Deployment of Attention

To recapitulate—the activation map combines information from bottom-up and top-down processes. The contributions of different processes are modulated by some understanding of the task demands. The result is a signal at each location that indicates the likelihood of a target at that location.⁶ Attention is deployed in order of decreasing activation. It shifts from peak to peak in the activation map until the target is found or until the search is terminated. If the search is terminated without a target's having been found, the simulation "guesses." It guesses "no" on the vast majority of trials, generating blank trials and misses. The few "yes" guesses, when wrong, produce the few false alarms seen in this type of experiment.

Search termination is discussed more extensively below and in Chun and Wolfe (1994).

As with many other serial search strategies, the model described here requires that locations that have been examined are somehow removed from further consideration. In the simulation, this is accomplished by setting the activation at a checked location to an arbitrarily low value. This is akin to the "inhibition of return" that has been proposed for these purposes (Klein, 1988; Mackeben & Nakayama, 1988; Posner & Cohen, 1984; Rafal, Calabresi, Brennan, & Sciolto, 1989; Tipper, Driver, & Weaver, 1991). Some such mechanism is likely to be active in visual search, although we have failed to replicate one of the more promising methods for measuring it (Wolfe & Pokorny, 1990). One could make the "inhibition of return" wear off after some period of time, allowing items to be revisited on long searches, but GS2 does not do this.

GS2 treats the deployment of attention as a serial step in processing. Attention is either at one location or another. If it is at location *x*, it can only be redeployed to *y* with the passage of some time. It is important to remember, however, that this serial/parallel distinction is somewhat fuzzy. For example, it seems likely that the attentional stage could be modeled as a limited-capacity parallel process with attentional resources divided between loci in amounts proportional to the relative activations at those loci. Search could proceed in parallel at all locations until one item confirmed its identity as the target or until all items above some activation threshold were rejected (see discussion of blank trials in the next section, and cf. Ratcliff, 1978). While the current version of guided search employs a serial, attentional stage, a limited-capacity parallel attentional stage would not be hard to accommodate within the framework of the model. The central idea of Guided Search, the parallel guidance of attention, could be preserved in either case.

Whether attentional deployment is serial or parallel, it is worth remembering that what is being deployed may also be a limited-capacity *parallel* process. To return to the example of reading; reading occurs at one restricted region at a time. Within that region, the act of reading may involve the use of a more local, parallel process of letter and/or word recognition (see, e.g., Moser, 1991). This would be a limited-capacity spatially parallel process. In order to process the next word, this parallel process would need to be redeployed in a presumably serial fashion. As another example, the feature-binding model of Hummel and Biederman (1992) can be thought of as another case of a limited-capacity parallel process. It acts in parallel to identify objects but will not work over too large or too crowded a visual field. It could be deployed in a serial fashion from locus to locus to perform its parallel identification of the object(s) at each locus.

The Termination of Unsuccessful Searches

In GS2, a search ends when a target is found or when the searcher is reasonably convinced that no target is likely

to be found. As search proceeds from item to item in order of decreasing activation, it will eventually reach items with activations so low that there is little chance that any item of lower activation could be a target item. Thus, it should be possible to set an *activation threshold*. Items with activations below threshold would not attract attention. In positing such a threshold, the model does not assume that the subject has any access to the actual activation levels produced by stimuli in a given experiment. Instead, the correct activation threshold is found by a simple adjustment procedure based on feedback (entirely automatically, we assume). We have modeled the adjustment as a simple staircase. If a trial search ends successfully (target present or absent), the threshold moves up N units of activation, causing subsequent unsuccessful searches to be abandoned more rapidly. If an error is made (hit or false alarm), the threshold moves down KN trials, causing subsequent unsuccessful trials to be abandoned more reluctantly. Like any similar staircase, this one, in effect, estimates one point on a psychometric function corresponding to the proportion of correct responses. The parameter, K , determines the estimated point, and this corresponds to the error rate. For example, a K of 14 yields an error rate of about 7%. A more detailed account of this aspect of the model, including simulation with systematic variation of the free parameters, can be found in Chun and Wolfe (1994). For simulations described here, the staircase parameter was generally fixed at 14, although we will show the effects of systematic variation of this parameter for one simulated search task below. Clearly, in the real world, the tolerance for error and, thus, the staircase parameter will vary with subjects and with instructions to the subjects.

Based on the observation that error rates rise with set size, GS2 has a second rule for search termination. Some very long searches are terminated when the simulation concludes that it almost never takes this long to find a target. Specifically, a random RT is generated on each trial, uniformly distributed between 60% and 100% of the maximum possible RT for the experiment (largest set size \times 50 msec/item). If the simulated search runs longer than this RT, the search is terminated.

All searches terminated without a target's explicitly being found result in a guessing response. Obviously, if a target has not been found, the most plausible guess is that it is not present. However, there *could* be a target lurking in the low-activation items, so the simulation guesses "no" on 97% of the trials and "yes" on the remaining 3%. Erroneous "no" guesses generate miss errors. Erroneous "yes" guesses generate false alarm errors. These are rare in the data and in the simulation.

Finally, we assume that a minimum of two items are examined on each trial before the trial is terminated. This makes the blank trials somewhat slower than the target trials, reflecting real data, and prevents an excess of errors at small set sizes (in which lower densities of items produce lower bottom-up activations).

Variability of Activations

It is evident that the size of the activations that guide attention in this model must have some degree of variability. First, they are based on the output of early vision processes that are limited by noise (Geisler, 1989). Second, if each specific stimulus always produced the same activation value, most "guided" searches would yield 0 msec/item $RT \times$ set size functions. For example, a search for a red vertical target among red horizontal and green vertical distractors would always produce the greatest activation at the target location and, thus, flat search slopes. The failure of predictions of this sort implies that the target item is not the item with the highest activation on every trial. Distractor items could have higher activations than the target if each activation level varied around a mean value. The mean would be determined in the manner described in the preceding sections, summing bottom-up and top-down contributions. The variability reflects the fact that signals in the nervous system are probabilistic in nature.

Thus, like many other visual tasks, visual search can be described as a signal detection problem. There are various items clamoring for attention. The activation of the distractors is the noise. The activation of the target is the noise plus a guiding signal. The efficiency of search is determined by the strength of the attention-attracting signal of the target item relative to the attention-attracting noise of the distractors.

The recasting of guided search as a standard signal detection problem is shown in Figure 4. (See Pavel, Econopoulou, & Landy, 1992, for related ideas.) Figure 4 shows the simplest case of a single target in a homogeneous set of distractors. Distractor activations are drawn from a normal population with mean, N . Target activations are drawn from a distribution with mean $N+S$, where S is the signal. Other search tasks will yield distractor distributions of different shape. For instance, consider a search for a target 30° from vertical among a large number of horizontal distractors and a few vertical distractors. In this case, the distractor distribution would be bimodal, with the horizontals having lower average activation than the verticals have. Though the shape of the activation distributions may change, the basic principles remain the same. On target trials, distractors will be examined if their activation is greater than the target activation. Over many trials, the number of distractors checked will be equal to the proportion of the distractor distribution that is above the average target activation (see Figure 4). On blank trials, distractors will be checked if they have activations above the activation threshold (complicated slightly by the time-out quitting rule). Misses will occur on target trials when the target has an activation *below* the activation threshold and on which the subject guesses "no." The proportion of misses corresponds to the proportion of the target distribution below the activation threshold. As the signal strength increases, the target and distractor activation distributions move apart.

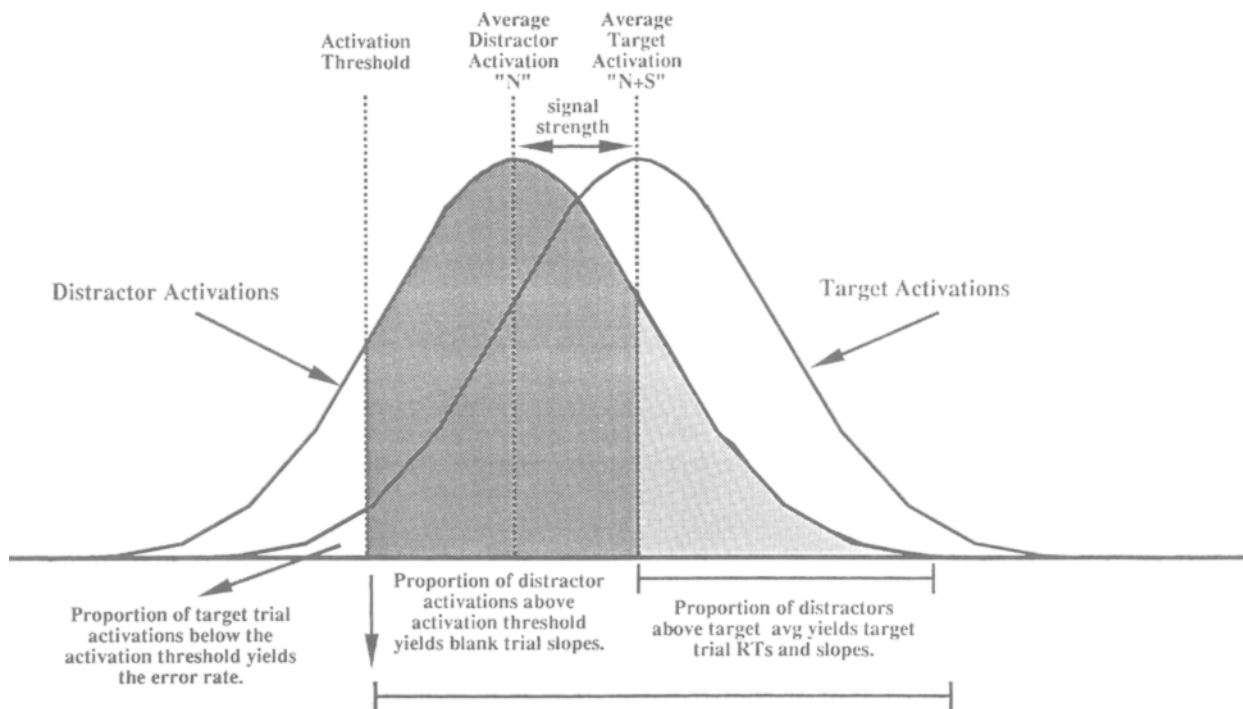


Figure 4. Visual search as a signal detection problem. Each distractor will have some activation, shown here as distributed normally. If preattentive information is available to distinguish the target from the distractors, the target activation will be drawn from a distribution with a higher mean activation (the signal). Response time (RT) on target-present trials is based on the proportion of distractors with activations higher than the mean target activation. RT on blank trials is based on the proportion of distractor activations above the activation threshold. Finally, miss rate is based on the proportion of target activations lying below the activation threshold.

Fewer of the distractors are checked and the slope of $RT \times \text{set size}$ functions decreases. If the activation threshold is raised, the error rate increases and blank trial slopes decrease—a classic speed-accuracy tradeoff.

The present version of the model makes the following simplifying assumptions about variability: (1) Input to the activation map comes from “neurons” that carry signals perturbed by noise. (2) The noise in all “neurons” is normally distributed around the same mean. (3) Noise in one “neuron” is independent of the noise in other “neurons.” (4) More salient stimuli (those producing larger activation signals) will cause more of these input “neurons” to become active. (5) The final activation in the activation map for a specific location is the average of all activated “neurons” at that location. The variability of the final activation will decline with the square root of the number of activated “neurons” (given normally distributed noise). The end result is that as signal strength increases, the number of recruited cells increases and the variability of activation decreases.

This decrease of variability with increasing signal strength serves a useful function in the model. If we return to Figure 4, the ratio of the area of the distractor activation distribution above the average target activation to the area of that distribution above the average target activation threshold gives the ratio of target to blank trial slopes. If variability in the target and distractor distributions is held constant over variations in signal strength, the blank:target

slope ratio varies inversely with signal strength. Looking retrospectively at data from a wide variety of search tasks, we have found no such systematic pattern in real slope ratios. We examined 708 sets of target and blank trial slopes and blank:target slope ratios from subjects tested on a wide variety of different search tasks in our lab. Out of these 708 sets, 187 had target trial slopes less than 5 msec/item. This subset showed a mean slope ratio of 1.87. Another 167 had target slopes greater than 20 msec/item with a mean ratio of 2.00. These were not significantly different. The only obvious difference is that slope ratios are 3.5 times more variable in the group with the shallow slopes (more discussion of this matter can be found in Chun & Wolfe, 1994). Thus, if the activation variability remained constant as activation strength rose, the simulation would produce a systematic change in slope ratios not seen in the experimental results. By assuming that larger signals recruit more “neurons” and that the resulting average signal has lower variability, we produce results comparable to those seen in the experimental data.

In the actual implementation, activation of a “neuron” in the activation map is normally distributed with a standard deviation of 100 units. As the signal strength increases, the standard deviation is $100/\sqrt{(\text{signal}/11)}$. This roughly satisfies the two relevant constraints: first, that the area of the distractor distribution above the average target activation be equal to the area between the average target activation and the activation threshold (see Fig-

ure 4), and, second, that the area of the target distribution below the activation threshold correspond to a reasonable error rate (here 7%). The simulation is quite tolerant of moderate variation in these parameters.

Since it is unlikely that noise ever drops to zero in the real world, a lower limit of 10 is placed on the target standard deviation.

Computing Reaction Time

In the simulation, we assume that it takes an average of 50 msec to attend to and examine each item. We also assume that this time per item is normally distributed, with a standard deviation of 25 msec. RT is simply the number of items checked, multiplied by the time to check each item; 400 msec are added as a baseline for "overhead" (motor programming, etc.). For N items, the RT has a mean of $400 + 50N$ and an SD of $25\sqrt{N}$.

Summary

The model and its simulation can be summarized as follows.

1. For each feature (color and orientation, in the simulation), the image is filtered by broadly tuned channels that correspond roughly to categorical attributes in the feature space ("steep," "shallow," "red," "green," etc.).

2. Based on the outputs of the filters, bottom-up activation of an item is the weighted average of the differences between that item and its neighbors. The contribution of each item to the bottom-up activation of another item is inversely related to *similarity* between the items in the relevant *feature* space, and inversely related to the *distance* between the items in *physical* space.

3. For each feature, the top-down activation of an item is determined by its match to the categorical properties of the designated target. Thus, if the target is tilted 20° to the right of vertical, top-down activation in orientation would be based on the "steepness" and the "right" tiltedness of each item.

4. To create the activation map, we take a weighted sum of all top-down and bottom-up activations. The weightings are based on task demands (e.g., if color is irrelevant to the task, its contribution to the activation map is reduced).

5. In the simulation, activations are rescaled so that the sum of activations on any given trial is zero. Larger activations recruit more "neurons." Each "neuron's" activity is normally distributed with the same variance. Activation at a location is the average of all active "neurons." The result is that activations are variable and their variance decreases as the signal strength increases.

6. Items are examined in a serial, self-terminating manner, beginning with the most active item and continuing in order of decreasing activation until either the target is found or no items remain with activations above the activation threshold.⁷ In GS2, rejected items are not revisited. If a search is terminated without a target's having been

found, the searcher guesses "no" 97% of the time and "yes" 3% of the time.

7. The simulated reaction time has a mean of $400 + 50N$ msec and an SD of $25\sqrt{N}$, where N is the number of items examined on the trial.

8. If an error is made, the activation threshold is lowered. If a trial is successfully completed, the activation threshold is elevated. The ratio of the up and down step sizes in this staircase determines the error rate and is set to 14:1 in the simulation.

There are a significant number of parameters to be set in this model, some more critical than others. For the remainder of this paper, *all* parameters are fixed unless specifically stated. If varied, they are varied only to test the effect of their variation, not to fit the data. Parts 2 and 3 of this paper will review some of the basic findings in visual search and show that this version of Guided Search can reproduce a wide range of those findings. A central claim of this paper is that *the simulation can produce believable results with this single set of parameters.*

Changes from GS1

The basic ideas of the original Guided Search model are preserved in GS2. Parallel feature processes gather information from across the visual field. That information is summed across features to create an activation map that embodies the system's best guesses about the location of a target item. The activations that are summed come from bottom-up, stimulus-driven properties of the display and from the subject's top-down knowledge of the demands of the task. Attention is deployed in a serial manner from item to item until a target item is found or until no likely target items remain. Combination of activation from more than one type of feature allows efficient search by the system as a whole for targets that no single component could find. GS2 differs from the original version in virtually *all* specific details. The most notable of these are the following:

1. Input is now presumed to be filtered by broadly tuned "categorical" channels.

2. Bottom-up activation is sensitive to distance and density effects. In GS1, each item was compared to every other item equally.

3. Unlike GS1, activation is larger for "stronger" stimuli in GS2 (e.g., better exemplars of a category).

4. In GS2, stimuli must be different by a minimum amount before bottom-up activation will be produced. GS1 did not have a threshold on these texton gradients, nor did it have a maximum bottom-up activation as GS2 does.

5. In GS2, top-down activation is categorical.

6. In GS1, the combination of top-down and bottom-up information into an activation map was not a weighted average with weights based on the task demands as it is in GS2.

7. The rules for termination of unsuccessful searches are made explicit in GS2.

8. The properties and sources of noise in activations that guide attention are more explicitly defined in GS2.

Relationship to Other Models

There are several other models and/or simulations of visual search behavior. A point-by-point comparison of each of these with GS2 is beyond the scope of this paper, but a few more general comments are in order. There is a family resemblance between GS2 and a group of models that includes Duncan and Humphreys's "attentional engagement theory" (Duncan & Humphreys, 1989), Asher Cohen's model (Cohen & Ivry, 1991; Cohen & Rafal, 1991), and the current version of feature integration (Treisman & Sato, 1990). Each has a parallel front end that guides or restricts the deployment of attention. These other models are basically verbal and qualitative in nature. Each is filled with good ideas. However, none has been simulated and none is presently described in enough detail to be simulated. For example, one cannot make strong predictions about slope ratios until one has a model of blank trial RTs. More generally, with models with this many "moving parts," it is hard to know whether a solution to one problem (such as "parallel" search for conjunctions) will cause the model to fail on another task (such as search asymmetries in feature searches).

A number of models have been simulated. Most of these are neural net models (see, e.g., Humphreys & Müller, 1993; Ross, Grossberg, & Mingolla, 1993; Sandon & Yanikoglu, 1990). Although GS2 is not a neural net model, there is no reason why it should not be. At the very least, the stages of the model from the input to the production of the activation map would seem to be prime candidates for this type of simulation. The existing simulations tend to be fairly limited in the data that they chose to model. Thus, for instance, Humphreys and Müller's very interesting SERR model is presently implemented on one set of form stimuli (Ts and Ls in various positions). The Ross et al. (1993) model simulates a few search tasks. Bundesen (1990) has attempted to model a very comprehensive set of data. However, that data set does not include the recent findings of efficient search for conjunctions (described below) and various other visual search results that are the starting point for the guided search model. It would be intriguing to see the 2.0 version of his model.

It is interesting that many of the models, particularly the neural net models, rely heavily on grouping processes as the engine of efficient search. This is quite different in spirit from GS2. In GS2, each item attracts attention on the basis of its properties. Bottom-up processes acknowledge the effects of neighbors but do not act to group similar items. In models such as SERR and that of Ross et al. (1993), search is more efficient when distractor items can be grouped. There is an implicit grouping in GS2. Similar items will have similar top-down activations. Spatially contiguous groups of similar items will produce lower bottom-up activations. However, GS2 does nothing explicit with those groupings. Attention is directed to single items. There are no group-wide operations as

there are in other models. There are data that produce problems for a reliance on grouping. These will be discussed below, as will evidence that grouping is important in *some* search tasks. An even-handed assessment might suggest that models like SERR overemphasize the role of grouping, whereas GS2 underestimates that role.

In our Guided Search project, we have attempted to go beyond existing models by constructing a comprehensive model of visual search behavior (e.g., feature integration) that is explicit enough to be simulated. The GS2 result is a model that, when simulated, can produce realistic results for several dozen specific search experiments with a single setting of its parameters. In the following two parts of the paper, a range of search results will be described, with real and simulated data compared whenever possible.

2. PROPERTIES OF THE PREATTENTIVE STAGE Comparing Experimental and Simulated Visual Search

Basic Features in Visual Search

In the following review of some of the literature on preattentive processing, it will be shown how the simulation responds to some standard sets of stimuli. In the current simulation, only color and orientation are implemented. In the first section of Part 2, orientation will be used as a generic *feature* to test the performance of the simulation on *feature searches* (searches in which the target is defined by variation within a single featural dimension). The remainder of Part 2 will review the status of other basic features in visual search.

Orientation

Search for a target orientation among distractors of a second orientation is parallel as long as the two orientations are sufficiently different (Landy & Bergen, 1991; Nothdurft, 1991b; Sagi, 1990). The most impressive data set on this subject may be found in Foster and Ward, 1991a. As the stimuli get closer in orientation, the slopes rise. The guided search simulation produces this result.

Simulation. Unless otherwise stated, all of the simulation data reported here come from a single run of the simulation with a fixed set of parameters. Each simulated search experiment was run for 100 practice trials and 1,000 real trials. Set sizes were 8, 12, 16, 20, 24, 32, 40, and 56. Other parameters were as described in Part 1. Figure 5 shows the slopes of target and blank trial slopes for distractor orientations from 5° to 55° and a target orientation of 0° (vertical). For large differences the slope is near zero, rising to a standard serial, self-terminating search at smaller differences.

Search asymmetries. A characteristic of feature searches is that they are asymmetrical. That is, a search for item *x* among *ys* does not yield the same result as a search for *y* among *xs*. In orientation, for example, Treisman has found and many have replicated that it is easier to search for a tilted line among vertical lines than vice

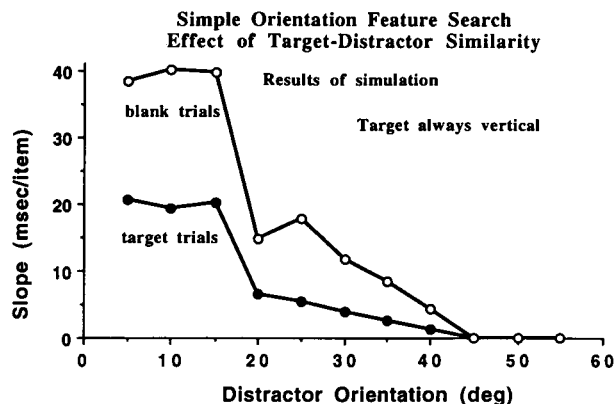


Figure 5. Simulated data for orientation feature search. The subject searches for a vertical target among homogeneous distractors of another orientation. In different 1,000-trial blocks, distractor orientation varies from 5° to 55°. Search changes from serial to "parallel" as target-distractor difference increases.

versa (Treisman & Gormican, 1988). In her account, the vertical line is the prototype and a tilted line is a deviation from the prototype. She argues that it is easier to search for a deviation among homogeneous prototype distractors than to search for a prototype among deviations. Her version is somewhat easier to explain in the color dimension (Treisman & Gormican, 1988, Experiment 7 for color). Red, green, and blue are prototype colors. Magenta, lime, and turquoise would be deviations. Consider a search for red among magenta distractors or vice versa. Red and magenta both contain "redness," the prototype color. Magenta also contains "blueness." A search for red among magenta items is a search for the item that lacks blueness. This is harder than a search for magenta

among red items, a search for the only item that has blueness.

The Guided Search account is similar, and it explains in mechanistic detail how this could come about. The broadly tuned channels that provide input to top-down and bottom-up processors define what Treisman is calling prototypical feature attributes. Consider a search for a 20° target among vertical distractors. The 20° target produces a strong "right" signal against no signal from the 0° distractors. Steepness information is not useful (vertical is steeper than 20°) but a "parallel" search can be based on the fairly large difference in "rightness." If target and distractor are exchanged, the situation alters. Now the target is vertical and the distractors are 20° to the right. The target is steeper than the distractor but the steepness difference between target and distractor is not great. The target produces no response in the left-right channel, so feature search has to be based on the small difference in steepness. The result is a less efficient search.

The simulated result can be seen in Figure 6. Here simulated data are presented for two conditions: 0° target among 20° distractors (square symbols) and 20° target among 0° distractors (circle symbols). Clearly, the standard search asymmetry is produced by the simulation. A similar result is seen in simulation for color. Search for orange among yellow distractors is easier than that for yellow among orange. In this case, the yellow items stimulate only a "yellow" categorical filter. They are prototypical. Orange stimulates "yellow" and "red" categorical filters. Thus, a search for orange can be based on a search for unique red activity, while a search for yellow must be based only on the greater yellow activity relative to the yellow activity generated by the orange distractors.

Asymmetry in conjunction search. Cohen (1993) has found asymmetries in conjunction search that are not

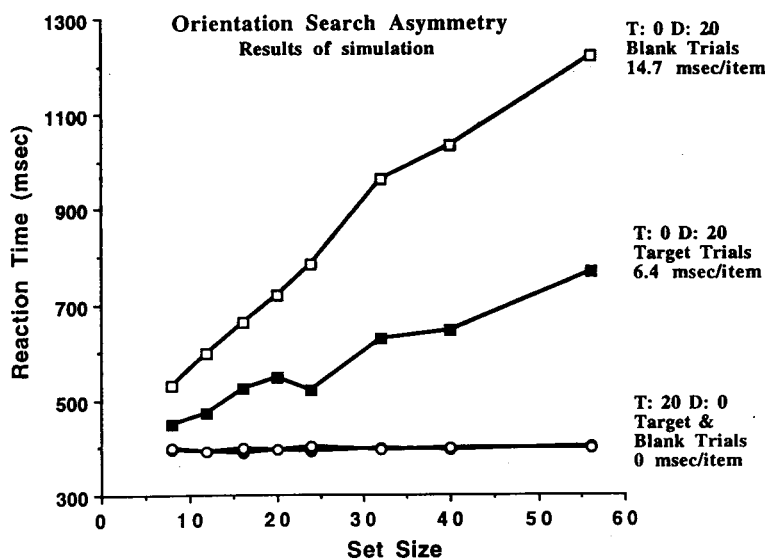


Figure 6. Simulated search for a vertical target among 20° distractors is harder than search for 20° among verticals. This simulated search asymmetry mirrors the human data. Each condition is run for 1,000 trials.

directly predicted by asymmetries in the component feature searches. It is possible, for example, to have symmetrical feature searches (all slopes near zero) with asymmetrical searches for conjunctions of those features. This occurs in the GS2 simulation as well. Consider two colors, yellow and yellow-green, and two orientations, 0° (vertical) and 30° (tilted). All of the feature searches produce simulated target trial slopes under 2 msec/item. (These simulations were performed with the parameters given previously but with 30 practice and 300 real trials and set sizes of 5, 10, and 15 items.) Now consider two conjunction tasks: either a search for a yellow-green tilted target or a search for a yellow vertical target. In both cases, distractors were yellow-green vertical and yellow tilted. Search for the yellow-green tilted target had a target trial slope of 6.3 msec/item. Search for the yellow vertical target had a target trial slope of 17.8 msec/item. Guidance in these conjunction searches is based only on top-down activity. Guidance in the feature searches includes a substantial contribution from bottom-up activity. The addition of a bottom-up signal is adequate in this case to push the feature search slopes to near zero. However, the conjunction search, based solely on the top-down activity, still shows the asymmetry between slower search for a prototype (yellow, vertical) and faster search for a deviation (yellow-green, tilted).

Cohen's (1993) data include cases in which the asymmetry for features runs opposite to the asymmetry for conjunctions of those features, something not yet seen in GS2 and perhaps dependent on the precise form of the broadly tuned input channels.

The preceding is not an account of all search asymmetries. For example, it is easier to find a big item than a small item or a moving item rather than a stationary item (Dick, Ullman, & Sagi, 1987). A more sensory explanation may account for search asymmetries of this sort. Big or moving stimuli produce bigger signals than small or stationary stimuli in the input channels of the visual system. These bigger signals attract attention. A clear case of an asymmetry of this sort is Yantis's finding that abrupt onsets (but not offsets) attract attention (Yantis, 1993; Yantis & Johnson, 1990; Yantis & Jones, 1991; Yantis & Jonides, 1990).

More complex orientation searches. When the distractors are not all of a single orientation, search can become very difficult (Moraglia, 1989). Wolfe, Friedman-Hill, et al. (1992) found that even two distractor orientations could be enough to make search for a target orientation exceedingly inefficient (see also Alkhateeb, Morris, & Ruddock, 1990). Part of this effect seems to be a form of peripheral masking of orientation unrelated to visual search (Andriesen & Bouma, 1976). With that factor controlled, however, search for one orientation among two or more orientations can be efficient if the distractors all lie to one *side* of the target in orientation space (analogous to D'Zmura's, 1991, finding for color) or if the target is the only item in an orientation *category*. The relevant categories seem to be "steep," "shallow," "left,"

"right," and "tilted" (Wolfe, Friedman-Hill, et al., 1992).

Figure 7 shows target trial data from simulations of orientation feature search experiments with heterogeneous distractors. The conditions are taken from Figure 4A of Wolfe, Friedman-Hill, et al. (1992). The simulation uses the same set sizes (4, 8, and 12) and the same number of trials (30 practice, 300 real) as did the original experiment. The results are shown in Figure 7A. Distractors that flank the target by 20° or 40° to either side produce slow, relatively inefficient search. In this general finding, simulation and data agree. However, human subjects did better when the target was vertical, whereas the simulation does worse. Two factors might contribute to this discrepancy: (1) The broadly tuned channels might be the wrong shape. If they were sharply peaked at vertical and horizontal, activations for vertical and horizontal would be boosted. (2) Alignment with the frame of the screen or the gravitational vector might aid search for vertical targets (Marendaz, Stivalet, Barraclough, & Walkowiak, 1993).

The simulation more faithfully reproduces the results of Wolfe, Friedman-Hill, et al. (1992, Experiment 3). In that experiment, the target differed by 40° in orientation from one of the two distractor orientations and 60° from the other. Search was efficient when the target was the *only steep* target and harder when it was the *steepest* or *steep-right* target. The simulation results (Figure 7B) are a good qualitative fit to the real results shown in Figure 6A of Wolfe, Friedman-Hill, et al. (1992).

In other work, we have found that symmetry relationships between items (Wolfe & Friedman-Hill, 1992a) and angular relations between items (Wolfe & Friedman-Hill, 1992b) also influence orientation feature searches. These factors are not included in the current simulation and pose a challenge to a truly quantitative version of the guided search model (or any other model, for that matter). The fine-grained performance of the model is dependent on the details of the parallel feature processes and the broadly tuned input channels. Those details are not known for most basic features. Thus, *any* model of visual search will necessarily have some omissions until additional data become available.

Color

The substantial set of experiments devoted to the search for basic features has produced a set of candidate features. Some, like orientation, are agreed upon by virtually all workers in the field. Others are more problematical. No one doubts that color is a basic feature in visual search (Bundesen & Pedersen, 1983; Carter, 1982; D'Zmura, 1991; Farmer & Taylor, 1980; Green & Anderson, 1956; Moraglia, 1989; Smith, 1962). Laboratory experiments and real world experience make it clear that color is one of the more effective ways for an item to attract attention. This is interesting, given that, unlike such visual properties as orientation, motion, size, and so forth, color is expendable. Black and white pictures contain almost

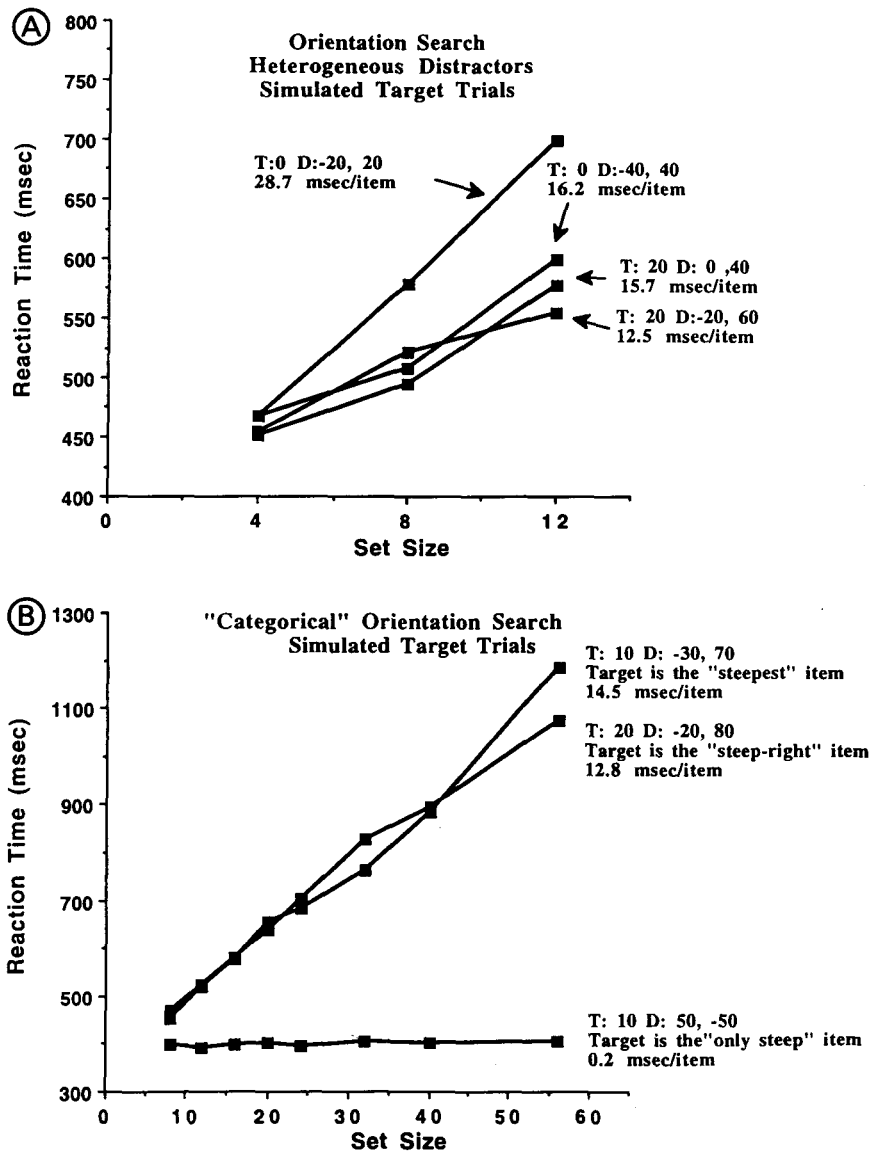


Figure 7. Simulated orientation searches: Target trials only. (A) Search for a target among two different distractor orientations is hard when distractors flank the target (e.g., 0° among -20° and 20°). Compare to Figure 4A of Wolfe, Friedman-Hill, Stewart, and O'Connell (1992). (B) Search among two flanking distractors is easier if the target has unique categorical status (e.g., 10° target among 50° and -50° is the only steep item). Compare to Figure 6A of Wolfe, Friedman-Hill, et al. (1992). Each curve represents 1,000 simulated trials.

as much information as color. Many species lack color vision and many humans have markedly deficient color vision. A color-blind individual can function quite normally. The same would not be true of an orientation-blind individual. For such a marginal ability, color holds a remarkably high place in the pantheon of basic features for visual search.

Few studies of search for color have been strongly grounded in color psychophysics. Two notable exceptions are the recent papers of D'Zmura (1991) and Nagy and Sanchez (1990). Using stimuli in the isoluminant plane of a 3-D color space, D'Zmura (1991) showed that par-

allel search is possible when the target color is on one side of a line drawn through that plane while the distractor colors are on the other side. This implies that parallel search is possible whenever there is only a single distractor color. Obviously, there must be a limit on this ability. Nagy and Sanchez (1990) examined this limit systematically. For a fixed target color, they varied the color of homogeneous distractors in order to find the minimum color differences that would support parallel search. This *preattentive jnd* (Pjnd) can be plotted as a contour surrounding the target's location in color space. The shape and size of the Pjnd contour are quite different from the

shape and size of the standard jnd contour around the same point (Pjnds' contours are larger and do not look like MacAdam ellipses). The same authors subsequently extended their work into the peripheral visual field (Nagy, Sanchez, & Hughes, 1990).

Several recent studies have shown that parallel search remains possible even when distractor colors are quite heterogeneous (e.g., 8 different colors; see Duncan, 1989; Smallman & Boynton, 1990; Wolfe et al., 1990). This demonstrates the existence of top-down activation for color, since any bottom-up information has been reduced or eliminated by the distractor heterogeneity. In the model and the simulation, this top-down activation is categorical. As a result, simulated search for a categorically unique "yellow" among "red" and "green" items is efficient (simulated target trial slope, 0.1 msec/item; blank trial, 0.0 msec/item), whereas search for "yellow-green" among "orange" (yellow-red) and "blue-green" is quite inefficient (simulated slopes of 16 and 27 msec/item) even though the "distances" in the simulated color space are identical in the two conditions. This is a good qualitative match to results for human subjects.

As discussed above, color search is asymmetrical (Treisman & Gormican, 1988), with search for orange among yellow items easier than search for yellow among orange items. According to GS2, search for orange is easier because it can make use of unique categorical information about "redness" that is not available in the search for yellow.

The treatment of color in the present simulation is quite schematic. It would be useful to be more precise about the definition of "color" in visual search. The work of D'Zmura (1991) and of Nagy and Sanchez (1990) brought the requisite precision to the chrominance plane of color space. What happens as luminance is varied? Are black and white "colors," or is luminance a separate dimension? In Smallman and Boynton, 1990, these achromatic colors behave like colors. The picture is somewhat more complicated in a paper by Callaghan (1984). In a variation on a texture segmentation task, she found that brightness variation interfered with tasks based on hue but not vice versa. Luschow and Nothdurft (1993) also show a difference between preattentive processing of color and luminance. More work is needed to clarify this point.

Motion

Only color and orientation are implemented as features in the present simulation. Nevertheless, it is worth briefly reviewing the other basic features that make up the preattentive stage of processing in this and other models. There is little doubt that motion is a basic feature (Dick et al., 1987; Driver, McLeod, & Dienes, 1992b; McLeod et al., 1988; Nakayama & Silverman, 1986; Nothdurft, 1993a; Stone & Mack, 1993; Treisman & Sato, 1990). Like color, it seems to have a very strong ability to group a set of items together (McLeod et al., 1988). Here the search asymmetry is unsurprising, and it is probably based on asymmetries in early visual processing. It is much harder to find a stationary item among moving items than

a moving item among stationary (Dick, 1989). Apparent motion supports parallel search, too (Dick et al., 1987). However, this holds only for short-range apparent motion. Long-range apparent motion targets give rise to serial search and appear to depend on detections of individual onsets and offsets, not on the perception of motion (Dick et al., 1987). Motion may not be a unitary dimension. There is evidence for separate processing of speed and direction (Driver, McLeod, & Dienes, 1992a).

While feature maps that are psychophysically separable need not have physiologically distinct substrates, the physiological underpinnings of preattentive motion processing may be localizable. McLeod, Driver, Dienes, and Crisp (1991) tested a patient with a bilateral lesion in extrastriate visual cortex in the presumed human homologue of monkey area MT, an area specialized for motion processing (Zeki et al., 1991). Unlike normal subjects, this patient cannot restrict visual attention to the moving items in arrays containing moving and nonmoving items.

Size

If targets and distractors are sufficiently different in size, parallel search is found (Treisman & Gelade, 1980). Size has been used as a variable in a variety of search experiments (e.g., Quinlan & Humphreys, 1987; Treisman & Sato, 1990), and the results support the assumption that size is a basic feature. There is no clear pattern of search asymmetry in size. Treisman and Gelade (1988) found that it was easier to find big items among small than vice versa, but this asymmetry vanished if the distractors were held constant across conditions. In that case, small items among medium ones and big items among medium ones produced similar slopes. However, all the slopes were steep in this experiment. Wolfe and Bose (1991) also did a series of experiments with three sizes: big, medium, and small. When big and medium were paired against each other, slopes were shallow with either size as target. When one size was used as the target and the other two as distractors, searches for big and small were easy, whereas searches for medium were quite inefficient. Again, it appears that search is easy when the target and the distractors are on opposite sides of a single line dividing the feature space. Treisman and Gelade (1980) performed a similar experiment and found that the slopes were steep when medium was the target and big and small were the distractors.

Spatial frequency can be considered an aspect of size, though the relationship of size and frequency has not been systematically examined. Spatial frequency behaves like a basic feature (Sagi, 1988) and medium spatial frequencies are hard to find among lower and higher frequencies (Wolfe & Bose, 1991).

Stereoscopic Depth

Parallel search is possible when the target item lies in one depth plane and the distractors lie in a second depth plane (Nakayama & Silverman, 1986; see also Andersen, 1990; Andersen & Kramer, 1993). It is also possible to search for stereoscopic tilt (tilt into the page pops out from

tilt out of the page; see Holliday & Braddick, 1991). It would be interesting to know whether parallel processing survives distractor heterogeneity in stereo. Intuition suggests that it might be possible to find the "closest" or the "farthest" item (with closer easier), and that all other stereoscopic searches with more than one depth plane would be serial. There is some limited work in this area (O'Toole & Walker, 1993), but the results are not clear at this point.

Dichoptic phenomena do not all necessarily support parallel search. Binocular rivalry occurs when different, unfusable stimuli are presented at corresponding loci in each eye (Blake, 1989; Breese, 1909). It is seen as an unstable alternation between the two monocular images. It is perceptually quite salient and is, in some sense, the counterpart of stereopsis (Wolfe, 1986). Moreover, it can be seen to occur across the visual field at many locations at once. Nevertheless, a rivalrous target among fused distractors is *not* found in parallel (Wolfe & Franzel, 1988). Finding a fused target among rivalrous distractors is, if anything, harder still.

Other Depth Cues

In a series of very clever experiments, Enns, Rensink, and their colleagues have shown that a variety of pictorial depth cues are processed in parallel for purposes of visual search. The line relations that allow a 3-D interpretation of 2-D line drawings (e.g., T junctions, Y junctions) can be found efficiently (Enns & Rensink, 1991). Parallel search is also possible if targets and distractors differ in 3-D orientation by line arrangement (Enns & Rensink, 1990b) or shading information (Enns & Rensink, 1990a; see Enns, 1992, for a theoretical review). In a related finding, Ramachandran (1988) has shown parallel search for bumps among holes (and vice versa, though there is an asymmetry) in shapes defined exclusively by shading information. Epstein and Babler (1990) have shown efficient search for slant in the third dimension. Occlusion cues are available preattentively (Enns & Rensink, 1992), as is slant from texture cues (Aks & Enns, 1993) and from shadow cues (Rensink & Cavanagh, 1993) (i.e., an implausible shadow pops out from among items with plausible shadows). There is evidence that the perception of these features is influenced by what can be considered cognitive factors (Brown, Enns, & Greene, 1993).

It would be interesting to pit different 3-D cues against each other. For example, could one search in parallel for bumps defined by shading among bumps defined by stereopsis and/or linear perspective cues (assuming proper control for other distinguishing features), or is the abstract 3-D structure and not its surface medium the basic feature? Evidence from the orientation domain supports the latter hypothesis. Oriented lines can be created in a number of ways: through luminance differences, color, texture, depth, and so forth. All of these support parallel search for orientation, suggesting that it is the abstract property that is coded (Bravo & Blake, 1990; Cavanagh, Arguin, & Treisman, 1990). (For an interesting exception, see O'Connell & Treisman, 1992.)

Binocular Lustre

If a spot is darker than the background in the image presented to one eye and brighter in the other eye, the resulting perception is one of lustre or shininess (Bülthoff & Blake, 1989; Helmholtz, 1866/1962). This binocular *lustre* can be found in parallel among matte distractors (Wolfe & Franzel, 1988). There is probably nothing special about the binocular aspect of binocular lustre. Any shiny spot probably stands out among matte spots, but the experiment has not been done. There is only weak (non-significant) evidence for a search asymmetry. A matte target can be found in parallel among lustrous distractors.

Vernier Offset

Fahle has shown that it is possible to search in parallel for a vernier offset in a line among unbroken lines (Fahle, 1990, 1991a, 1991b). Parallel search remains possible even if the orientation of the lines is varied, reducing the chance that this task is a variant of an orientation search (Wilson, 1986). It is not possible to identify the direction of the offset in parallel.

Curvature

Treisman and Gormican (1988) found that curved lines could be found in parallel among straight distractors. There is an asymmetry: straight targets among curved distractors yield less efficient search. It could be that curvature itself is not a basic feature. Curves are loci of significant change in orientation and might be detected as texton gradients in an orientation map rather than in a map of their own. However, Wolfe, Yee, and Friedman-Hill (1992) did a series of experiments with targets and distractors that had similar local variation in contrast. Curved targets were still efficiently found among noncurved distractors (see also Fahle, 1991b). Angles can be found among curves more efficiently than can curves among angles (Cheal & Lyon, 1992).

Terminators

There is a small group of candidate features that fall under the general heading of form primitives. These are the most problematical of the basic features, with debates swirling around most of them (see the section on form, below). The problem is related to the general problem of determining a set of primitive features for objects and forms. Obviously, some of the preceding features (size, orientation, curvature, vernier offset) can be considered to be aspects of form perception. Julesz suggests adding line terminators to the list (Julesz, 1984; Julesz & Bergen, 1983). Treisman and Gormican (1988), using a C versus O search, have found support for the featural status of line termination (or its opposite, closure). Taylor and Badcock (1988) have found important limitations on the use of terminators as a feature. Julesz, using an E versus S task, had argued that a terminator number could be used to support parallel search. However, Taylor and Badcock found that serial search was required for a target with seven terminators among distractors with only two. They argue that only the simple presence of termi-

nators is detected preattentively. To muddy the topic still further, Cheal and Lyon (1992) got a different asymmetry. Target trial slopes for an E among Ss (three terminators versus two) were steeper than those for an S among Es. This asymmetry was maintained when stimuli were rotated 90°.

Intersections

Julesz (1984) also suggested intersections as a possible feature or texton. Treisman and Gormican (1988) did not find convincing evidence of featural status of intersections in a visual search task. Nothdurft (1991a) used masking experiments in a texture segmentation task to argue against featural status for intersections.

It is worth noting that a number of efforts have been made to explain all form-based *texture segmentation* on the basis of the outputs of spatial filters (oriented Gabor filters, linear center-surround receptive fields, etc.). The general conclusion is that segmentation of pairs like +s and Ls could be accomplished without the need to invoke special mechanisms for intersections, corners, or terminators (Bergen & Adelson, 1988; I. Chen & DeValois, 1993; Graham, Beck, & Sutter, 1992; Graham, Sutter, Venkatesan, & Humaran, 1992; Gurnsey & Browse, 1989; Keeble & Morgan, 1993; Malik & Perona, 1990; Rubenstein & Sagi, 1990; Voorhees & Poggio, 1988). Julesz and Kröse (1988) responded to Bergen and Adelson and to Voorhees and Poggio by filtering a + and an L texture to remove the large spatial frequency energy differences that were supposed to permit texture segmentation. They reported that segmentation was still easy without this information. The issue appears to remain unresolved. One can question the applicability of texture segmentation research to the issue of the status of intersections as visual search features. The connection requires the assumption that "effortless" texture segmentation and "parallel" visual search are roughly equivalent indicators of the featural status of an attribute. As will be discussed below, this is probably not strictly true.

The "Problem" of Form

The first seven "features" listed here seem to be relatively uncontroversial. Numbers 8-11 are all aspects of form, and it is not clear that these make up either the necessary or the sufficient list of form features. For example, Pomerantz and Pristach (1989) have presented some data that suggest that there may be other form features. They demonstrated that the search for "(" among "("s is fairly difficult. However, adding another ")" to both stimuli makes this a relatively easy search for "(" among "("s. A series of experiments of this sort convinced them that emergent features were waiting to be discovered. They eliminated closure, intersections, and terminators as explanations of their results. They did not consider a filtering account for their results but these findings do suggest that other aspects of perceptual organization, perhaps some of the old Gestalt rules, may have roles in parallel visual search.

In a somewhat different context, L. Chen (1982) suggested a topologically based set of form features. Rubin and Kanwisher (1985) objected to the original study, arguing that Chen's topological "holes" were confounded with other candidate features such as perimeter length and area. L. Chen and Zhou have responded to these criticisms with better controlled experiments (L. Chen, 1990; Zhou, L. Chen, & Zhang, 1992; Zhou, Zhang, & L. Chen, 1993). Recently, Enns and Rensink (1991, Experiment 1) presented new experimental evidence against this conjecture, but the issue remains unresolved.

Heathcote and Mewhort (1993) find that relative position can support efficient search (e.g., finding a target that is black on the left and white on the right among distractors that are white on the left and black on the right). They argue that this is a problem for models such as Guided Search or feature integration, but it seems more plausible to argue that this is further evidence that (1) we do not understand preattentive processing of form very well and (2) processing may include phase information (Tadmor & Tolhurst, 1993).⁸ Their finding does raise difficulties for the claim that shading is a basic feature (Ramachandran, 1988), because the black-white versus white-black condition was run as a control (producing inefficient search) to show that bumps were not being distinguished from holes on the basis of simple luminance polarity.

The extensive data set of Cheal and Lyon (1992) encapsulates the problem of form in one paper. They describe a number of search tasks that produce intermediate slopes, suggesting guidance of attention. However, no simple set of form features emerge. For example, a rotated S among Ms was relatively efficient—apparently more efficient than a search for an upright S among Es (rotated Ms). Why? Any feature that one invokes to explain a result of this sort has an ad hoc flavor, yet it would appear that there is more "signal" in the S versus M case than in the S versus E case.

A variety of other candidate form primitives *do not* support parallel search. Treisman and Gormican (1988) failed to find support for the featural status of "juncture," "convergence," or "containment" (whether a dot was inside or outside a figure), and Brown, Weisstein, and May (1992) failed to find evidence for the featural status of volumetric shapes such as Biederman's (1987) geons.

Learning Effects

A possible solution to the problem of form is to argue that new features can be learned. In much of the early work on visual search, letters and numbers were used as stimuli (Duncan, 1980; Egeth, Jonides, & Wall, 1972; Eriksen & Eriksen, 1974; Schneider & Shiffrin, 1977; Treisman & Gelade, 1980). Obviously, in many cases letters are simply a convenient way to embody different form-primitives. For instance, X versus O could involve orientation, curvature, terminators, and so forth. However, there is evidence from experiments with these massively overlearned stimuli which suggests that it may be possible to create new "features" (Schneider & Shiffrin,

1977), including possibly an ability to distinguish letters from numbers in parallel (Egeth et al., 1972; Jonides & Gleitman, 1972). Recently, Wang, Cavanagh, and Green (1992) have shown that some mirror reversed letters can be found efficiently among homogeneous arrays of non-reversed letters. Search for the normal letter among mirror-reversed letters is not as efficient (see also Wang & Cavanagh, 1993).

There are at least three ways to interpret these findings: (1) attention may be guided on the basis of subtle differences in form; (2) new features can be learned; or (3) there is preattentive processing of some semantic information. The first option is implicit in guided search as described above. In order to search successfully, the subject must weight the output of top-down and bottom-up processes in a manner that produces a good guiding signal. If you are looking for red things, you use the output of the channel broadly tuned for red, and so forth. When the signal strength is large, we may assume that a successful weighting function can be obtained quickly. When a potential signal is small, it may take many trials for the weighting function to become optimized. (We plan to implement a learning function of this sort in the next generation of the model.) Almost any pair of stimuli seems likely to produce different responses in *some* channel. Unfortunately, these experiments do not cast a lot of light on the nature of preattentive processing, because it is hard to determine what differences between target and distractor are driving the search. For example, upon closer examination of what we thought was a search for a target orientation among two distractor orientations, we found that subjects were able to use the size of the emergent angles formed by neighboring lines as a cue to target presence (Wolfe & Friedman-Hill, 1992b).

Tuning the weighting function seems more likely than creation of new feature maps (Option 2), although this depends on the definition of a new feature map. A strong claim for a new feature map could be made if a new feature "learned" in visual search could then be used in other tasks. Vieira and Treisman tested this prediction. They practiced subjects on visual search for novel shapes until the subjects performed "parallel" search, but found that these shapes did not behave like features in other tasks (e.g., apparent motion) (Treisman, Vieira, & Hayes, 1992; Vieira & Treisman, 1988). Francolini and Egeth (1979) came to a similar conclusion in the case of the learned distinction between letters and digits (see also Kelly, Harrison, & Hodge, 1991).

Finally, it is possible that there is preattentive processing for semantic information (Johnston, Hawley, & Farnham, 1993). If extensive parallel processing of semantic information could be demonstrated, the need for guided search would be seriously called into question. Why guide attention on the basis of simple features if you can search in parallel for items defined by their meaning? At present, however, the evidence for parallel semantic processing is fairly thin.

The Structure of Items

It may be useful to distinguish between parallel processing of form *features* and processing of another set of properties having to do with the *structural* relationships between those features. For instance, Elder and Zucker (1993) used pairs of lines that were either bent toward each other, something like this—"(')"—or bowed out, something like this—"(')". The tops and bottoms of the lines could be connected to form something like a search for a wasp waist among barrels. They systematically varied the degree of closure of the items in the display and found that search efficiency improved as the features were more and more bound to each other. Baylis and Driver (1993) have shown that the relative position of two points is more rapidly determined if the points lie on a single object than if the equivalent points lie on different objects.

Preattentive processing is sensitive to the hierarchical relationships within items. We have found that preattentive processing can distinguish between properties of whole items and properties of their component parts. Thus, it is possible to search efficiently for conjunctions of two colors if it is the color of a whole item that is conjoined with the color of a part of the item (e.g., find the red cube with a yellow spot on it among red cubes with blue spots and blue cubes with yellow spots). However, it is not possible to search efficiently for conjunctions of the colors of two parts (e.g., find the cube with the red and yellow spots among cubes with red and blue spots and cubes with blue and yellow spots) (Wolfe, Friedman-Hill, & Bilsky, 1994; Wolfe et al., 1990). A similar sensitivity to scale information at early stages in visual processing is presumed to underlie the global precedence effect (Kinchla, 1977; Kinchla & Wolfe, 1979; Navon, 1977; Robertson, Egly, Lamb, & Kerth, 1993) and may be reflected in a part-whole global/local division of labor in neuropsychological data (Farah, 1992; Robertson & Lamb, 1991).

Where Are the Feature Maps for Visual Search?

It is tempting to think that the basic features of visual search and the basic features of early visual processing are the same. We have chromatic mechanisms in early vision, and color is a feature for search. We have channels for spatial frequency. Spatial frequency and/or size is a basic feature for search. There are orientation channels and there is parallel search for orientation, and so on. This attempt to equate parallel processing in search with the early stages of feature extraction is almost undoubtedly incorrect. True, the features of visual search must reflect this earlier processing. After all, one cannot search for information that is not extracted from the input in the first place. However, search appears to operate over a representation of features that is generated later in the sequence of processing steps. For example, consider visual search for orientation. Orientation channels

in early vision are much more narrowly tuned than those measured in visual search (Foster & Ward, 1991a, 1991b; Thomas & Gille, 1979). The orientation of a target in visual search can be derived from a variety of other properties like texture difference (Bravo & Blake, 1990), color, or depth (Cavanagh et al., 1990). This places the extraction of searchable orientation after the extraction of these other stimulus properties in the sequence of processing. Similar results are found with other features. In color, discrimination functions obtained in search tasks differ in scale and shape from those determined in standard color psychophysics tasks (Nagy et al., 1990).

Some information that is available in early vision is *not* available to visual search. For example, consider a task in which the subject needs to find the dot that is white in the left eye and black in the right eye among spots that are black in the left and white in the right. The background is gray. These stimuli will all yield an impression of binocular lustre. Subjects cannot find the target even though the task would be trivial with one eye closed. If the left eye is closed, the target is a black dot among a field of white dots (Wolfe & Franzel, 1988). This information is available in early vision where eye-of-origin information is preserved but is not part of the representation used to support visual search.

Observations of this sort point to a relatively late locus for the parallel feature maps. Can anything be said about their physiological locus? The foregoing arguments make V1, primary visual cortex, an unlikely locus. Extrastriate cortex is much more promising. There is good evidence that different regions of extrastriate cortex are specialized for the processing of different features (Desimone, Schein, Moran, & Ungerleider, 1985; Hubel & Livingstone, 1987; Van Essen & Maunsell, 1983; Zeki, 1978; Zeki et al., 1991). More importantly for the present discussion, attention and task demands influence responses in extrastriate cortex more than at earlier stages of processing (Chelazzi, Miller, Duncan, & Desimone, 1993; Desimone & Ungerleider, 1989; Haenny, Maunsell, & Schiller, 1988; Haenny & Schiller, 1988; Maunsell, Sclar, & Nealey, 1988; Moran & Desimone, 1985; Spitzer, Desimone, & Moran, 1988). Another similarity lies in the ability of extrastriate cortex and visual search to make use of features derived from other features. Thus, subjects can search for oriented lines that are defined by color differences, texture differences, or subjective contours formed by line termination (Cavanagh et al., 1990; Gurnsey, Humphrey, & Kapitan, 1992). Cells in V2 can see such subjective contours, which are apparently unseen by cells in V1 (Peterhans et al., 1986; Von der Heydt, Peterhans, & Baumgartner, 1984).

This is not to say that V1 (or even LGN; see Nothdurft, 1990) have nothing to do with processing the basic features in visual search. Obviously, the visual search basic features are based on processing at earlier stages, and, certainly, building blocks of texture segmentation appear to be available quite early. Cells in V1 will respond more strongly to an oriented line if that line is surrounded by lines of a different orientation rather than by lines of the

same orientation, even if those surrounding lines are outside the classical receptive field (Knierim & Van Essen, 1992) and they can respond to oriented textures even when they cannot respond to the individual lines making up the texture (Nothdurft & Li, 1984).

It would be naive to think of a one-to-one mapping between feature maps in visual search and specific regions of extrastriate cortex. First, the mapping of specific features onto specific areas of cortex is not straightforward even for promising cases such as motion and color (e.g., Logothetis & Charles, 1990). Second, the existence of an independent feature map for a specific feature in an attention model, even if backed by convincing data, by no means requires an independent physiological locus for that map. Orientation and size, for instance, act as basic, apparently separable features in visual search. At the same time, it is unlikely that one will find an extrastriate area specialized for size and insensitive to variation in orientation or vice versa. The two variables seem to be coded by the same cells, but, in visual search, the information can be used as if it were represented in two independent feature maps.

Finally, in thinking about the physiological locus of psychophysical modules such as those proposed for visual search, it is well to remember that we are not birds, cats, or even monkeys—close relatives though we may be. There is some evidence for species differences in search behavior. Allan and Blough found that pigeons showed no search asymmetry for form search tasks that produced strong asymmetries in human data (Allan & Blough, 1989). Von der Heydt and Dursteler (1993) found that their macaques could do conjunction searches with RT \times set size function slopes near zero while their humans could not. Granted, *our* subjects can do some conjunctions with slopes near zero (see Part 3), and their monkeys had many, many trials but the possibility remains that there are differences between species that will complicate any effort to localize the processes of visual search.

The Nature of Basic Features in Visual Search

Adelson and Bergen (1991) make a useful distinction between two tasks in vision: the measurement of “stuff,” such as redness or upward motion, and the labeling of “things,” like a red thing moving up. They argue that the purpose of early vision is to make measurements of visible stuff or “substances” (Adelson & Bergen, 1991). Visual search appears to use the end product of this early vision analysis of the substances of the visual world. The basic features used in visual search can be described as the visual properties of bits of surfaces: color (not component wavelengths), relative (not absolute) size and orientation, position in depth, lustre, motion, and so forth. Another way of saying this would be to propose that many perceptual constancies (color, size, etc.) have been calculated prior to the level of representation of the basic features in visual search.

While the set of basic features in visual search may describe the “stuff” of visual perception, visual search it-

self is the search for “things.” It is the search for a target *item*; the red vertical thing, not redness and verticality. Thus, when subjects search for a conjunction of two features, guided search is much more efficient if those two features are attributes of one thing than if the attributes are spatially separated (see, e.g., Grabowecky & Khurana, 1990). Other examples were discussed in the section on item structure, above. We may conclude that search straddles the “stuff-thing” boundary as it straddles the parallel-serial bottleneck. Guided search could be described as a search for a *thing* based on information about its constituent *stuff*.

Visual Search and Texture Segmentation

In the hunt for basic features, texture segmentation has been suggested as equivalent to search (Julesz, 1984; Rubenstein & Sagi, 1990; Treisman, 1985, 1986b). However, the two tasks do not always produce the same results (Nothdurft, 1991a). Wolfe (1992) has shown that elements that can support highly efficient visual search (e.g., some conjunctions) do not form good textures. Conversely, some elements that form good textures (e.g., elements formed of the conjunction of two orientations) do not permit efficient search. This is not entirely surprising. Search and texture segmentation are not only two different paradigms, they are laboratory versions of two different visual tasks. To continue with Adelson and Bergen’s “stuff-thing” metaphor, search involves looking for some *thing* while textures segment at the points where the *stuff* changes. Perhaps, texture segmentation occurs on the “stuff” side of the “stuff-thing” boundary. This may be seen in the near absence of top-down influences in texture segmentation. While top-down processes can guide search toward an item defined by a conjunction of two features (Wolfe et al., 1989), they cannot produce segmentation of a region defined by that conjunction (Treisman, 1986a). To give one specific example, Nothdurft (1993b) has shown that texton gradients in either color or orientation can form emergent figures. Three non-colinear gradients form an immediately perceived triangle. However, if the vertices of the triangle are defined by *conjunctions* of color and orientation, we have shown that no emergent triangle is seen (Wolfe, Chun, & Friedman-Hill, 1993). The identical conjunctive stimuli will produce search with RT \times set size slopes near zero. This issue is discussed at more length in Wolfe (1992; see Ben-Av, Sagi, & Braun, 1992, for a related point).

To summarize, the parallel feature maps seem to lie toward the end of the early visual processes concerned with visual stuff. The dozen or so basic features are the properties that define visual surfaces and that divide one surface area from the next in texture segmentation. The apparatus of guided search uses knowledge about these surface properties to spatially restrict the hunt for a specific thing.

III. GUIDING ATTENTION

Comparing Experimental and Simulated Visual Search Conjunction Searches

The original motivation for the guided search model was the finding that searches for conjunctions of two basic features did not need to be strictly serial as had been suggested in the original feature integration theory. As discussed above, there is now a large body of data showing that conjunctions can yield RT \times set size target trial slopes that range from apparently “parallel” near 0 msec/item to apparently serial around 20–30 msec/item. The critical variables influencing slopes appear to include stimulus salience (e.g., color saturation) (Duncan & Humphreys, 1989, 1992; Treisman & Sato, 1990; Wolfe et al., 1989), the choice of features (Treisman & Sato, 1990), and, perhaps, the density of the items (high density may impair efficient search; see Cohen & Ivry, 1991).

In guided search, efficient conjunction search is made possible by the combination of attention-guiding activation from two or more basic features. In a search for a red vertical item among red horizontal and green vertical distractors, the color processor would provide activation to guide attention toward red items while the orientation processor would guide attention toward vertical items. The combination of these activations, degraded by the noise described above, would guide attention toward red vertical items. Guided Search, therefore, holds that preattentive processing of two or more features must occur at the same time. An alternative, in the example given here, might be to select the red items and subsequently process those items for orientation (Egeth et al., 1984). In the Subset Searches section of Part 4, evidence is presented against this alternative. For the present, it is worth noting that there is evidence that two different features can be preattentively processed at the same time (Arguin & Cavanaugh, 1988; Downing & Gossman, 1970).

Figure 8 shows performance of the GS2 simulation for conjunction search. It mimics the performance of human subjects. Figure 8A shows RT \times set size functions for 10 repetitions of the simulation. Each run represents 100 practice and 1,000 real trials. The simulated target is red vertical and the distractors are red horizontal and green vertical. The resulting slopes average 7.7 msec/item for target trials and 15.8 msec/item for blank trials. Error rates average 6.8% (5% misses, 1.8% false alarms). Slope ratios average 2.1:1.

These relatively shallow slopes for conjunction searches are comparable to those reported for human subjects (see, e.g., Table 1 in Wolfe et al., 1989), and the 2:1 slope ratio is comparable to the average in human experimental data. The 10 repetitions of the simulation produce similar results each time. By contrast, results from human subjects on conjunction searches are quite variable. Table 1 of Wolfe et al. (1989) shows a comparable *average* target trial slope of 7.5 msec/item. However, the

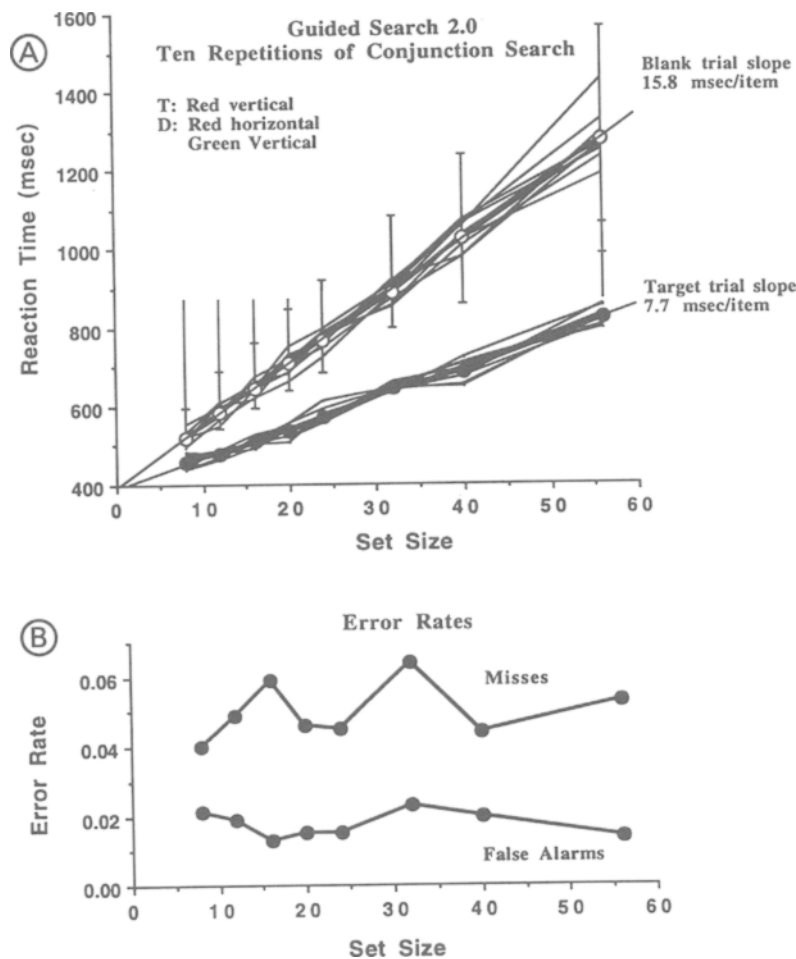


Figure 8. (A) Results for 10 repetitions of a conjunction search for red vertical among red horizontal and green vertical items. Circles are average response times. Error bars are average standard deviations at each point (not *SD* of the averages). Results are comparable to published conjunction data (e.g., Treisman & Sato, 1990, Table 2). (B) Error rates.

range from 2.0 to 20.2 msec/item is much greater than the simulated results. The obvious difference between the simulation and reality is that the simulation is run with all parameters fixed. That is the equivalent of 10 repetitions on a single subject. Moreover, 1,000 trials were simulated and only 300 trials were run on the human subjects. We have only the beginnings of an understanding of the source or sources of individual differences in search performance (O'Neill, Wolfe, & Bilsky, 1993; Tipper & Baylis, 1987). However, as a first approximation, we can simulate intersubject variability by changing the magnitude of the top-down activation, in effect changing the signal strength. In the simulation in Figure 8, top-down signals for color and orientation were fixed at 100 arbitrary units. In a separate simulation, those signals were chosen randomly from a normal distribution with a mean of 100 and *SD* = 30. The simulation was run for 300 trials, using the set sizes of the Wolfe et al.'s (1989) original experiment: 8, 16, 32. Introducing these "individual

differences" does make the simulation more closely resemble the real data. In 10 runs of this simulation, target trial slopes varied from 3.2 to 11.9 msec/item (average, 7.7) and blank trial slopes varied from 10.7 to 25.1 msec/item (average, 16.6). Slope ratios varied from 1.3 to 3.3 (average, 2.5). This is not to say that differences in signal strength are the explanation of individual differences. It is merely worth noting that variation in these parameters produces variations in the data similar to those seen in real populations. The true roots of individual variation are an interesting topic for further research. Within the context of Guided Search, it should be possible to distinguish between variability caused by differences in signal strength (alters target trial slopes) and, for instance, differences in the adjustment of the activation threshold after an error (little effect on target trial slopes).

To return to the conjunction search simulation of Figure 8: A different aspect of variability in the data is also captured by this simulation. As discussed by Ward and

McClelland (1989), a simple model that proposes a serial self-terminating search should predict a greater RT variance for target trials than for blank trials, because on blank trials, one should always examine the entire subset of candidate target items, whereas on target trials, one examines a variable number of items before finding the target. Unlike the simple model, the actual data almost always show that the blank trial variances are higher than the target trial variance. In GS2, all searches are serial self-terminating searches through a subset of the items. However, because the activation threshold varies with time, the size of the subset varies with time. Because the size of the subset varies and because the time for a redeployment of attention is assumed to be somewhat variable, the RT varies. The result, shown in the error bars in Figure 8, is that the model produces the qualitatively correct pattern of variances. Variances are larger for blank than for target trials and grow as set size grows. This matter is considered further in the discussion of serial searches below.

The pattern of error rates, shown in Figure 8B, mirrors some but not all properties of actual data. The overall percentage of errors is reasonable, as is the division between more common misses and less common false alarms. However, in a set of three conjunction tasks run on 30 subjects each, we found that the rate of misses tended to increase with set size while false alarms fell. The present simulation produced little systematic variation with set size. The matter is discussed further in the section on error rates.

As noted, there is a general consensus that relatively shallow slopes for conjunctions require highly salient differences between target and distractors (e.g., Duncan & Humphreys, 1989; Treisman & Sato, 1990). Figure 9

shows a simulation of a search for a red vertical among green vertical and red lines tilted 40° from vertical (low salience) compared to having red distractors 90° from vertical (high salience). Reducing salience increases slopes. This is a close match to the results from an actual experiment of this sort shown in Figure 3 of Wolfe et al. (1989).

In a finding that is not directly modeled by the simulation, Wolfe et al. (1989), Dehaene (1989), and Quinlan and Humphreys (1987) have shown that triple conjunctions are easier than comparable double or standard conjunctions. In a triple conjunction task, the target is defined by three features. The addition of a third source of top-down activation makes search more efficient. Triple conjunctions are not possible in the current simulation, because only color and orientation are simulated. However, it is clear that a simulation with an additional feature would show the correct pattern of results. The addition of a third feature in a triple conjunction would simply increase signal strength and thus increase the efficiency of search.

Various studies have shown an effect of distractor ratios in conjunction search (Egeth et al., 1984; Poisson & Wilkinson, 1992; Zohary & Hochstein, 1989). In brief, the standard condition in which two types of distractors are equally common is especially taxing because there is no useful bottom-up information. If, in a search for red vertical items, there are only 25% red oblique items and 75% green vertical, search is easier because the red items form a small, salient subset. Figure 10 shows the results of a simulation based on Poisson and Wilkinson (1992). In this standard search for a conjunction of color and orientation, the target was red vertical, and the distractors were red, tilted 60° and yellow-green, vertical. To use the same graphing format as Poisson and Wilkinson

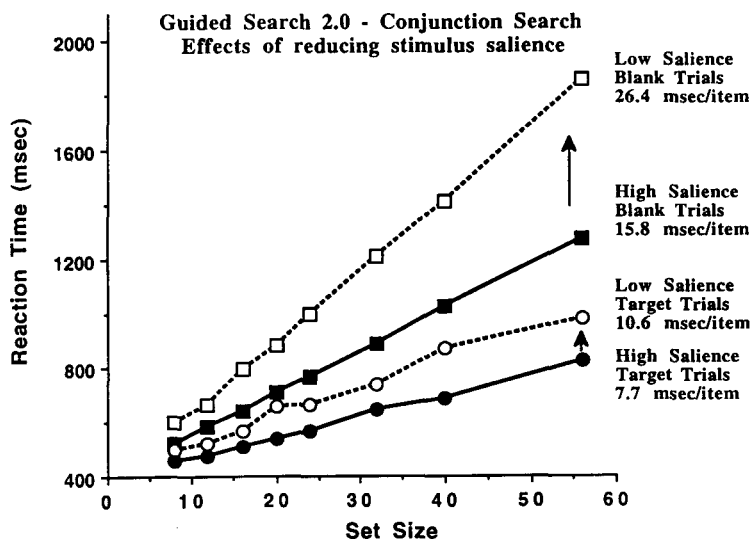


Figure 9. "Guided" search for conjunctions depends on the salience of the stimuli used in both real and simulated data. Here the effects of reduced salience are shown for simulated data. This condition is modeled after Figure 3 in Wolfe, Cave, and Franzel (1989).

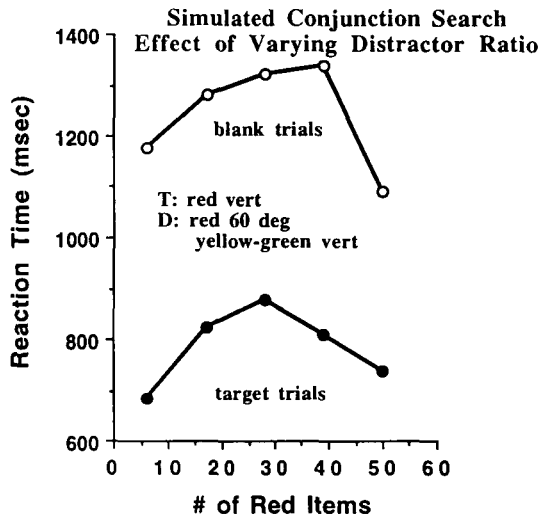


Figure 10. Varying the ratio of different distractor types alters search performance in both real and simulated data, because more bottom-up information is available when there are only a few of one kind of distractor (e.g., a small red subset). Response times are for Set Size 56 only. Simulation is based on Poisson and Wilkinson (1992). See their Figure 2.

did, RT for the largest set size (56) is plotted against the number of red distractor items. The pattern of results seen here is qualitatively similar to the pattern reported by Poisson and Wilkinson. The longest RTs are seen when the two distractor types have equal probability. The human results are skewed to favor color subsets over orientation subsets. That would be accomplished in GS2 by having the activation due to color be larger than that due to orientation.

Search for conjunctions of two instances of a single type of feature are generally inefficient. Thus, search is serial when the target is a red and green item among distractors that are (1) red and blue and (2) green and blue (Wolfe et al., 1990). In the current version of Guided Search, this is modeled by making it impossible to specify two colors or two orientations for search at the same time. There are alternative accounts. It could be that it is possible to give two colors top-down activation at the same time but that an item that contains red *and* green gets no more activation than items that contain either red *or* green. At present the data are unclear on this matter (Wolfe, Yu, Pruszenski, & Cave, 1988). As discussed above, searches for two colors are efficient when subjects search for the color of a whole item conjoined with the color of a component part (e.g., find the red house with yellow windows among red houses with blue windows and blue houses with yellow windows) (Wolfe et al., 1994). Interestingly, this does not appear to work for orientation. Search for orientation \times orientation conjunctions seems to be inevitably inefficient (e.g., search for a vertical thing with an oblique part among horizontals with oblique parts and verticals with horizontal parts). The ability to use con-

figural information in color but not in orientation is an aspect of human data not simulated in GS2.

Serial Searches

Conjunction searches can vary. Efficient searches and shallow slopes are found when strong guidance is available from the parallel processes. Inefficient searches and steep slopes are found when guidance is weak or absent. When no information from the parallel processes can be used to distinguish between targets and distractors, there is no possibility of guided search, and search becomes a serial, self-terminating search through virtually the entire set of items. (In the present model, search becomes serial, though there are other alternatives; see Humphreys & Müller, 1993, and the discussion of this matter earlier in this paper.) We can simulate this search by making the simulated color and orientation of the target identical to the color and orientation of the distractors. In this case, there is no guidance based on basic feature information. The target would have to be defined by properties not processed in parallel. Items have different activations due to the random fluctuations in activation discussed previously. Attention is still deployed from peak to peak in the activation map, but the peaks are random. The resulting search proceeds in this random fashion until attention stumbles on the item labeled as the target in the simulation or until the search is abandoned. Figure 11A shows the averaged results for 10 repetitions and compares that average to the average data from 10 real subjects performing a T versus L task (data from Wolfe & Pokorny, 1990). The bottom set of curves in Figure 11A show the standard deviations for real and simulated data; Figure 11B shows simulated error rates. These results reveal both the strengths and the limitations of the current version of Guided Search.

Simulated RTs had 150 msec added to match the mean target RTs with the real RTs. Clearly the simulated target trial slope matches the human data. Blank trial slopes are somewhat shallower in the simulation. One problem with standard serial self-terminating search models is at least partially overcome in this simulation. If a search were truly serial and self-terminating, then, for a given set size, *all* blank trial searches should take the same time, the time required to exhaustively examine all items. Variances of blank trial RTs should be near zero (Ward & McClelland, 1989). This does not happen in real serial searches where blank trial variances are larger than target trial variances. Dotted lines at the bottom of Figure 11A show the standard deviation for each set size, averaged over 10 repetitions of the simulation. Solid lines show actual data. For target trials (solid symbols), simulated and real standard deviations are comparable. They show similar increases with set size, although the simulated SDs are somewhat smaller than those in the real data. Blank trial SDs are smaller and rise more slowly with set size in the simulation than in the real data (open symbols).

A possible explanation can be found in a comparison of real and simulated RTs and error rates. Simulated tar-

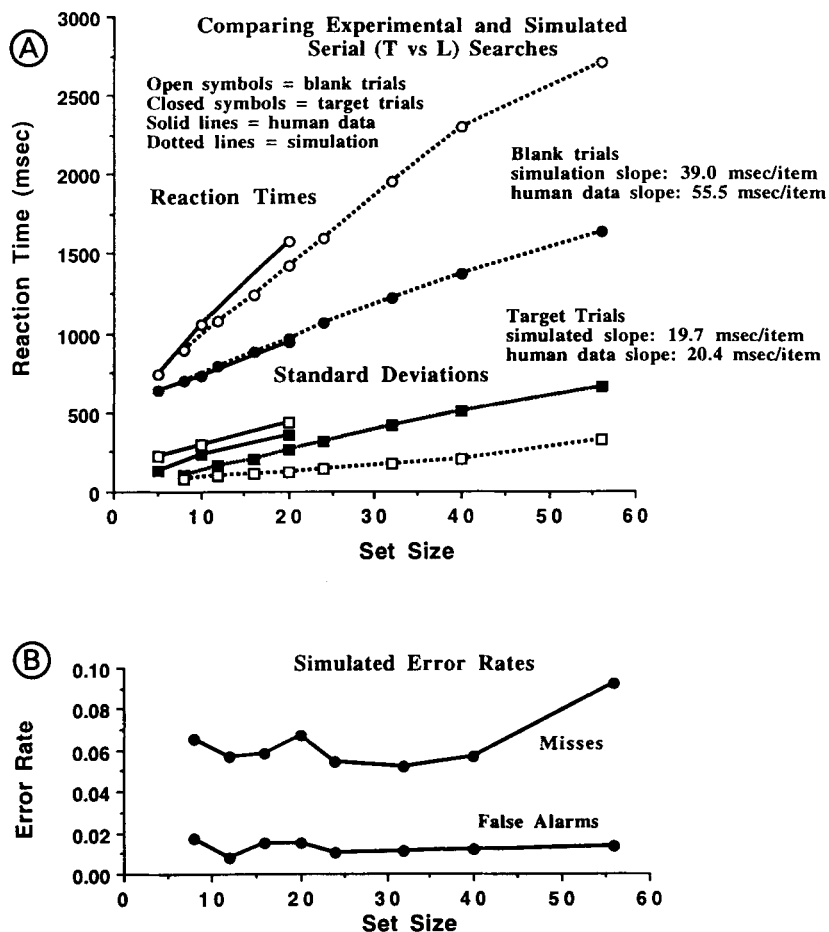


Figure 11. Real and simulated results for serial search. Task for human subjects is a search for a T in one of four rotations among similarly rotated Ls. A 150-msec constant has been added to the simulated response times (RTs). Avg RTs and slopes agree well. Standard deviations (plotted at bottom of panel A) are comparable for target trials but clearly incorrect for blank trials. Error rates are somewhat higher for the simulation (B) than for this set of subjects.

get trial RTs are very similar to real target RTs. The only difference is a 150-msec constant difference that indicates only that the constant added to RTs in the simulation was too small by about 150 msec. The slope of the RT \times set size function for blank trials is shallower for simulated than for real data, and the error rates are somewhat greater. This suggests that real subjects were being more cautious than the simulation. It is possible that real subjects *revisited* previously rejected items on some trials while the simulation never did so. This would produce the longer real RTs on blank trials. It would produce greater variability in blank trial RTs, and it would reduce the error rate. Confirmation of this rechecking hypothesis requires an effective method for tracking the movements of attention during a single trial (Klein, 1988; Wolfe & Pokorny, 1990). Rechecking could be introduced into Guided Search. In the present version, the activation at a locus is suppressed if an item is attended and rejected.

In humans, that suppression must dissipate over time, allowing the item to regain enough activation to attract attention a second time.

Blank Trials and Speed–Accuracy Tradeoffs

The simulation gives up on unsuccessful searches when there are no activations above the activation threshold or when the search has gone on too long. The error rate and the slope of blank trial RT \times set size functions are dependent on the position of the activation threshold. If the threshold is too liberal, error rates rise and RTs fall. If it is too conservative, RTs rise as error rates fall—a classic speed–accuracy tradeoff. The activation threshold is set by a staircase procedure that makes the threshold more liberal after successful trials and more conservative after errors. The staircase step size determines the average position of the threshold. Thus, the simulation will produce speed–accuracy tradeoffs when the staircase parameter is

systematically varied. This was discussed in earlier work (Cave & Wolfe, 1990) and is more extensively considered in Chun and Wolfe (1994).

The simulation results for systematic variation of the staircase parameter are shown in Figure 12. In this simulation, the staircase parameter is systematically varied from 2 to 16. The search task is a color \times orientation conjunction (T, red vertical; D, red horizontal, green vertical). As the parameter increases, miss rates drop from 31.2% to 4.5% and blank trial slopes for this serial search increase from 5.5 to 16.8 msec/item while target trial slopes increase more modestly, rising only from 4.0 to a plateau around 7–8 msec/item. False alarm rates are not altered by variations of this parameter. When Chun and Wolfe (1994) manipulated the payoffs and penalties for speed and accuracy of real subjects, a similar result was obtained. For the other runs of the simulation, this parameter was set to 14. It is worth noting that simulated slopes do not change markedly for parameter settings from 10 to 16, suggesting that the precise setting of this parameter is not critical.

Errors

The pattern of errors changes with the search task. In a large data set from 30 subjects, we found that overall error rate increased as $RT \times$ set size slopes increased (the opposite of a speed-accuracy tradeoff). We also found that error rate became more dependent on set size as $RT \times$ set size slopes increased. For simple feature searches, error rate is independent of set size. For less efficient searches, error rate rises with set size. The simulation re-

produces both of these effects though not perfectly. This can be seen in Figure 13.

The data for serial and conjunction searches are based on 10 runs of the simulation and are the overall error rates for the data shown in Figures 8 and 11, respectively. The data for feature search are based on a single 1,000-trial run. Error rates rise with task difficulty, though the difference between feature and conjunction searches is greater than that seen in real data and the difference between conjunction and serial searches is less than that seen in real data. Error rates increase with set size for the serial search but only at the largest set size. They should increase with set size for conjunction search but do not.

Two parameters govern error rate. One is the activation threshold that is adjusted with each trial. The second is a timing threshold (Have I been searching for an unreasonably long time?). This threshold varies randomly within a range and is not adaptively adjusted in the model. The probable cause of these deviations from the experimental data is the lack of an intelligent way to adjust this parameter.

Summary of Simulation Results

With a single set of parameters, GS2 can reproduce a substantial body of visual search data.

1. It does simple feature searches "in parallel."
2. As the similarity between target and distractor features increases, those parallel searches decrease in efficiency, eventually becoming serial.
3. The simulation shows search asymmetries similar to those in the real data.

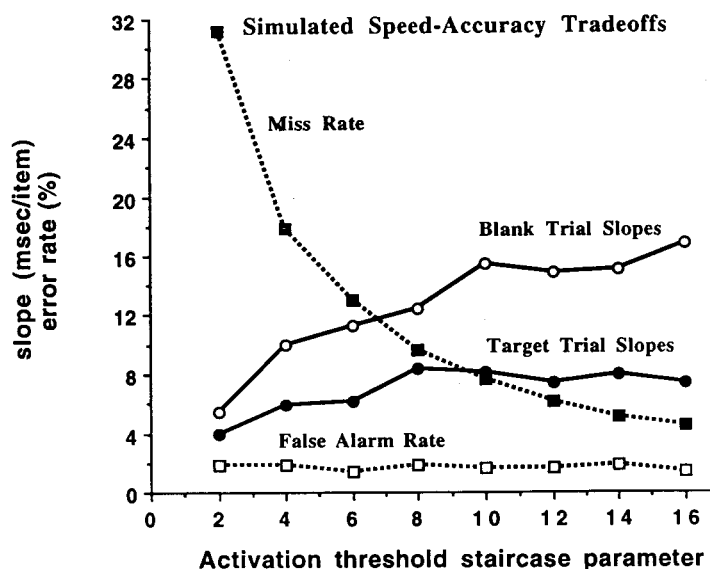


Figure 12. In all previous illustrations of the simulation, the staircase parameter that affects miss rate has been held fixed at 14. Here it varies. As the parameter increases, error rate decreases and blank trial slopes (on this conjunction task) increase: a standard speed-accuracy tradeoff.

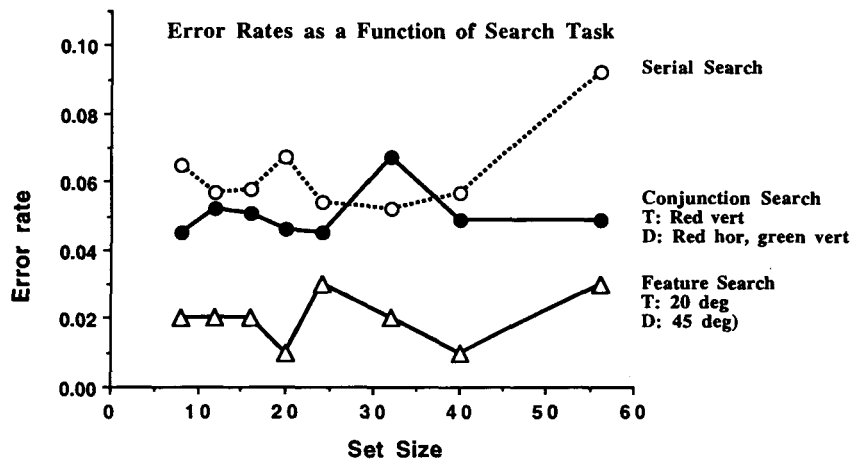


Figure 13. Pattern of error rates as a function of set size for different search tasks. Error rates rise as task difficulty rises. Dependence of error rate on set size is seen only for harder (serial) tasks.

4. With two distractors flanking the target (e.g., search for 10° among -50° and 50°), search is efficient only if the target possesses a unique categorical attribute ("steep" in the example).

5. For salient stimuli, simulated conjunctions of color and orientation produce relatively efficient RT \times set size functions as top-down information about each feature guides attention to the conjunction of two features.

6. As in feature searches, efficiency of search increases as target/distractor similarity decreases.

7. Variability of RTs in the simulation mirrors that seen in real data for conjunction search.

8. If no parallel guidance is possible, search is serial and self-terminating, with slopes comparable to those from human subjects.

9. However, the variability of RTs in simulated serial tasks is too low.

10. For all these various tasks, the simulation uses a single staircase procedure to set its activation threshold so as to produce an acceptable error rate.

11. The staircase parameters can be varied to produce speed-accuracy tradeoffs.

12. The pattern of error rates across tasks mirrors the pattern seen in experimental data, though imperfectly.

4. OMISSIONS, PROBLEMS, AND SPECULATIONS

Obviously, the GS2 simulation does not attempt to address all of the visual search literature. The lacunae fall into three categories: There are some aspects of the data that are addressed by the guided search model that were simply *omitted* from the simulation. There are *problems* with the Guided Search simulation. Real data and the simulation simply do not agree on some points. Of course, it is the model and not the data that must be considered inadequate in these cases. Finally, there are a number of

topics that neither model nor simulation address. For some of these, *speculations* can be offered. Some of the omissions, problems, and speculations are considered in the final section of this paper.

Omissions

Part-Whole Relationships

As was discussed above, it seems impossible for attention to be guided to a conjunction of two colors when those two colors are the colors of two parts of a single item (Wolfe et al., 1990). However, when one color is the general color of a whole item and the other color is the color of a constituent part, guided color \times color conjunction search becomes possible (Wolfe et al., 1994). This suggests some preattentive separation between the representation of whole properties and the representation of their parts. No effort was made to simulate this.

Continuous Stimuli

Related to the preattentive processing of relationships within an item is the preattentive ability to divide the input into a set of items in the first place. In standard laboratory visual search tasks, the items are neatly isolated on a homogeneous background. Obviously, this is not the case in the real world. Do theories like guided search apply to the real world of spatially continuous stimuli? We created stimuli that resemble continuous aerial views of terrain and had subjects search for "polluted" rivers, blue lakes, etc. Conjunction search with these complex stimuli was highly efficient, replicating the standard lab result. However, "serial" searches for stimuli like "S-shaped" river segments were much less efficient than their already inefficient laboratory counterparts (Wolfe, 1994). This suggests that even classic serial searches rely on some parallel preprocessing—the processing required to locate the "items" to be serially searched. This is not part of the current simulation. What is needed is a more realistic early vision "front end" of the simulation.

Efficient search in continuous stimuli is a challenge for models that rely on grouping of similar, contiguous distractors (Ross et al., 1993) because the intervening continuous background would interfere with simple definitions of contiguity. This is not to say that all the blue lakes or yellow rivers might not be grouped—only that it will require a clever grouping mechanism to do the task. GS2 would do the task with the same simple, location-based mechanisms described above. A location with “yellow” and “vertical” remains “yellow” and “vertical” on a continuous background.

The Dynamics of Activation

In the simulation, top-down and bottom-up activations are computed once for each trial, and they remain constant for the duration of the trial. In tasks in which the stimuli are available until the subjects respond, this is undoubtedly incorrect. It is more reasonable but also more computationally laborious to assume that activations are continuously updated. For most tasks simulated here, this is unlikely to be important. Given stimuli that remain unchanged for the duration of a trial, changes in the activations should be limited to noise fluctuations. Updating activation could be important in cases where inhibition of return wears off before the end of a search (Posner & Cohen, 1984). In these cases, attention could unknowingly revisit previously rejected items (see discussion of serial search, above). For other tasks, such as the subset searches described below, the updating of activations during a single trial is an important part of guidance.

Subset Searches

In the guided search architecture described in the first portion of this paper, parallel feature modules pass information only to the activation map and are independent of each other. *Subset searches* reveal that parallel processing in one feature map can be based on information from another. In the standard subset search condition, the target might be a red line of orientation x° among green of x° and red of y° . x and y here denote orientations that change randomly from trial to trial. The only constraint is $\text{abs}(x-y) > 30^\circ$. To describe the task another way, subjects search for the odd man out in the red subset. If feature maps were entirely independent, subjects could activate all red items, but the colorblind orientation process could do nothing when faced with an array of half x° items and half y° items. The expected result would be a serial search through the red subset (in the manner proposed for standard conjunctions by Egeth et al., 1984). In fact, in a variety of these tasks, subjects are clearly *not* performing a serial, self-terminating search through the reds. In some versions, there is no effect of set size. In most cases, target and blank trial RTs and slopes do not differ. It appears that information about the red subset becomes available to the orientation processor. As parallel orientation information is updated, it becomes pos-

sible for the orientation process to operate only over the red subset and to ignore the distractors of irrelevant color (Friedman-Hill & Wolfe, 1992, in press).

Subset search reveals an aspect of guidance that may be of considerable use in real world searches. For example, it should be possible to search in parallel for something moving in the green grass and not be disturbed by other things moving in the blue sky.

This two-step process (get the red items into some sort of group, then do orientation) is a staple of a number of models of standard conjunction search (e.g., Ross et al., 1993). This mandatory grouping is probably not required for conjunction search. We can compare the results of standard conjunction data with our subset search data where the subjects are forced to get the red items first. The pattern of RTs is quite different in the two cases. Standard conjunction search produces twice the slope on blank trials as on target trials and longer blank RTs than target RTs at all set sizes. Subset searches produce blank and target RTs that are approximately the same as one another. Additionally, subset search RTs are 150–200 msec slower on target trials than standard conjunction search RTs—reflecting, we assume, the time required to select the red items and resubmit them for orientation processing.

Preattentive Details

Another substantial omission in the current simulation has been mentioned previously. The true performance of a real or simulated guided search mechanism will be dependent on the properties of the initial processing of basic visual properties such as color, orientation, and so forth. Thus, it will make a difference if the “signal” produced by a red stimulus is, for some reason, bigger than the signal produced by a blue stimulus, and it will make a difference if orientation channels have a 10° or 20° bandwidth, and so on. Correction of this omission awaits more data.

Regional Variation

All else being equal, RTs are faster and error rates are lower for targets presented near fixation (Cheal & Lyon, 1989; Chun & Wolfe, 1994; Efron, Yund, & Nichols, 1987; Sekuler & Ball, 1986) and also for targets in the right visual field (Yund, Efron, & Nichols, 1990). This could be modeled in a number of ways. For instance, there could be a difference in baseline levels of activation biasing movements of attention toward some locations. Alternatively, there could be differences in bottom-up or top-down processing depending on location. (Note that according to Efron et al., 1987, there is a sex difference in regional variation.)

Most interestingly, there is evidence that the attentional visual field is not the same as the classical visual field (Ball, Roenker, & Bruni, 1990) and that cognitive load and factors such as age can change the ability to attend to loci away from fixation (Ball, Beard, Roenker, Miller,

& Griggs, 1988; Ball, Owsley, Sloane, Roenker, & Bruni, 1993; Ball et al., 1990; Graves et al., 1993; Sekuler & Ball, 1986).

Briefly Presented Stimuli

In the visual search paradigm simulated here, stimuli are visible until the subject responds. In many visual search experiments, stimuli are presented briefly (e.g., Bergen & Julesz, 1983; Kröse & Julesz, 1989; Sagi & Julesz, 1985a, 1985b). This method has various advantages. For instance, it rules out systematic eye movements. The dependent measure in these experiments is percent correct, rather than RT. Easy searches produce high percent correct at all set sizes, whereas less efficient searches produce high percent correct only at low set sizes. The GS2 simulation can be easily modified to simulate the brief presentation paradigm. Assume 150 msec of processing time. At this SOA, Guided Search proposes that the subject will be able to examine about three items. To simulate this paradigm, the simulation gets a "hit" if it finds the target among the three highest activations. Otherwise, it guesses with the ratio of "no" to "yes" guesses driven by the guessing parameter that otherwise drives search termination. (I.e., if your error rate is low, guess "no" because you are getting all the yesses without guessing. If the error rate is high, guess "yes" half the time and "no" half the time.) Figure 14 shows results for a simulation of this sort of feature, conjunction, and serial searches. The results are comparable to those obtained from human subjects. Feature search can be done in a flash. Conjunction search is accurate at low set sizes and falls as set size increases. Serial search is harder still.

Problems

Although the simulation produces a fairly faithful reproduction of the results of many human visual search tasks, there are points of disagreement between the simulation and the real data. Some of these are quantitative. For example, the model produces search asymmetries, but the magnitudes of those asymmetries often differ significantly from the data. Many of these quantitative differences are, no doubt, due to the unreal nature of the feature maps. If the analysis of basic features is unrealistic, the precise output of the simulation will be incorrect.

The more serious errors are those that are *qualitative*, because these suggest errors in the general form of the model. Two are discussed here.

Error Rates

As noted at the end of Part 3, the patterns of error rate and distribution as a function of search task are not as close to the human data as one would desire. The probable cause is the failure to adaptively adjust the timing parameter that causes some long trials to be abandoned early.

Variability

As Ward and McClelland (1989) and others have noted, any model that proposes a serial self-terminating search for certain visual search tasks runs into difficulties with the variability of the experimental data. These models predict less variability on blank trials than on target trials while the data tend to show the opposite. This was a failing of the original version of guided search. It is substantially but not completely ameliorated in GS2. For conjunction searches, the simulation now produces greater

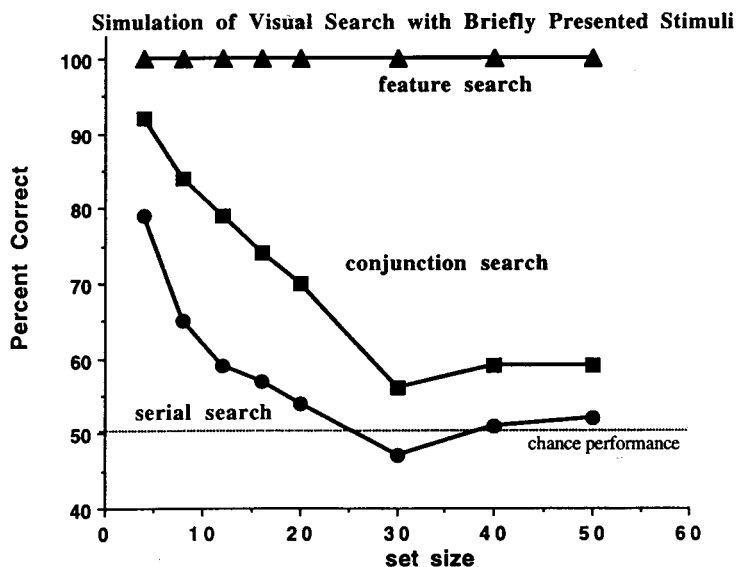


Figure 14. Simulated results of search for briefly presented stimuli. Here the dependent measure is percent correct as a function of set size. With a 150-msec duration, the simulation shows no decline in percent correct in a "parallel" task, some decline in a "guided" (conjunction) task, and more decline in a serial (T versus L) task.

variability on the blank trials. However, as discussed above, pure serial, self-terminating searches still produce too little blank trial variability.

There are at least three sources of variability in the experimental data that are not incorporated into the current simulation. First, the simulation assumes that subjects can perfectly monitor the progress of their searches. That is, the simulation never rechecks an item once it has been rejected. This is probably incorrect. Second, the simulation fixes parameters that probably vary in real experiments. For example, in the staircase that sets the activation threshold for termination of blank trials, the size of the steps is fixed. In real experiments, these parameters and, thus, the RTs for blank trials may vary during an experimental session. Third, the simulation treats all locations in the visual field as equal. Spatial variability would lead to increased temporal variability. Each of these sources of variability will have their primary impact on blank trial RTs and will, thus, raise blank trial RT variability more than target trial variability. As with the account of error rates, this account may seem plausible, but a real test will have to wait until it is simulated.

Speculations

Finally, there are some phenomena of visual search that need to be explained in any complete model of search but about which Guided Search can only speculate at this point.

Illusory Conjunctions

The phenomenon of illusory conjunction certainly seems like one that ought to be useful in constraining models of visual search. Illusory conjunctions occur most readily when an array of stimuli is presented briefly and the subject is asked to state if a particular conjunction was in the display. Subjects will often report the presence of, say, a nonexistent red square if "red" and "square" are present at separate locations in the display (Treisman & Schmidt, 1982). This has sometimes been considered as evidence that basic features are free floating prior to the application of attention, but subsequent research suggests that this may be somewhat of an overstatement (Cohen & Ivry, 1989). Certainly guided search cannot assume free-floating features in preattentive vision. If the combination of information from different feature maps produces guided search for conjunctions, it follows that those feature maps must have preattentively available information about spatial location.

Two observations make illusory conjunctions difficult to fit into a general model of search. First, they are most reliably produced with brief stimuli, meaning that they are rarely "seen" (or, at least, reported) until after the stimulus is gone (though see Treisman & Schmidt, 1982, for a discussion of this point). Second, it is possible to get illusory conjunctions of entities that do not look like basic features (Fang & Wu, 1989; Intraub, 1985; Prinzmetal & Keyser, 1989; Treisman & Souther, 1986; Virzi & Egeth, 1984). These observations suggest (1) that illusory conjunctions may be a family of effects involv-

ing erroneous combination of different sorts of information, and (2) that these errors may require a degraded representation of the input.

Once the stimulus has been removed, the internal representation of that stimulus must begin to decay. Perhaps illusory conjunctions are a product of that decay. While feature information may be registered in a spatially organized map, the quality of the spatial information may degrade after the stimulus vanishes. Thus, after a few moments, if queried about the presence of a red square, attention may find redness and squareness in similar locations and may construct a red square that was not originally present. Such an account would seem to predict that illusory conjunctions should be more common between features that are near neighbors in the display. This is sometimes the case, but certainly not always (Cohen & Ivry, 1989). It is not a serious problem. The degrading of information after stimulus offset can provide several mechanisms for illusory conjunction. For instance, the stimuli will activate internal representations of their "types": "red," "square," "tilted," etc. These types have an existence separate from the specific "tokens" that elicit them (Kanwisher & Driver, 1992; Kanwisher & Potter, 1989, 1990). After the stimulus is gone, the spatial referents of types could become confused, allowing "red" and "square" to become attached to a single location when, in the actual display, they were elicited by quite different items (see Kanwisher, 1991, for similar ideas in a somewhat different framework.).

Density Effects

Regarding a possibly related matter, Cohen and Ivry (1991) report that efficient search for conjunctions occurs only when items are widely spaced and not when they are more densely packed. They propose coarse and fine-grained feature integration processes to account for the data. In the context of guided search, one could speculate that the "hill" of activation produced by each item in each feature map has some girth as well as some height. Thus, if stimuli were tightly packed, the "blue" hill from a blue "O" could overlap with the "X" hill from a yellow "X." This could disrupt guidance and could lead to illusory conjunctions as discussed above. If stimuli were farther apart, there would be fewer opportunities for overlap of the representations. We are presently attempting to study this matter but have found it difficult to replicate the basic Cohen and Ivry results. With rather similar stimuli, our subjects stubbornly continue to produce efficient conjunction searches even for densely packed stimuli.

Nakayama and his colleagues have reported another density effect (Bravo & Nakayama, 1992; Maljkovic & Nakayama, 1992). When a subject needs to report on the form of an item identified by a unique color, a strong density effect makes the task easier with larger set sizes when the target color is unpredictable (bottom-up information only). The density effect decreases and disappears as target color predictability increases (adding top-down information).

Are All Feature Maps Created Equal?

In most models of the parallel, preattentive stage, processing is assumed to occur in a similar fashion for different features. Thus, for instance, in the absence of evidence one way or the other, we assume that top-down processing is categorical for all features because it is categorical for orientation (and, probably, for color and size). However, there are some indications that different features may be processed in qualitatively different ways. For example, Nothdurft (1991b) has done a series of experiments in which subjects view a field of line segments of one orientation. If three segments are rotated to a different orientation, they are immediately seen as the vertices of a triangle, and subjects can determine the orientation of this virtual triangle in a brief exposure. The vertices of the triangle can be three lines of the same orientation or three lines of different orientations. The contrast with the background creates the triangle. We have replicated this result (Chun et al., 1992; Wolfe et al., 1993). However, when we tried the same experiment with color rather than orientation, we found that three items of one color on a background of another color made a more easily perceived triangle than did three items, each of a unique color. The cost of color heterogeneity in this task suggests, though it does not prove, that the preattentive representation of color may be different from the preattentive representation of orientation. Specifically, "red," "green," and "blue" may have semi-independent representations in a way that "vertical," "horizontal," and "oblique" do not. Nothdurft (1993b) has recently reported similar results.

Detection Versus Identification

Because all responses in Guided Search are made after attention finds the target item, the model, in its present form, makes no provision for responses based directly on the contents of preattentive feature maps. Such responses are certainly implicit if not explicit in models that talk about parallel searches and serial searches as two qualitatively different types of search. In Guided Search, there are no "parallel" searches in this sense. A search that produces $RT \times \text{set size}$ slopes near zero is a search in which the activation of the target item is higher than that of any distractors on all or nearly all trials. As a result, attention goes directly to the target location. Some of the more interesting evidence in favor of two different search mechanisms comes from evidence that it is possible to detect a target without knowing what or perhaps where it is (Bundesen, 1991; Sagi & Julesz, 1985b). As with illusory conjunctions, these effects are most commonly found with stimuli that are presented only briefly. We may speculate that they occur when the target activation in the activation map is large enough to almost certainly signal the presence of a target but when the information about the source of that activation has been degraded after stimulus offset. One prediction would be that these effects would only occur in searches that pro-

duce near zero $RT \times \text{set size}$ functions in a standard search task. If the target is not the first attended item, then, in the brief presentation case, there will not be enough information to decide whether subsequent items are or are not targets.

CONCLUSIONS

The goal of the Guided Search project is to create an explicit model of human visual search behavior. With a fixed set of parameters, GS2 can reproduce the results of many standard laboratory search tasks.

1. Search is "parallel" for feature searches with large target-distractor differences.
 2. Feature search becomes less efficient as the target-distractor difference declines.
 3. Feature search becomes less efficient as distractor inhomogeneity increases.
 4. Even with inhomogeneous distractors, feature search remains efficient if the target is categorically unique.
 5. Search asymmetries are reproduced for feature searches and can be produced for conjunction searches even in the absence of asymmetrical feature searches.
 6. Conjunction search can be efficient, though generally not as efficient as simple feature search.
 7. Efficiency of conjunction search declines as stimulus salience declines.
 8. Standard conjunction search becomes more efficient if the two distractor types are not present in exactly equal numbers.
 9. In the absence of basic feature information, search is "serial."
 10. Slope ratios are near 2:1 for all search tasks except those with target trial slopes near zero.
 11. Variance of RT increases with set size and is greater for blank trials than for target trials (except for the most "serial" searches, as noted above).
 12. Miss and false alarm error rates are plausible.
 13. Speed-accuracy tradeoffs can be produced by manipulating the computerized subject's search-termination criterion.
 14. Credible individual differences in performance can be modeled by introducing noise into otherwise fixed simulation parameters.
- Finally, some propaganda in support of GS2. What has this model got that other models do not have?
- Current data.* A number of other models seem to take as their start and end the data that undergird Treisman's original version of feature integration theory. For example, there are models that still insist on serial search for all conjunctions. This is now a decade out of date, and new models should be based on new data (as, indeed, Treisman's models are).

A large body of data. As summarized above, the simulation reproduces data from a wide range of search tasks. Many simulations content themselves with a "feature"

search, a "conjunction" search, and a "serial" search. The standard test bed for guided search is a set of 35 different search tasks (not all discussed in this paper).

Explicitness. Simulation of GS2 requires that underlying assumptions, parameters, and so forth be made explicit. A number of more qualitative models in the literature contain some intriguing ideas. There is nothing inherently wrong with qualitative models, but the models of visual search are becoming sufficiently complex so that the qualitative description may not capture their behavior. The parts of the models interact, sometimes in unexpected ways. Only by implementing the model can we be reasonably sure that it does what we say it does.⁹

The Guided Search project is not finished. Returning to the example that opened this paper, recall that hypothetical stack of journals that collapsed as you left your office. One day, perhaps, your robot will clean up the mess before you get back and it will be its Guided Search module that will locate the copy of *Psychonomic Bulletin & Review* that you were reading and put it back at the top of the stack.

REFERENCES

- ADELSON, E. H., & BERGEN, J. R. (1991). The plenoptic function and the elements of early vision. In M. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 3-20). Cambridge, MA: MIT Press.
- AHISSAR, M., & HOCHSTEIN, S. (1992). Perceptual learning: Interactions between task and stimulus specificities. *Investigative Ophthalmology & Visual Science*, **33**, 1262. (Abstract)
- AKS, D. J., & ENNS, J. T. (1993). Early vision's analysis of slant-from-texture. *Investigative Ophthalmology & Visual Science*, **34**, 1185. (Abstract)
- ALKHATEEB, W. F., MORRIS, R. J., & RUDDOCK, K. H. (1990). Effects of stimulus complexity on simple spatial discriminations. *Spatial Vision*, **5**, 129-141.
- ALLAN, S. E., & BLOUGH, D. S. (1989). Feature-based search asymmetries in pigeons and humans. *Perception & Psychophysics*, **46**, 456-464.
- ANDERSEN, G. J. (1990). Focused attention in three-dimensional space. *Perception & Psychophysics*, **47**, 112-120.
- ANDERSEN, G. J., & KRAMER, A. F. (1993). Limits of focused attention in three-dimensional space. *Perception & Psychophysics*, **53**, 658-667.
- ANDRIESEN, J. J., & BOUMA, H. (1976). Eccentric vision: Adverse interactions between line segments. *Vision Research*, **16**, 71-78.
- ARGUIN, M., & CAVANAGH, P. (1988). Parallel processing of two disjunctive targets. *Perception & Psychophysics*, **44**, 22-30.
- 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., OWSLEY, C., SLOANE, M. E., ROENKER, D. L., & BRUNI, J. R. (1993). Visual attention problems as a predictor of vehicle crashes among older drivers. *Investigative Ophthalmology & Visual Science*, **34**, 3110-3123.
- 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.
- BAYLIS, G. C., & DRIVER, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 451-470.
- BEN-AY, M. B., SAGI, D., & BRAUN, J. (1992). Visual attention and perceptual grouping. *Perception & Psychophysics*, **52**, 277-294.
- BERGEN, J. R., & ADELSON, E. H. (1988). Early vision and texture perception. *Nature*, **333**, 363-364.
- BERGEN, J. R., & JULESZ, B. (1983). Rapid discrimination of visual patterns. *IEEE Transactions on Systems, Man, & Cybernetics*, **SMC-13**, 857-863.
- BIEDERMAN, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.
- BLAKE, R. (1989). A neural theory of binocular rivalry. *Psychological Review*, **96**, 145-167.
- BRAVO, M., & BLAKE, R. (1990). Preattentive vision and perceptual groups. *Perception*, **19**, 515-522.
- BRAVO, M. J., & NAKAYAMA, K. (1992). The role of attention in different visual-search tasks. *Perception & Psychophysics*, **51**, 465-472.
- BREESE, B. B. (1909). Binocular rivalry. *Psychological Review*, **16**, 410-415.
- BROWN, J. M., ENNS, J. T., & GREENE, H. (1993). Preattentive processing of line junctions can be altered by perceptual set. *Investigative Ophthalmology & Visual Science*, **34**, 1234. (Abstract)
- BROWN, J. M., WEISSTEIN, N., & MAY, J. G. (1992). Visual search for simple volumetric shapes. *Perception & Psychophysics*, **51**, 40-48.
- BÜLTHOFF, H. H., & BLAKE, A. (1989). Does the seeing brain know physics? *Investigative Ophthalmology & Visual Science*, **30**, 262. (Abstract)
- BUNDESEN, C. (1990). A theory of visual attention. *Psychological Review*, **97**, 523-547.
- BUNDESEN, C. (1991). Visual selection of features and objects: Is location special? A reinterpretation of Nissen's (1985) findings. *Perception & Psychophysics*, **50**, 87-89.
- BUNDESEN, C., & PEDERSEN, L. F. (1983). Color segregation and visual search. *Perception & Psychophysics*, **33**, 487-493.
- CALLAGHAN, T. C. (1984). Dimensional interaction of hue and brightness in preattentive field segregation. *Perception & Psychophysics*, **36**, 25-34.
- CARTER, R. C. (1982). Visual search with color. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 127-136.
- CAVANAGH, P., ARGUIN, M., & TREISMAN, A. (1990). Effect of surface medium on visual search for orientation and size features. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 479-492.
- CAVE, K. R., & PASHLER, H. (1994a). *Visual selection mediated by location 1: Selecting successive visual objects*. Manuscript submitted for publication.
- CAVE, K. R., & PASHLER, H. (1994b). *Visual selection mediated by location 2: Selecting noncontiguous locations*. 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.
- CHEAL, M., & LYON, D. (1989). Attention effects on form discrimination at different eccentricities. *Quarterly Journal of Experimental Psychology*, **41A**, 719-746.
- CHEAL, M., & LYON, D. (1992). Attention in visual search: Multiple search classes. *Perception & Psychophysics*, **52**, 113-138.
- CHELAZZI, L., MILLER, E. K., DUNCAN, J., & DESIMONE, R. (1993). Neural basis for visual search in inferior temporal cortex. *Nature*, **363**, 345-347.
- CHEN, I., & DEVALOIS, R. L. (1993). Even and odd symmetric mechanisms are equally involved in texture discrimination. *Investigative Ophthalmology & Visual Science*, **34**, 1289. (Abstract)
- CHEN, L. (1982). Topological structure in visual perception. *Science*, **218**, 699-700.
- CHEN, L. (1990). Holes and wholes: A reply to Rubin and Kanwisher. *Perception & Psychophysics*, **47**, 47-53.
- CHUN, M. M., & WOLFE, J. M. (1994). *Just say no: How are visual searches terminated when there is no target present?* Manuscript submitted for publication.
- CHUN, M. M., WOLFE, J. M., & FRIEDMAN-HILL, S. R. (1992). Texture gradients group within but not across feature maps. *Investigative Ophthalmology & Visual Science*, **33**, 960. (Abstract)
- COHEN, A. (1993). Asymmetries in visual search for conjunctive targets. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 775-797.
- COHEN, A., & IVRY, R. B. (1989). Illusory conjunction inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 650-663.

- COHEN, A., & IVRY, R. B. (1991). Density effects in conjunction search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 891-901.
- COHEN, A., & RAFAL, R. D. (1991). Attention and feature integration: Illusory conjunctions in a patient with a parietal lobe lesion. *Psychological Science*, **2**, 106-110.
- DEHAENE, S. (1989). Discriminability and dimensionality effects in visual search for featural conjunctions: A functional pop-out. *Perception & Psychophysics*, **46**, 72-80.
- DESIMONE, R., SCHEIN, S. J., MORAN, J., & UNGERLEIDER, L. G. (1985). Contour, color and shape analysis beyond the striate cortex. *Vision Research*, **25**, 441-452.
- DESIMONE, R., & UNGERLEIDER, L. G. (1989). Neural mechanisms of visual processing in monkeys. In F. Boller & J. Grafman (Eds.), *Handbook of neuropsychology* (pp. 267-299). Amsterdam: Elsevier.
- DICK, M. (1989). *Parallel and serial processes in motion detection*. Unpublished doctoral dissertation, Weizmann Institute, Rehovot, Israel.
- DICK, M., ULLMAN, S., & SAGI, D. (1987). Parallel and serial processes in motion detection. *Science*, **237**, 400-402.
- DOWNING, B. D., & GOSSMAN, J. R. (1970). Parallel processing of multi-dimensional stimuli. *Perception & Psychophysics*, **8**, 57-60.
- DRIVER, J., MCLEOD, P., & DIENES, Z. (1992a). Are direction and speed coded independently by the visual system? Evidence from visual search. *Spatial Vision*, **6**, 133-147.
- DRIVER, J., MCLEOD, P., & DIENES, Z. (1992b). Motion coherence and conjunction search: Implications for guided search theory. *Perception & Psychophysics*, **51**, 79-85.
- DUNCAN, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review*, **87**, 272-300.
- DUNCAN, J. (1989). Boundary conditions on parallel processing in human vision. *Perception*, **18**, 457-469.
- DUNCAN, J., & HUMPHREYS, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, **96**, 433-458.
- DUNCAN, J., & HUMPHREYS, G. W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 578-588.
- D'ZMURA, M. (1991). Color in visual search. *Vision Research*, **31**, 951-966.
- ERFON, R., YUND, E. W., & NICHOLS, D. R. (1987). Scanning the visual field without eye movements: A sex difference. *Neuropsychologia*, **25**, 637-644.
- EGETH, H. E., JONIDES, J., & WALL, S. (1972). Parallel processing of multielement displays. *Cognitive Psychology*, **3**, 674-698.
- EGETH, H. E., VIRZI, R. A., & GARBART, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 32-39.
- ELDER, J., & ZUCKER, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, **33**, 981-991.
- ENNS, J. T. (1992). Sensitivity of early human vision to 3-D orientation in line-drawings. *Canadian Journal of Psychology*, **46**, 143-169.
- ENNS, J. T., & RENSINK, R. A. (1990a). Scene based properties influence visual search. *Science*, **247**, 721-723.
- ENNS, J. T., & RENSINK, R. A. (1990b). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, **1**, 323-326.
- ENNS, J. T., & RENSINK, R. A. (1991). Pre-attentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, **98**, 335-351.
- ENNS, J. T., & RENSINK, R. A. (1992). An object completion process in early vision. *Investigative Ophthalmology & Visual Science*, **33**, 1263. (Abstract)
- EPSTEIN, W., & BABLER, T. (1990). In search of depth. *Perception & Psychophysics*, **48**, 68-76.
- ERIKSEN, B. A., & ERIKSEN, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, **16**, 143-149.
- FAHLE, M. (1990). Parallel, semi-parallel, and serial processing of visual hyperacuity. *Human Vision & Electronic Imaging: Models, Methods, & Applications*, **1249**, 147-159.
- FAHLE, M. (1991a). A new elementary feature of vision. *Investigative Ophthalmology & Visual Science*, **32**, 2151-2155.
- FAHLE, M. (1991b). Parallel perception of vernier offsets, curvature, and chevrons in humans. *Vision Research*, **31**, 2149-2184.
- FANG, S.-P., & WU, P. (1989). Illusory conjunctions in the perception of Chinese characters. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 434-447.
- FARAH, M. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, **1**, 165-169.
- FARMER, E. W., & TAYLOR, R. M. (1980). Visual search through color displays: Effects of target-background similarity and background uniformity. *Perception & Psychophysics*, **27**, 267-272.
- FOLK, C. L., REMINGTON, R. W., & JOHNSTON, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance*, **18**, 1030-1044.
- FOSTER, D. H., & WARD, P. A. (1991a). Asymmetries in oriented-line detection indicate two orthogonal filters in early vision. *Proceedings of the Royal Society London: Series B*, **243**, 75-81.
- FOSTER, D. H., & WARD, P. A. (1991b). Horizontal-vertical filters in early vision predict anomalous line-orientation frequencies. *Proceedings of the Royal Society London: Series B*, **243**, 83-86.
- FRANCOLINI, C. M., & EGETH, H. E. (1979). Perceptual selectivity is task dependent: The pop-out effect pops out. *Perception & Psychophysics*, **25**, 99-110.
- FRIEDMAN-HILL, S. R., & WOLFE, J. M. (1992). Activation vs. inhibition in visual search. *Investigative Ophthalmology & Visual Science*, **33**, 1356. (Abstract)
- FRIEDMAN-HILL, S. R., & WOLFE, J. M. (in press). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception & Performance*.
- GEISLER, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, **96**, 267-314.
- GRABOWECKY, M., & KHURANA, B. (1990). Features were meant to be integrated. *Investigative Ophthalmology & Visual Science*, **31**, 105. (Abstract)
- GRAHAM, N., BECK, J., & SUTTER, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: Effects of sign and amount of contrast. *Vision Research*, **32**, 719-743.
- GRAHAM, N., SUTTER, A., VENKATESAN, C., & HUMARAN, M. (1992). Non-linear processes in perceived region segregation: Orientation selectivity of complex channels. *Ophthalmic & Physiological Optics*, **12**, 142-146.
- GRAVES, M. A., BALL, K. K., CISELL, G. M., WEST, R. E., WHORLEY, K. D., & EDWARDS, J. D. (1993). Auditory distraction results in functional visual impairment for some older drivers. *Investigative Ophthalmology & Visual Science*, **34**, 1418. (Abstract)
- GREEN, B. F., & ANDERSON, L. K. (1956). Color coding in a visual search task. *Journal of Experimental Psychology*, **51**, 19-24.
- GURNEY, R., & BROWSE, R. A. (1989). Asymmetries in visual texture discrimination. *Spatial Vision*, **4**, 31-44.
- GURNEY, R., HUMPHREY, G. K., & KAPITAN, P. (1992). Parallel discrimination of subjective contours defined by offset gratings. *Perception & Psychophysics*, **52**, 263-276.
- HAENNY, P. E., MAUNSELL, J. H. R., & SCHILLER, P. H. (1988). State dependent activity in monkey visual cortex: II. Visual and non-visual factors. *Experimental Brain Research*, **69**, 245-259.
- HAENNY, P. E., & SCHILLER, P. H. (1988). State dependent activity in monkey visual cortex: I. Single cell activity in V1 and V4 on visual tasks. *Experimental Brain Research*, **69**, 225-244.
- HEATHCOTE, A., & MEWHORT, D. J. K. (1993). Representation and selection of relative position. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 488-516.
- HELMHOLTZ, H. VON (1962). *Treatise on physiological optics* (Vol. 3; J. P. C. Southall, Trans.). New York: Dover. (Original work published 1866)
- 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.
- HOLLIDAY, I. E., & BRADDICK, O. J. (1991). Pre-attentive detection of a target defined by stereoscopic slant. *Perception*, **20**, 355-362.
- HUBEL, D. H., & LIVINGSTONE, M. S. (1987). Segregation of form,

- color, and stereopsis in primate area 18. *Journal of Neuroscience*, 7, 3378-3415.
- HUMMEL, J. E., & BIEDERMAN, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, 99, 480-517.
- HUMPHREYS, G. W., & MÜLLER, H. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology*, 25, 43-110.
- HUMPHREYS, G. W., QUINLAN, P. T., & RIDDOCH, M. J. (1989). Grouping processes in visual search: Effects with single and combined-feature targets. *Journal of Experimental Psychology: General*, 118, 258-279.
- INTRAUB, H. (1985). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology: Human Perception & Performance*, 11, 431-442.
- IVRY, R. B., & COHEN, A. (1990). Dissociation of short- and long-range apparent motion in visual search. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 317-331.
- JOHNSTON, W. A., HAWLEY, K. J., & FARNHAM, J. M. (1993). Novel popout: Empirical boundaries and tentative theory. *Journal of Experimental Psychology: Human Perception & Performance*, 19, 140-153.
- JONIDES, J., & GLEITMAN, H. (1972). A conceptual category effect in visual search: O as letter or digit. *Perception & Psychophysics*, 12, 457-460.
- JONIDES, J., & YANTIS, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics*, 43, 346-354.
- JULESZ, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neuroscience*, 7, 41-45.
- JULESZ, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics*, 54, 245-251.
- JULESZ, B., & BERGEN, J. R. (1983). Textons, the fundamental elements in preattentive vision and perceptions of textures. *Bell Systems Technical Journal*, 62, 1619-1646.
- JULESZ, B., & KRÖSE, B. (1988). Features and spatial filters. *Nature*, 333, 302-303.
- KAHNEMAN, D., & TREISMAN, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29-61). Orlando, FL: Academic Press.
- KANWISHER, N. (1991). Repetition blindness and illusory conjunctions: Errors in binding visual types with visual tokens. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 30-47.
- KANWISHER, N., & DRIVER, J. (1992). Objects, attributes, and visual attention: Which, what, and where. *Current Directions in Psychological Science*, 1, 26-31.
- KANWISHER, N., & POTTER, M. C. (1989). Repetition blindness: The effects of stimulus modality and spatial displacement. *Memory & Cognition*, 17, 117-124.
- KANWISHER, N., & POTTER, M. C. (1990). Repetition blindness: Levels of processing. *Journal of Experimental Psychology: Human Perception & Performance*, 16, 30-47.
- KARNI, A., & SAGI, D. (1990). Texture discrimination learning is specific for spatial location and background element orientation. *Investigative Ophthalmology & Visual Science*, 31, 562. (Abstract)
- KARNI, A., & SAGI, D. (1992). Later but (almost) forever—The time course of learning of a visual skill. *Investigative Ophthalmology & Visual Science*, 33, 1356. (Abstract)
- KEEBLE, D. R., & MORGAN, M. J. (1993). A linear systems approach to texture perception. *Investigative Ophthalmology & Visual Science*, 34, 1237. (Abstract)
- KELLY, P. L., HARRISON, D. W., & HODGE, M. H. (1991). The category effect in visual selective attention. *Bulletin of the Psychonomic Society*, 29, 71-74.
- KINCHLA, R. A. (1977). The role of structural redundancy in the perception of visual targets. *Perception & Psychophysics*, 22, 19-30.
- KINCHLA, R. A. (1992). Attention. *Annual Review of Psychology*, 43, 711-742.
- KINCHLA, R. A., & WOLFE, J. M. (1979). The order of visual processing: "Top-down," "bottom-up," or "middle-out." *Perception & Psychophysics*, 25, 225-231.
- KLEIN, R. (1988). Inhibitory tagging system facilitates visual search. *Nature*, 334, 430-431.
- KNIERIM, J. J., & VAN ESSEN, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert monkey. *Journal of Neurophysiology*, 67, 961-980.
- KRÖSE, B. A. J., & JULESZ, B. (1989). The control and speed of shifts of attention. *Vision Research*, 29, 1607-1619.
- KWAK, H.-W., DAGENBACH, D., & EGETH, H. [E.] (1991). Further evidence for a time-independent shift of the focus of attention. *Perception & Psychophysics*, 49, 473-480.
- LANDY, M. S., & BERGEN, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, 31, 679-691.
- LENNIE, P., TREVARTHEN, C., VAN ESSEN, D., & WASSLE, H. (1990). Parallel processing of visual information. In L. Spillman & J. S. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 103-128). San Diego, CA: Academic Press.
- LOGOTHETIS, N. K., & CHARLES, E. R. (1990). V4 responses to gratings defined by random dot motion. *Investigative Ophthalmology & Visual Science*, 31, 90. (Abstract)
- LUSCHOW, A., & NOTHDURFT, H. C. (1993). Pop-out of orientation but not pop-out of motion at isoluminance. *Vision Research*, 33, 91-104.
- MACH, E. (1959). *The analysis of sensations* (C. M. Williams, Trans.). New York: Dover. (Original work published 1886)
- MACKEBEN, M., & NAKAYAMA, K. (1988). Fixation release facilitates rapid attentional shifts. *Investigative Ophthalmology & Visual Science*, 29, 22. (Abstract)
- MALIK, J., & PERONA, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, 7, 923-932.
- MALJKOVIC, V., & NAKAYAMA, K. (1992). Color and position priming in the deployment of attention. *Investigative Ophthalmology & Visual Science*, 33, 1354. (Abstract)
- MARENDAZ, C., STIVALET, P., BARRACLOUGH, L., & WALKOWIAC, P. (1993). Effect of gravito-inertial cues on visual search for orientation. *Journal of Experimental Psychology: Human Perception & Performance*, 19, 1266-1277.
- MAUNSELL, J. H. R., SCLAR, G., & NEALEY, T. A. (1988). Task-specific signals in area V4 of monkey visual cortex. *Society for Neuroscience Abstracts*, 14, 10.
- MCLEOD, P., DRIVER, J., & CRISP, J. (1988). Visual search for conjunctions of movement and form is parallel. *Nature*, 332, 154-155.
- MCLEOD, P., DRIVER, J., DIENES, Z., & CRISP, J. (1991). Filtering by movement in visual search. *Journal of Experimental Psychology: Human Perception & Performance*, 17, 55-64.
- MORAGLIA, G. (1989). Display organization and the detection of horizontal line segments. *Perception & Psychophysics*, 45, 265-272.
- MORAN, J., & DESIMONE, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782-784.
- MORDKOFF, J. T., & YANTIS, S. (1993). Dividing attention between color and shape: Evidence of coactivation. *Perception & Psychophysics*, 53, 357-366.
- MOSER, M. C. (1991). *The perception of multiple objects*. Cambridge, MA: MIT Press, Bradford Books.
- NAGY, A. L., & SANCHEZ, R. R. (1990). Critical color differences determined with a visual search task. *Journal of the Optical Society of America A*, 7, 1209-1217.
- NAGY, A. L., SANCHEZ, R. R., & HUGHES, T. C. (1990). Visual search for color differences with foveal and peripheral vision. *Journal of the Optical Society of America A*, 7, 1995-2001.
- NAKAYAMA, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 411-422). Cambridge: Cambridge University Press.
- NAKAYAMA, K., & SILVERMAN, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature*, 320, 264-265.
- NAVON, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9, 353-383.
- NEISSER, U. (1967). *Cognitive psychology*. New York: Appleton, Century, Crofts.
- NOTHDURFT, H. C. (1990). Texture discrimination by cells in the cat lateral geniculate nucleus. *Experimental Brain Research*, 82, 48-66.

- NOTHDURFT, H. C. (1991a). Different effects from spatial frequency masking in texture segregation and texture detection tasks. *Vision Research*, **31**, 299-320.
- NOTHDURFT, H. C. (1991b). Texture segmentation and pop-out from orientation contrast. *Vision Research*, **31**, 1073-1078.
- NOTHDURFT, H. C. (1993a). Matching luminance pop-out with that of orientation or motion. *Investigative Ophthalmology & Visual Science*, **34**, 1054. (Abstract)
- NOTHDURFT, H. C. (1993b). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research*, **33**, 1937-1958.
- NOTHDURFT, H. C., & LI, C. Y. (1984). Representation of spatial details in textured patterns by cells of the cat striate cortex. *Experimental Brain Research*, **57**, 9-21.
- O'CONNELL, K. M., & TREISMAN, A. (1992). *Shared orientation coding for lines, dot pairs, and edges*. Manuscript submitted for publication.
- O'NEILL, P., WOLFE, J. M., & BILSKY, A. B. (1993). Individual differences in visual search. *Investigative Ophthalmology & Visual Science*, **34**, 1236. (Abstract)
- O'TOOLE, A. J., & WALKER, C. L. (1993). Disparity as a visual primitive: The competing role of surface percepts. *Investigative Ophthalmology & Visual Science*, **34**, 1187. (Abstract)
- PASHLER, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics*, **41**, 191-201.
- PASHLER, H. (1988). Cross-dimensional interaction and texture segregation. *Perception & Psychophysics*, **43**, 307-318.
- PAVEL, M., ECONOPOULY, J., & LANDY, M. S. (1992). The psychophysics of rapid visual search. *Investigative Ophthalmology & Visual Science*, **33**, 1355. (Abstract)
- PETERHANS, E., VON DER HEYDT, R., & BAUMGARTNER, G. (1986). Neuronal responses to illusory contour stimuli reveal stages of visual cortical processing. In J. D. Pettigrew, K. J. Sanderson, & W. R. Levick (Eds.), *Visual neuroscience* (pp. 343-351). Cambridge: Cambridge University Press.
- POISSON, M. E., & WILKINSON, F. (1992). Distractor ratio and grouping processes in visual conjunction search. *Perception*, **21**, 21-38.
- POMERANTZ, J. R., & PRISTACH, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 635-649.
- POSNER, M. I., & COHEN, Y. (1984). Components of attention. In H. Bouma & D. G. Bowhuis (Eds.), *Attention and performance X* (pp. 55-66). Hillsdale, NJ: Erlbaum.
- PRINZMETAL, W., & KEYSAR, B. (1989). Functional theory of illusory conjunctions and neon colors. *Journal of Experimental Psychology: General*, **118**, 165-190.
- 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.
- RAFAL, R. D., CALABRESI, P. A., BRENNAN, C. W., & SCIOLTO, T. K. (1989). Saccade preparation to recently attended locations. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 673-685.
- RAMACHANDRAN, V. S. (1988). Perception of shape from shading. *Nature*, **331**, 163-165.
- RATCLIFF, R. (1978). A theory of memory retrieval. *Psychological Review*, **85**, 59-108.
- RENSINK, R., & CAVANAGH, P. (1993). Processing of shadows at preattentive levels. *Investigative Ophthalmology & Visual Science*, **34**, 1288. (Abstract)
- ROBERTSON, L. C., EGLY, R., LAMB, M. R., & KERTH, L. (1993). Spatial attention and cuing to global and local levels of hierarchical structure. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 471-487.
- ROBERTSON, L. C., & LAMB, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology*, **23**, 299-330.
- ROCK, I. (1974, January). The perception of disoriented figures. *Scientific American*, pp. 78-85.
- ROSS, W. D., GROSSBERG, S., & MINGOLLA, E. (1993). A neural model of visual search. *Investigative Ophthalmology & Visual Science*, **34**, 1235. (Abstract)
- RUBENSTEIN, B. S., & SAGI, D. (1990). Spatial variability as a limiting factor in texture discrimination tasks: Implications for performance asymmetries. *Journal of the Optical Society of America A*, **7**, 1632-1643.
- RUBIN, J. M., & KANWISHER, N. (1985). Topological perception: Holes in an experiment. *Perception & Psychophysics*, **37**, 179-180.
- SAGI, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. *Perception & Psychophysics*, **43**, 601-603.
- SAGI, D. (1990). Detection of an orientation singularity in Gabor textures: Effect of signal density and spatial-frequency. *Vision Research*, **30**, 1377-1388.
- SAGI, D., & JULESZ, B. (1985a). Fast noninertial shifts of attention. *Spatial Vision*, **1**, 141-149.
- SAGI, D., & JULESZ, B. (1985b). "Where" and "what" in vision. *Science*, **228**, 1217-1219.
- SANDON, P. A., & YANIKOGLU, B. A. (1990). Visual search as constraint propagation. In *Proceedings of the 12th Annual Conference of the Cognitive Science Society* (pp. 574-581). Hillsdale, NJ: Erlbaum.
- SCHNEIDER, W., & SHIFFRIN, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review*, **84**, 1-66.
- SEKULER, R., & BALL, K. K. (1986). Visual localization: Age and practice. *Journal of the Optical Society of America A*, **3**, 864-868.
- SMALLMAN, H. S., & BOYNTON, R. M. (1990). Segregation of basic color in an information display. *Journal of the Optical Society of America A*, **7**, 1985-1994.
- SMITH, S. L. (1962). Color coding and visual search. *Journal of Experimental Psychology*, **64**, 434-440.
- SPITZER, H., DESIMONE, R., & MORAN, J. (1988). Increased attention enhances both behavioral and neuronal performance. *Science*, **240**, 338-340.
- STERNBERG, S. (1969). High-speed scanning in human memory. *Science*, **153**, 652-654.
- STONE, W., & MACK, A. (1993). Perception of grouping based on motion without attention. *Investigative Ophthalmology & Visual Science*, **34**, 1233. (Abstract)
- TADMOR, Y., & TOLHURST, D. J. (1993). Both the phase and the amplitude spectrum may determine the appearance of natural images. *Vision Research*, **33**, 141-145.
- TAYLOR, S., & BADCOCK, D. (1988). Processing feature density in preattentive perception. *Perception & Psychophysics*, **44**, 551-562.
- THEEUWES, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics*, **50**, 184-193.
- THEEUWES, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics*, **51**, 599-606.
- THOMAS, J. P., & GILLE, J. (1979). Bandwidths of orientation channels in human vision. *Journal of the Optical Society of America*, **69**, 652-660.
- TIPPER, S. P., & BAYLIS, S. C. (1987). Individual difference in selective attention: The relation of priming and interference to cognitive failure. *Personality & Individual Differences*, **8**, 667-675.
- TIPPER, S. P., BREHAUT, J. C., & DRIVER, J. (1990). Selection of moving and static objects for the control of spatially directed attention. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 492-504.
- TIPPER, S. P., DRIVER, J., & WEAVER, B. (1991). Object centered inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology*, **43A**, 289-298.
- TOWNSEND, J. T. (1971). A note on the identifiability of parallel and serial processes. *Perception & Psychophysics*, **10**, 161-163.
- TOWNSEND, J. T. (1976). Serial and within-stage independent parallel

- model equivalence on the minimum completion time. *Journal of Mathematical Psychology*, **14**, 219-239.
- 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.
- TREISMAN, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, & Image Processing*, **31**, 156-177.
- TREISMAN, A. (1986a, May). Features and objects in visual processing. *Scientific American*, pp. 114B-125.
- TREISMAN, A. (1986b). Properties, parts, and objects. In K. R. Boff, L. Kaufmann, & J. P. Thomas (Eds.), *Handbook of human perception and performance* (pp. 35.1-35.70). New York: Wiley.
- TREISMAN, A. (1988). Features and objects: The 14th Bartlett Memorial Lecture. *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 & Performance*, **16**, 459-478.
- TREISMAN, A., & SCHMIDT, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, **14**, 107-141.
- TREISMAN, A., & SOUTHER, J. (1986). Illusory words: The roles of attention and of top-down constraints in conjoining letters to form words. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 3-17.
- TREISMAN, A., VIEIRA, A., & HAYES, A. (1992). Automaticity and preattentive processing. *American Journal of Psychology*, **105**, 341-362.
- TSOTSOS, J. K. (1990). Analyzing vision at the complexity level. *Behavioral & Brain Sciences*, **13**, 423-469.
- VAN ESSEN, D. C., & MAUNSELL, J. H. R. (1983). Hierarchical organization and functional streams in the visual cortex. *Trends in Neuroscience*, **6**, 370-375.
- VIEIRA, A., & TREISMAN, A. (1988, November). *Automatic search: Changing perceptions or procedures?* Paper presented at the meeting of the Psychonomic Society, Chicago.
- VIRZI, R. A., & EGETH, H. E. (1984). Is meaning implicated in illusory contours? *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 573-580.
- VON DER HEYDT, R., & DURSTELER, M. R. (1993). Visual search: Monkeys detect conjunctions as fast as features. *Investigative Ophthalmology & Visual Science*, **34**, 1288. (Abstract)
- VON DER HEYDT, R., PETERHANS, E., & BAUMGARTNER, G. (1984). Illusory contours and cortical neuron responses. *Science*, **224**, 1260-1262.
- VOORHEES, H., & POGGIO, T. (1988). Computing texture boundaries from images. *Nature*, **333**, 364-367.
- WANG, Q., & CAVANAGH, P. (1993). Acquired familiarity effects in visual search with Chinese characters. *Investigative Ophthalmology & Visual Science*, **34**, 1236. (Abstract)
- WANG, Q., CAVANAGH, P., & GREEN, M. (1992). Familiarity and pop-out in visual search. *Investigative Ophthalmology & Visual Science*, **33**, 1262. (Abstract)
- WARD, R., & MCCLELLAND, J. L. (1989). Conjunctive search for one and two identical targets. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 664-672.
- WHITE, J. M., LEVI, D. M., & AITSEBAOMO, A. P. (1992). Spatial localization without visual references. *Vision Research*, **32**, 513-526.
- WILSON, H. R. (1986). Responses of spatial mechanisms can explain hyperacuity. *Vision Research*, **26**, 453-469.
- WILSON, H. R., LEVI, D., MAFFEI, L., ROVAMO, J., & DEVALOIS, R. (1990). The perception of form: Retina to striate cortex. In L. Spillman & J. S. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 231-272). San Diego, CA: Academic Press.
- WOLFE, J. M. (1986). Stereopsis and binocular rivalry. *Psychological Review*, **93**, 269-282.
- WOLFE, J. M. (1992). "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research*, **32**, 757-763.
- WOLFE, J. M. (1994). Visual search in continuous, naturalistic stimuli. *Vision Research*, **34**, 1187-1195.
- WOLFE, J. M., & BOSE, M. (1991). *Visual search for the "medium" stimulus*. Unpublished research report. Cambridge, MA: Massachusetts Institute of Technology.
- WOLFE, J. M., & CAVE, K. R. (1989). Deploying visual attention: The guided search model. In T. Troscianko & A. Blake (Eds.), *AI and the eye* (pp. 79-103). Chichester, U.K.: Wiley.
- 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 & Performance*, **15**, 419-433.
- WOLFE, J. M., CHUN, M. M., & FRIEDMAN-HILL, S. R. (1993). Making use of texture gradients: Visual search and texture grouping exploit the same parallel processes in different ways. *Spatial Vision*, **7**, 90.
- WOLFE, J. M., CHUN, M. M., & FRIEDMAN-HILL, S. R. (in press). Making use of texture gradients: Visual search and perceptual grouping exploit the same parallel processes in different ways. In T. Papathomas & A. Gorea (Eds.), *Linking psychophysics, neurophysiology, and computational vision: A volume in honor of Bela Julesz*.
- WOLFE, J. M., & FRANZEL, S. L. (1988). Binocularity and visual search. *Perception & Psychophysics*, **44**, 81-93.
- WOLFE, J. M., & FRIEDMAN-HILL, S. R. (1990, November). *Must we attend to every abrupt onset?* Paper presented at the meeting of the Psychonomic Society, New Orleans.
- WOLFE, J. M., & FRIEDMAN-HILL, S. R. (1992a). On the role of symmetry in visual search. *Psychological Science*, **3**, 194-198.
- WOLFE, J. M., & FRIEDMAN-HILL, S. R. (1992b). Visual search for orientation: The role of angular relations between targets and distractors. *Spatial Vision*, **6**, 199-208.
- WOLFE, J. M., FRIEDMAN-HILL, S. R., & BILSKY, A. B. (1994). Parallel processing of part-whole information in visual search tasks. *Perception & Psychophysics*, **55**, 537-550.
- 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 & Performance*, **18**, 34-49.
- WOLFE, J. M., & POKORNY, C. W. (1990). Inhibitory tagging in visual search: A failure to replicate. *Perception & Psychophysics*, **48**, 357-362.
- WOLFE, J. M., YEE, A., & FRIEDMAN-HILL, S. R. (1992). Curvature is a basic feature for visual search. *Perception*, **21**, 465-480.
- WOLFE, J. M., YU, K. P., PRUSZENSKI, A. D., & CAVE, K. R. (1988, November). *A preattentive feature process can execute only one command at a time*. Paper presented at the meeting of the Psychonomic Society, Chicago.
- 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 & Performance*, **16**, 879-892.
- YANTIS, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, **2**, 156-161.
- YANTIS, S., & JOHNSON, D. N. (1990). Mechanisms of attentional priority. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 812-825.
- YANTIS, S., & JONES, E. (1991). Mechanisms of attentional selection: Temporally modulated priority tags. *Perception & Psychophysics*, **50**, 166-178.
- YANTIS, S., & JONIDES, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception & Performance*, **16**, 121-134.
- YUND, E. W., EFRON, R., & NICHOLS, D. R. (1990). Detectability gradients as a function of target location. *Brain & Cognition*, **12**, 1-16.
- ZEKI, S. M. (1978). Uniformity and diversity of structure and function in Rhesus monkey prestriate visual cortex. *Journal of Physiology*, **277**, 273-290.
- ZEKI, S. M., WATSON, J. D. G., LUECK, C. J., FRISTON, K. J., KENNARD, C., & FRACKOWIAK, R. S. J. (1991). A direct demonstration

of functional specialization in human visual cortex. *Journal of Neuroscience*, **11**, 641-649.

ZHOU, W., CHEN, L., & ZHANG, X. (1992). Topological perception: Holes in illusory conjunction and visual search. *Investigative Ophthalmology & Visual Science*, **33**, 958. (Abstract)

ZHOU, W., ZHANG, X., & CHEN, L. (1993). Shape transformation in illusory conjunctions. *Investigative Ophthalmology & Visual Science*, **34**, 1082. (Abstract)

ZOHARY, E., & HOCHSTEIN, S. (1989). How serial is serial processing in vision? *Perception*, **18**, 191-200.

NOTES

1. Three points are worth making here. First, the claim that a limited capacity process must be deployed in a serial manner across the visual field is *not* a claim that the process is itself serial. For example, face recognition may well be a parallel process in the sense that eyes, nose, mouth, etc. are all considered at once. It is, however, a limited-capacity parallel process that must be deployed from face to face. One cannot process a room full of faces in parallel. Second, turning to visual search, the general notion of a spatially restricted deployment of attention does not require that the limited-capacity stage be limited to processing of one item at a time (though that assumption is made in the current version of Guided Search). We will return to this topic later. Third, the sequence of a parallel stage followed by a serial stage should not be taken as a one-way relationship between two independent processing stages. While it is clear enough that the initial (e.g., retinal) stages of processing are parallel, there is no sharp border between parallel and serial stages. As we will see, the two types of processes work together at the same time to solve visual problems (Friedman-Hill & Wolfe, in press).

2. One can spend a lot of time trying to determine whether the second stage of processing is a truly serial, item-by-item affair, or whether it is a limited-capacity parallel process. Alternatives include an ability to process more than one item at once (Cave & Pashler, 1994a, 1994b; Mordkoff & Yantis, 1993; Nakayama, 1990; Pashler, 1987) or even an ability to process all items at once with speed or depth of processing dependent on the allocation of a limited resource (Kinchla, 1977, 1992). The current version of Guided Search makes a choice. Items are processed one at a time. However, the basic issues in Guided Search and the underlying strategy would remain similar even if it could be shown that the limited-capacity stage still processes all items at once. The critical issue is how to deploy limited resources in an intelligent fashion. Currently, resources are deployed from item to item in decreasing order of the likelihood that they contain a target. Much the same result could be obtained by deploying a limited resource to all items with the *amount* of resource deployed at each locus dependent on the likelihood that the locus contains a target.

3. There are, of course, separate parallel pathways carrying information in the early visual pathways. The most important of these are the M and P (transient and sustained, color-opponent and broad-band) pathways and the ON and OFF pathways (Lennie, Trevarthen, Van Essen, & Wässle, 1990). However, as will be discussed later, these distinctions do not appear to have much impact on visual search, probably because visual search operates over representations generated later in visual processing.

4. The activation map is a modeling convenience, not a physiological hypothesis. There is no need to bring all activation together in a single locus. What is required is that the movement of attention be based on a weighted combination of the feature-specific activations.

5. It is probably a mistake to think of attention "moving" from location to location in some strict analogy with eye movements or mental imagery. More plausibly, it is deployed at location *x*, then disengaged from *x*, and redeployed at *y* without necessarily traversing intermediate points. The literature on this topic is beyond the scope of this paper. If one prefers a parallel, limited-capacity model of the deployment of attention, the role of the activation map changes. In a parallel model, the activation map would guide the apportionment of attention with more of the limited resource devoted to high-activation locations and less to lower activation locations.

6. There has been considerable discussion about whether attention is directed to locations or to objects. As the issue is framed in the literature, the answer is probably "objects" (Kahneman & Treisman, 1984; Tipper, Brehaut, & Driver, 1990; Tipper et al., 1991; Treisman, 1986a, 1988), but for the current version of Guided Search, the issue is not critical. In an array of items, activation will be higher at the items than in intervening blank space. Objects and locations will be, in effect, the same thing.

7. If there happened to be a tie in the activations of two items, the order of checking would be random.

8. Heathcote and Mewhort (1993) would disagree with this last point. They do not believe that their data can be explained by phase-sensitive mechanisms. In particular, they note that red-blue targets can be found amid blue-red distractors and that no cells have been found in the cortex that respond optimally to that polarity of edge. This is not a real objection to a phase account, since the red-green cell that they invoke would certainly respond differentially to red-blue versus blue-red edges.

9. We are currently revamping the simulation to run on a Sun (or equivalent) workstation and would be happy to share the code (in C) with interested readers who would care to try it on their favorite search paradigm.

(Manuscript received October 21, 1993;
revision accepted for publication January 29, 1994.)