8

Extending Guided Search: Why Guided Search Needs a Preattentive "Item Map"

Jeremy M. Wolfe

Guided Search is a model that seeks to explain how humans find one visual stimulus in a world filled with other, distracting stimuli. Complex stimuli like faces and words can only be identified one at a time. Attention must be deployed to an item before it can be fully identified. The heart of Guided Search is the idea that relatively simple processes can be used to guide attention intelligently so that, for example, in a search for a round skin-colored face one does not waste time examining an elongated green tree. Attention could be guided from location to location or from item to item in the visual field. The experiments discussed in this chapter argue for item-by-item guidance.

There is wide agreement on the existence of bottlenecks in human information processing (Kahneman, 1973; Neisser, 1967).[1] Humans cannot handle all of the demands placed on them at the same time. This is not a new observation. For example, in Shakespeare's *As You Like It,* Celia meets Orlando in the woods. When Rosalind, who loves Orlando, asks Celia for

---

I thank Gordon Logan, Lex Van der Heijden, Claus Bundesen, Patricia O'Neill, Greg Gancarz, and Sara Bennett for comments on a draft of this chapter and Alex Bilsky, Stacia Friedman-Hill, and Charles Pokorny for help collecting data. This work was supported by grants from the National Institutes of Health, National Eye Institute (EY05087), and the Air Force Office of Scientific Research (93NL105).

[1] Wide, but not universal, agreement (see Van der Heijden, chap. 17, this volume).

information about this meeting, she gives instructions for response comparable to the most arduous laboratory attention tasks:

> What did he when thou saw'st him? What said he? How looked he? Wherein went he? What makes he here? Where remains he? How parted he with thee? And when shalt thou see him again? Answer me in one word. (*As You Like It* 3.2.218–222)

Celia, a paleocognitive scientist, recognizes a bottleneck and responds

> you must borrow me Gargantua's mouth first; "tis a word too great for any mouth of this age's size." (*As You Like It* 3.2.223–224)

The play, having other business at hand, does not pursue the locus of this bottleneck. Celia seems to propose a motor or final common path limit. But the task could be limited at the level of response selection (Pashler, 1992, 1994) or memory scanning (Sternberg, 1969). Bottlenecks limit the flow of information at multiple loci in processing. The study of visual processing has provided one of the clearest examples of a processing bottleneck. There is fairly wide agreement that the initial stages of visual processing are spatially parallel, with all loci in the visual field processed at once and with no cost at Locus 1 for concurrent processing at Locus 2. At some point in the journey from image to perception and action, however, there is a bottleneck. Attention selects some stimuli for more processing than others. One can debate whether this attentional selection is early (e.g., Mack, Tang, Tuma, & Kahn, 1992) or late (e.g., Johnston, Hawley, & Farnham, 1993). One can also ask whether the selection implies serial processing of single locations or items (Treisman & Gelade, 1980), or whether selection implies differential allocation or processing resources to multiple loci in parallel (e.g., Kinchla, 1977). There seems little point, however, in arguing about the existence of some attentional selection.

The constriction in processing is quite severe. Consider a search for an *S* among mirror-reversed *S*s as shown in Figure 1. This is most likely a serial self-terminating search (Kwak, Dagenbach, & Egeth, 1991). Whether it is serial or limited capacity parallel, the search proceeds at a rate equivalent to the serial processing of one item every 40–60 ms. Looking out at any visually rich scene, it is clear that this is a potentially crippling limit. At 25 items per second, it could take many seconds to process a complex scene when the demands for action might require a much shorter response time.

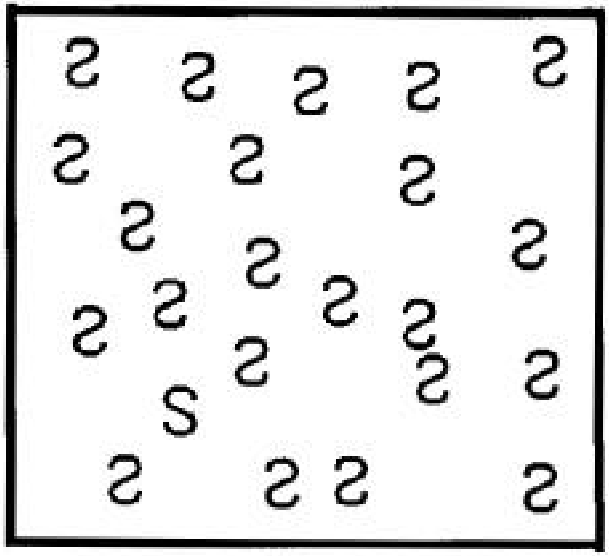


Figure 1

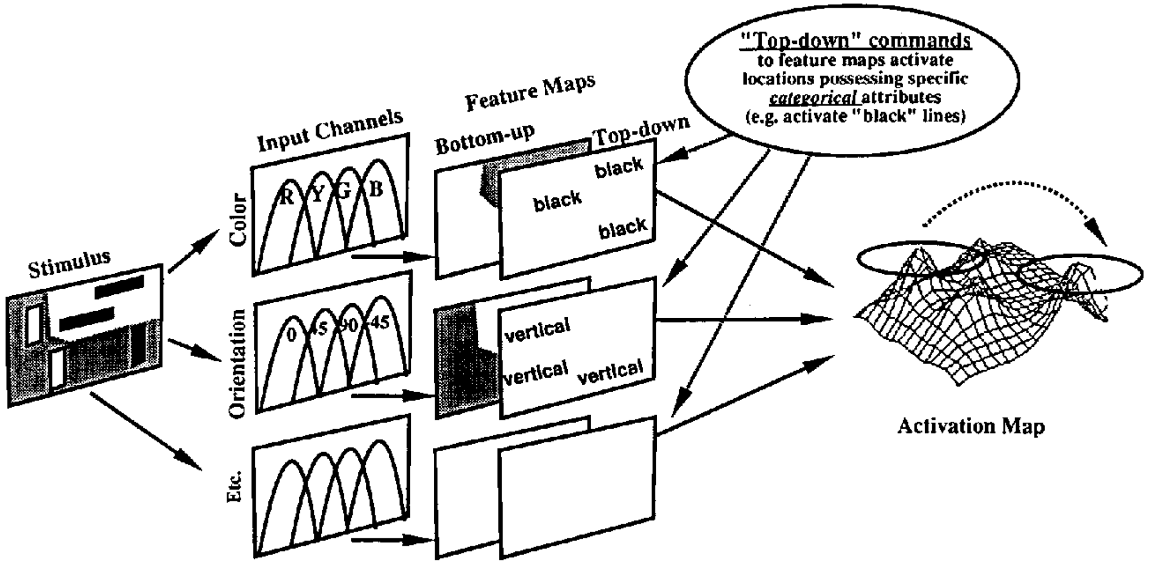
Search for an *S* among mirror-reversed *S*s is serial and self-terminating.

Obviously, the search in Figure 1 is an unusual case. Most searches, in the world or in the laboratory, do not require serial examination of all items. Some items and loci are removed from consideration prior to the bottleneck. Put another way, the parallel processes lying before the bottleneck guide the deployment of attention. The rules of that guidance are the subject of the Guided Search model developed in our laboratory (Cave & Wolfe, 1990; Wolfe, 1992b, 1994a; Wolfe, Cave, & Franzel, 1989). This chapter briefly reviews Guided Search 2.0 (GS2), the most recently published version of the model, and discusses data that illustrate some of the shortcomings of the model and one direction for its future development.

GUIDED SEARCH 2.0

Basic Features

A full account of GS2 is found in Wolfe (1994a). Figure 2 shows the basic architecture of the model. The input stimulus is filtered through broadly tuned channels for a limited set of basic features. This occurs in parallel across the visual field. There are perhaps a dozen basic features



The Stimulus is filtered through broadly-tuned "categorical" channels.

The output produces feature maps with activation based on local differences (bottom-up) and task demands (top-down).

A weighted sum of these activations forms the Activation Map. In visual search, attention deploys limited capacity resources in order of decreasing activation.

Figure 2

The architecture for Guided Search 2.0. From "Guided Search 2.0: A Revised Model of Visual Search," by J. M. Wolfe, 1994, *Psychonomic Bulletin and Review, 1,* p. 205. Copyright 1994 by the Psychonomic Society. Reprinted with permission.

(see Wolfe, 1994a, for a review), including such obvious candidates as color, orientation, and motion and some less obvious properties such as "lustre" (Wolfe & Franzel, 1988) and pictorial depth cues (Enns & Rensink, 1990). These less basic "basic features" make an important point. Parallel processing of visual information extends beyond early vision to a level that handles more complex properties of surfaces (e.g., He & Nakayama, 1992, 1994) and their relationship to one another (e.g., occlusion; Enns & Rensink, 1992). Further evidence that visual search uses relatively "late" parallel representations comes from experiments that show that "basic features" in visual search can be built up of primitives from other feature spaces (e.g., oriented items defined by color, texture, motion, and so on; Bravo & Blake, 1990; Cavanagh, Arguin, & Treisman, 1990).

Bottom-Up Activation

For purposes of the guidance of attention, there are two consequences of the parallel processing of basic features: *bottom-up*, stimulus-driven activation and *top-down*, user-driven activation. Bottom-up activation is based on local differences. A locus where color is changing will receive bottom-up activation in a color processor. An item of unique orientation will receive a large amount of bottom-up activation in an orientation processor because it will be different from all of its neighbors (see Figure 3a). The strength of bottom-up activation is a function of the magnitude of the difference between an item and its neighbors (see Figure 3b). The

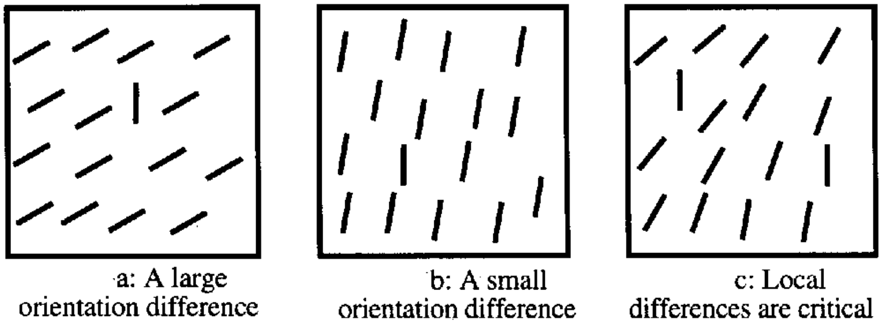

a: A large orientation difference
b: A small orientation difference
c: Local differences are critical

Figure 3

Illustrations of bottom-up activation.

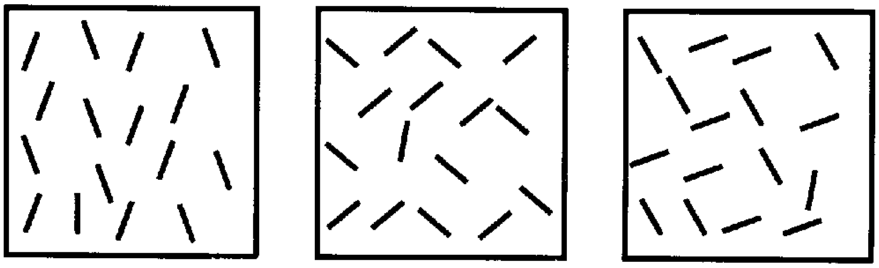
identity of more remote items is less important (see Figure 3c; Nothdurft, 1991, 1993a, 1993c). If the difference is large enough, attention will be guided to the target immediately on every trial and reaction time (RT) will be independent of set size. Bottom-up activation does not specify the source of the activation. A red item surrounded by green items produces a peak of activation that in principle could be identical to a peak produced by vertical surrounded by horizontal items.

The magnitude of the difference required to produce parallel search for target amidst homogeneous distractors can be called a "preattentive just noticeable difference" (pJND). Interestingly, the pJNDs are much larger than classical JNDs and are not related to those JNDs by a simple scale factor (Nagy & Sanchez, 1990). This local differencing aspect of parallel processing is what Julesz referred to as the calculation of "texton gradients" (Julesz, 1986). Different features appear to have different abilities to attract attention bottom-up, with abrupt onset and/or the creation of new objects being, perhaps, the most forceful (Jonides & Yantis, 1988; Yantis, 1993; Yantis & Jones, 1991).

Top-Down Activation

Guidance of attention that was entirely stimulus driven would not be particularly useful. It is important to be able to guide attention toward currently relevant features in the visual input. Otherwise, a search for a coin dropped on the sidewalk might be continuously disrupted by a more salient, flashing neon sign. It is obvious that people can bring search under volitional control. It is less obvious, but nevertheless true, that our ability to command parallel processing is limited in interesting ways. For instance, in Figure 4a, it is hard to search for the vertical item among distractors tilted 20° to either side. It is hard, not because subjects cannot discriminate 0° from ±20°, but because top-down control is limited to a very small set of apparently categorical items. In orientation, these appear to correspond roughly to the terms *steep*, *shallow*, *left*, *right*, and *tilted* (Wolfe, Friedman-Hill, Stewart, & O'Connell, 1992). It may be that the limit reflects an ability to select only a single, broadly tuned input channel in top-down processing (Foster & Ward, 1991).

To give an example, in Figure 4a, all of the items are steep. In Figure



a: It is hard to find "vertical" among other "steep" items.

b: It is quite easy to find "steep" among "shallow" items.

c: It is hard to find the "steepest" item.

Figure 4

Illustrations of top-down processing of orientation.

4b, search is easier because the target is the only steep item (10° among ±50°). In Figure 4c, search is hard again even though the angular distance from target to distractors is the same as in Figure 4b (10° among −30° and +70°). The search in Figure 4c is hard because the target is not categorically unique. It possesses steep and right properties but is presented among shallow right and steep left distractors.

Another important limitation on top-down processing is an inability to search efficiently for an item defined by a conjunction of two instances of a single type of feature. Thus, it is possible to search for a red or a green item among items of various colors, but search for a conjunctively defined item that is red *and* green among items that contain red *or* green, but not both, is very inefficient (Wolfe et al., 1990). This is different from the case of across-feature conjunctions (e.g., Color × Orientation—find the item that is red and vertical among items that are either red or vertical but not both.) Across-feature conjunctions can be very efficient (see Wolfe et al., 1989, and discussion below). This is not just a problem with color processing. Size × Size and Orientation × Orientation conjunctions are also very inefficient (Bilsky & Wolfe, 1994; and see Logan, 1994, for a class of difficult Shape × Shape conjunctions). There is one situation in which search for within-feature conjunctions is also efficient. This important exception is discussed below.

The Activation Map and the Mechanics of Search

In GS2, top-down and bottom-up activation build up in each feature processor independently. These independent activations are then combined into a general *activation map*. In visual search, attention is then deployed to the most active locus. If that does not contain a target, attention is redeployed to the next most active locus, until the target is found or the search is terminated. The activity in the activation map cannot be a simple sum (or average) or all the independent activations. It must be weighted to emphasize useful information and limit the impact of useless activation. For example, consider a search for a red line among orange lines of various orientations. There will be substantial bottom-up activity in the orientation processor that is irrelevant to the task. Efficient search, which is possible in such cases, requires that the orientation activation not mask the color activation. The weighting of color information would be high in this case and the weighting of orientation information would be low.

Pokorny and I used a different task to test the ability to ignore irrelevant activations. In this case, the activation came from abrupt onsets of the sort that Yantis and his colleagues have found can attract attention even when they are irrelevant to the task (Yantis & Jones, 1991; Yantis & Jonides, 1990). We asked if such onsets *must* attract attention. In the baseline condition, our subjects searched for a *T* among *Ls*. The *T* and *Ls* could be in any of four 90° rotations from upright. This is a standard serial search task yielding RT × Set Size slopes of about 20 ms/item on target trials and 40 ms/item on blank trials. In the abrupt-onset condition, white spots appeared on the screen in the regions between the *Ts* and *Ls* at a rate of one every 40 ms. If subjects had been unable to keep the activations due to these highly salient abrupt onsets out of the activation map that guides attention, search would have been massively impaired. As can be seen in Figure 5, there is a 50-ms increase in RT across set sizes, but the slopes remain the same in the two conditions, strongly suggesting that the abrupt onsets were not allowed to interfere with search. (Note that Yantis has also reported that abrupt onsets do not capture attention if they are known not to mark the target; Yantis & Jonides, 1990.)

In the current simulation of Guided Search, the weights that modulate the input from the feature processors to the activation map are set at

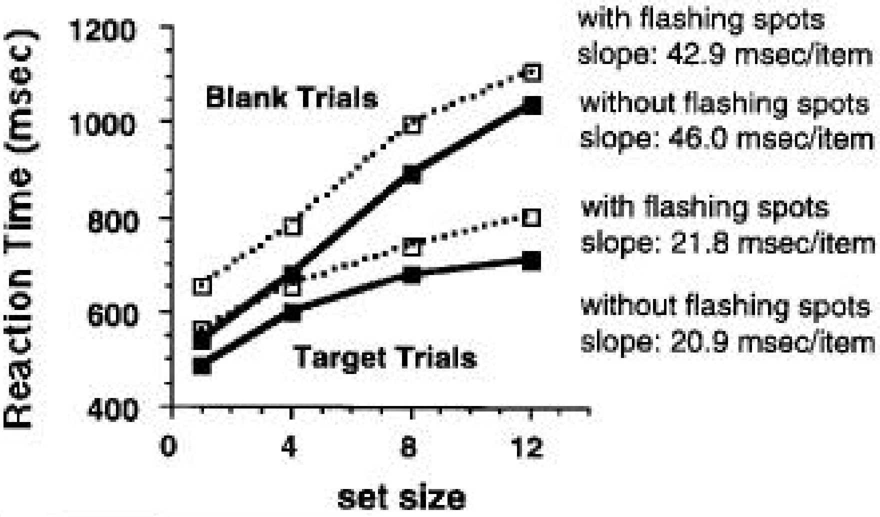


Figure 5

Results of searching for a *T* among *Ls* with and without flashing spots. Flashing spots, appearing at a rate of one every 40 ms, do not produce an increase in the slope of Reaction Time × Set Size functions for a serial search for a *T* among *Ls*.

the beginning of an "experiment" by instruction. That is, if the simulation is given a search for a red line target among orange lines of random orientation, it determines at the outset to largely ignore orientation information. This is not terribly realistic. It is more reasonable to imagine that the subject implicitly learns which preattentive processes contain signal and which, for all their activation, contain merely noise. Recently, Gancarz and I have given the simulation this learning ability. After each target-present trial, the simulation receives feedback that allows it to compare the activation of the target item with the average activation of other items. In this manner, it creates a signal-to-noise ratio for each parallel processor. The contribution of a processor to the overall activation map is increased if the ratio is greater than one and decreased otherwise. With appropriate limits to prevent the contributions from getting too large or too small, this simple learning mechanism produces a good approximation of the desired results. Processes that contain noise make little contribution. Processes that contain usable signals make stronger contributions.

This does assume feedback after each trial. It would be interesting to see if the performance of our subjects changed if we eliminated the feedback.

The architecture of feature processes feeding an activation map makes it possible to guide attention to likely loci for targets defined by conjunctions of two or more features even though no parallel process, by itself, is sensitive to conjunctive properties. Take a search for a red vertical line among 50% red horizontal and 50% green vertical distractors. Top-down activation in a color processor can activate all red items. Top-down processing in an orientation processor can activate all vertical (steep) items. The combination of these two sources of information will produce greater activation of an item that is both red and vertical. Indeed, if the system were noise free, this guidance of attention would be perfect and search for conjunctions should be no harder than search for salient single features. Search for conjunction can be highly efficient, with slopes near 0 ms/item (e.g., Wolfe, 1992a), but, in general, conjunction searches are somewhat less efficient than feature searches with slopes in the vicinity of 5–12 ms/item (Alkhateeb, Morland, Ruddock, & Savage, 1990; Cohen & Ivry, 1991; Dehaene, 1989; Egeth, Virzi, & Garbart, 1984; McLeod, Driver, Dienes, & Crisp, 1991; Quinlan & Humphreys, 1987; Treisman & Sato, 1990; Wolfe et al., 1989). Apparently, there is noise that will produce peaks of activation in the activation map that do not correspond to a target location. Search for a conjunction therefore becomes a serial search through a subset of the items. When differences between targets and distractors are large, that subset is quite small. The signal is almost always larger than the noise. As a result, slopes of RT × Set Size functions are near zero. In more demanding conjunction tasks, attention is incorrectly deployed to a greater number of distractors. The number of incorrect deployments, averaged across many trials, is proportional to set size, and so the RT × Set Size functions rise more steeply, approaching the slopes of serial search in the limit.

In Guided Search, parallel, serial, and guided searches are not qualitatively different. They lie on a continuum defined by the signal-to-noise ratio in the activation map. Thus, a *parallel* search is a search where the activation signal is so much larger than the background noise that attention is deployed first to the target location on all target-present trials. The subset in which attention is deployed contains only a single item: the target. The resulting RT × Set Size function is independent of set size. A fully serial search is a search with no parallel guidance. If preattentive processes cannot differentiate between targets and distractors (e.g., the searches in Figure 1 or Figure 5), search will be random over the entire set of items. If serial examination of items proceeds at a rate of one item every 40–60 ms, this will produce slopes of 20–30 ms/item on target trials and 40–60 ms/item on blank trials. In a guided search, the activation signal from the preattentive processes biases deployment of attention toward the target item but some distractor locations develop comparable levels of activation. The result is a serial search through a subset of the items.

Search Termination

Serial self-terminating searches produce 2:1 slope ratios. In a retrospective analysis of a large body of search data, Chun and I found that the average slope ratio between blank and target trials was about 2:1 across a wide range of tasks including conjunction searches (Chun & Wolfe, in press). Understanding search termination is easy enough in true serial searches. Target trial searches end when the target is found and blank trials end when all items have been examined and rejected. Serial search through a subset is more problematic. What should happen in guided searches (or, for that matter, parallel searches)? When should the subject abandon an unsuccessful search? Within the framework of the Guided Search model, there are at least two plausible mechanisms for quitting. As attention is deployed from peak to peak in the activation map, subjects could quit when none of the remaining peaks are high enough to warrant examination. That is, if the subject learns (implicitly, it is assumed) that virtually no target in the present task ever has an activation below some value, then search can be abandoned when no peaks remain above that value. Of course, the few targets with these low activations will not be found, generating miss errors. Alternatively, subjects could give up and guess when a search has gone on too long. That is, if the subject knows (again, implicitly) that almost any target should have been found within *N* ms, then search can be terminated if it goes on longer than *N* ms. If the probability of guessing increases with time, and most of the guesses are *nos*, the result will be error rates that increase somewhat with set size and a few false-alarm errors. There is evidence for activation and timing thresholds in search termination (Chun & Wolfe, in press). These timing and activation thresholds are fixed but change with feedback, becoming more conservative when errors are made and more liberal if correct responses are made. The GS2 simulation (Chun & Wolfe, in press; Wolfe, 1994a) incorporates both rules and does a reasonable job of reproducing the error rates, blank trial RTs, and 2:1 slope ratios seen in the human data.[2]

A PREATTENTIVE ITEM MAP

The GS2 model can simulate a wide range of tasks from feature searches to conjunction searches to serial searches. It yields results that are very similar to the results obtained in experiments on groups of human subjects. Much more detail about the GS2 model and simulation can be found elsewhere (Wolfe, 1994a). GS2, however, has its limitations. Specifically, for this chapter the focus on preattentive processing of features has obscured the need for preattentive processing of items to which to attach those disembodied features. In Triesman's original feature integration theory (Treisman & Gelade, 1980), features are processed preattentively. "Items" in the form of "object files" await the application of attention (Kahneman & Treisman, 1984). Most other work on visual search is similar in that it treats the preattentive feature as channels or filters for relatively primitive properties of the input. In the remainder of this chapter I argue for the need for a preattentive representation of items.

Top and Bottom

One place to begin is with a problem with the analysis of orientation reviewed above. As noted for purposes of the guidance of attention, orientation seems to be specified in terms of simple categories such as *steep* and *shallow*. This poses a problem as illustrated in Figure 6.

A line rotated through 180° is indistinguishable from an unrotated version. This is not true for most objects rotated through 180°. Both arrows in Figure 6b are steep, but to state the obvious, they are not the same.

In Figure 7, the search for an up arrow appears to be quite easy, and, indeed, informal experiments on two well-practiced subjects yielded target trial slopes under 5 ms/item.[3] Of course, this and other related exper-

---

[2] To get the slope ratios to work, Guided Search assumes that the variability of an activation signal decreases as the activation level rises. Details can be found in Chun and Wolfe (in press) and Wolfe (1994a).

[3] For a similar example with somewhat different purpose, see Enns and Rensink (1993).

Here, 20° = 200°.

Here, it does not.

Figure 6

Visual search results, based on search for line segments, ignore the possibility that top and bottom are features.

iments need to be run on more subjects, and there are a host of control experiments that one would want to do before making any definitive statement about the status of preattentive processing of top and bottom. However, for the sake of argument, suppose that these results are confirmed. How should they be understood? It could be that the standard work on orientation has been done with impoverished stimuli and that the orientation is better represented over 360° than over 180°. The alternative is that top and bottom are separable from orientation. After all, all of the

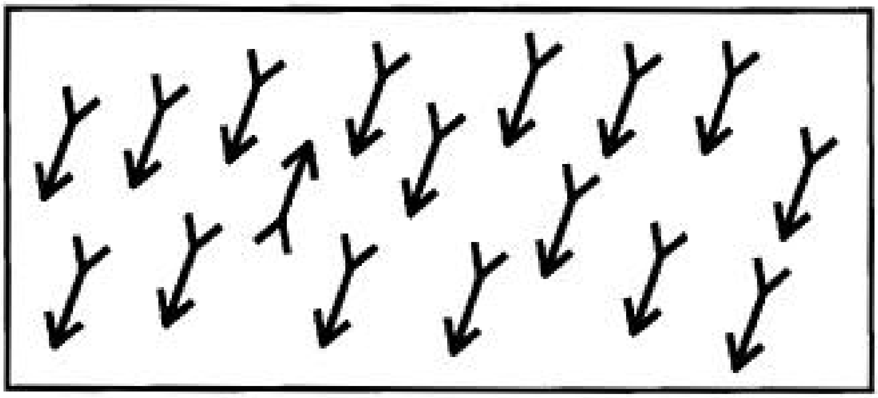


Figure 7

In this search, it appears that the up arrow pops out of a field of down arrows.

items in Figure 7 are "steep" in some meaningful sense. If top and bottom are represented in a different preattentive process, what is that process? Unlike other basic features, top and bottom must be the top and bottom *of something*. Color, orientation, motion, and so forth can all be surface properties that may or may not describe a whole item. Top and bottom are qualitatively different. Imagine a swirling amorphous cloud of features filling the visual field. Imagine that no definable items or objects can be seen. You could still see red. You could still see a vertical. However, you could not see top without also seeing an item that could have a top.

Parts and Wholes

Top and bottom are properties of items. There is other evidence that the structure of items is available preattentively. The inefficiency of Color × Color Conjunctions was described above (Wolfe et al., 1990). To reiterate with a different example of that class of search, it is hard to find a house painted half red and half yellow among houses painted red and blue and houses painted blue and yellow ($M = 38.9$ ms/item for target trials; Wolfe Friedman-Hill, & Bilsky, 1994). As promised earlier, there is a case where search for Color × Color conjunctions is efficient. If subjects search for a red house with yellow windows among red houses with blue windows and blue houses with yellow windows, slopes of RT × Set Size functions are comparable to those seen in standard across-feature conjunctions ($M = 13.1$ ms/item for target trials; Wolfe et al., 1994). It is the structure of the items that appears to be critical. Color × Color conjunctions are efficient when the color of the whole item is conjoined with the color of a part of that item. Conjunctions of the colors of two parts are inefficient.

Part and *whole* are terms of convenience here. Subjects and experimenters tend to spontaneously describe the stimuli in these experiments in part–whole terms, but we are not in a position to say with assurance that the preattentive representation is in terms of parts and whole items. It seems clear that something about the structure of the items is critical here, but we do not fully understand the details. In one series of experiments we found that *surroundedness* is a relevant aspect of the structure (Wolfe et al., 1994). The more one area surrounds another, the more the former acts as a whole and the later as a part. Stimuli for an experiment illustrating this point are shown in Figure 8 with the slopes from the resulting searches given below each type of item.

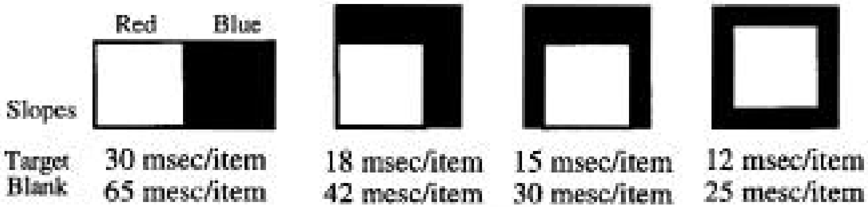
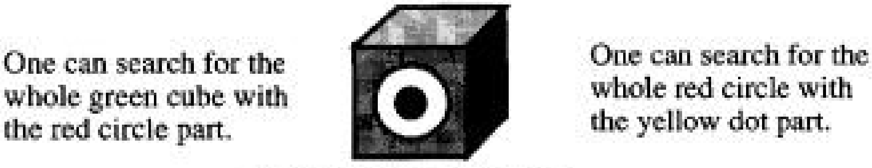


Figure 8

The more that a region of one color surrounds a region of another color, the easier it is to search for a target defined by a Color × Color conjunction.

In all cases, subjects searched for the red–blue target among red–yellow and yellow–blue distractors. The different-colored regions were of identical area and the colors were equiluminous. Other experiments have shown that the division of items into parts and wholes is under some degree of volitional control. Part and whole labels cannot be assigned at random, but using stimuli like the one shown in Figure 9, we (Wolfe et al., 1994) were able to assign a *part* label to the large circle in one condition (search for the green box with the red circle) and assign a *whole* label to the same large circle in another condition (search for the red circle with the yellow spot).

Like Color × Color conjunctions, Size × Size conjunctions are influenced by the structure of the items. It is hard to find the item with big and small parts but relatively easy to find a big item with small parts (Bilsky, Wolfe, & Friedman-Hill, 1994). Interestingly, Orientation × Orienta-



A green cube with a red circle and a yellow dot.

Figure 9

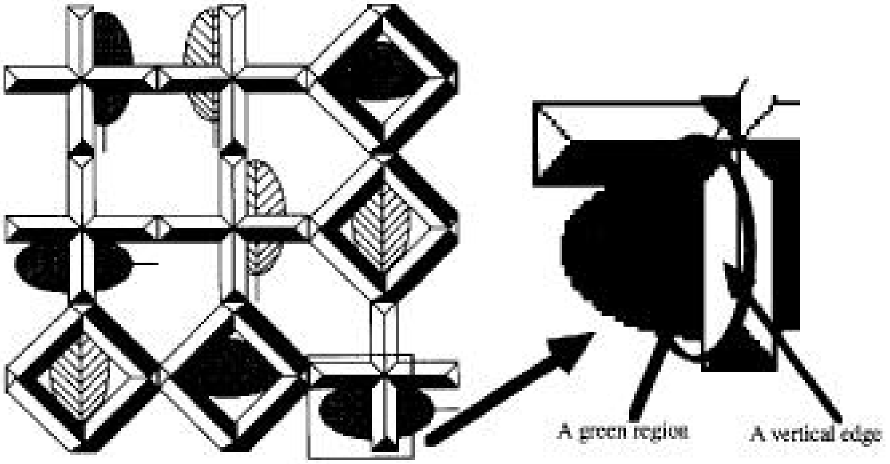
Stimulus used to show that a stimulus component that is a part in one visual search can be the whole item in another.

tion conjunctions do not show a similar sensitivity to part–whole structure. Every one of the large number of Orientation × Orientation conjunctions that we have tried have proven to be highly inefficient (Bilsky et al., 1994). Perhaps absolute orientations of parts and wholes are not coded because they do not stay invariant in the environment. A red item with yellow parts will be red with yellow if it falls over. The same cannot be said of a vertical item with a horizontal part. The relative angular relationship between parts or between parts and wholes may be more robust, and there is evidence for preattentive coding of angular relations (Wolfe & Friedman-Hill, 1992).

Continuous Stimuli and the Ownership of Borders

In most visual search experiments, preattentive processing could be by items, locations, or patches of disembodied features. When the display consists of stimuli presented in isolation on a homogeneous background, division into items is easy and not particularly useful. Attention can be guided to a red vertical item in isolation by finding redness and verticalness in the same place. Real-world stimuli, however, do not offer items in neat isolation on blank backgrounds. Real-world stimuli are continuous. They are also difficult to use in controlled visual search experiments. We (Wolfe et al., 1994) have devised a class of stimuli that are more continuous and naturalistic than standard laboratory stimuli. At the same time, they preserve our ability to place targets and distractors at arbitrary locations from trial to trial. The stimuli are created out of a set of square "tiles." The tiles are drawn so that any regular packing of them will produce a continuous image. Using a set that looked something like aerial views of terrain, I demonstrated that the basic findings of laboratory search studies could be replicated with stimuli of this sort (Wolfe, 1994b).

A different example is shown in Figure 10. Efficient search (~6 ms/item) for conjunctions of color and orientation can be done with stimuli of this sort and it makes little difference if the leaves are in front of or behind the lattice. As illustrated on the right of the figure, this raises a problem for Guided Search and other models. The figure on the right is merely a blow-up of a piece of the stimulus on the left. It points out that because the green leaf is occluded by a vertical segment of the lattice, there is a green vertical contour in the image. If the target is green vertical, why is there no evidence that this spurious green vertical acts as a distractor?



Search for the green (gray) vertical leaf.

Why isn't this "green vertical"?

Figure 10

Continuous stimuli create spurious conjunctions. Here the occlusion of a green leaf produces a green vertical contour in the image. However, these spurious conjunctions do not seriously interfere with search for targets defined by the true conjunctions of these features.

Naive observers have trouble seeing that this is a potential problem. The introspective answer is clear (and probably correct). This green vertical is not a problem because, in a real sense, it does not exist. The greenness and the verticalness are parts of two different items. They seem to lie in two different depth planes. Featural attributes are not disembodied in perception, they are "owned" by objects or items (Nakayama & Shimojo, 1990; Nakayama, Shimojo, & Silverman, 1989). The vertical feature is owned by the lattice. The green feature is owned by the leaf. Obviously, for this to work in an account of guided visual search, these owners must exist preattentively. They do not need to exist in their final form as *lattice* or *leaf*, but they do need to exist as items that can have featural labels attached to them.

The Item Map: Proposal and Implications

To summarize, preattentive processing of top and bottom, if it exists, implies preattentive processing of items that can have tops and bottoms.

Preattentive sensitivity to the part–whole structure of items, which certainly does exist, implies preattentive processing of items that can have parts and wholes. The failure of spurious conjunctions of features to interfere with visual search for items defined by the same conjunction of features implies that basic features can be preattentively assigned to items. My proposal, therefore, is that preattentive processing is not limited to parallel extraction of energy in different sets of feature-tuned channels. Preattentive processing appears to include an initial division of the stimulus into items for subsequent attentive examination.

At this stage, it is hard to be very specific about the nature of the preattentive item representation. Many experiments are needed before it is clear how, for example, parts and wholes are represented. Whether there is a limit on the number of preattentively available items is also got to be determined. Perhaps only a few items can be parsed at a time. A version of this idea can be found in Pylyshyn's finger of instantiation (FINST) model. Specifically, in discussing subitizing, Trick and Pylyshyn suggested that a limited-capacity, preattentive process can handle around five items at once (Trick & Pylyshyn, 1994). Their model does not address the specific search issues discussed here, but in effect, they are proposing that about five of Treisman's "object files" (Kahneman & Triesman, 1984) can be set up at once.

The notion of a fixed or nearly fixed number of items (or FINSTs, if they are the same thing) seems problematic here. Evidence for preattentive item processing (e.g., the part–whole experiments described above) comes from displays with many more than five items. Perhaps the critical idea to borrow from the FINST model is that the item map may be parallel but of limited capacity. Faced with the need to process a host of items in a search for a green leaf, that capacity is spread thin, creating a crude division of the field into items. Different tasks would change the deployment of this limited resource. Faced with the demand to keep track of the movements of a subset of moving items, perhaps the limit is reduced to about five (Yantis, 1992). Faced with a very small number of items, perhaps the capacity is adequate to allow all of the items to be represented in a manner that allows them to make contact with stored representations in memory. For instance, three letters might all activate their respective representations in memory and activations of flanking items might interfere with response to a target letter (Eriksen & Hoffman, 1973).[4] Negative priming might be a similar example where two items are so fully processed that the distractor representation must be inhibited in order to respond to the target (Tipper, 1985, 1992).

The idea of a resource that can be applied broadly with weak effect or narrowly with strong effect is not new. See Nakayama (1990) or Treisman (1993) for a couple of recent examples. It must also be stressed that the connection between preattentive processing of items in visual search and effects like the flanker effect, subitizing, and negative priming is highly speculative.

Is This Just "Late Selection" Revisited?

The proposal that people preattentively parse the visual scene into items has an air of "late selection" about it. However, preattentive items are not merely late selection with another name. It may be that items are found preattentively but, in visual search experiments, there appear to be profound limits on available information about those items. Classic late selection has items reaching their representations in memory in parallel. However, it is clear that the preattentive representation of items is not this elaborate. For example, there is no parallel processing of faces (Nothdurft, 1993b), let alone of face recognition. At an even more basic level, there may be no preattentive representation of spatial relationships such as up, down, left, and right. The orientation of a single element may be known, but subjects show no sign of parallel processing of relationships such as "Is the '+' above the '−'?" (Logan, 1994). This may explain why preattentive processes seem unable to distinguish between *Ts* and *Ls* (Bergen & Julesz, 1983).

It may be that this view of preattentive processing provides a link between early and late selection models. If, as discussed above, the preattentive item map is a limited capacity parallel processor, perhaps it can fully process a small number of items, allowing for the late selection results that are seen in flanker tasks, negative priming, and, perhaps, novel pop-out (Hawley, Johnston, & Farnham, 1994; Johnston et al., 1993). A

---

[4] I recognize that it is quite a reach from visual search for red houses with yellow windows to the flanker effect, but it seemed appropriate in light of the occasion.

proposal of this sort has recently been put forward by Lavie and Tsal (1994), but the matter remains open.

SUMMARY

Guided Search 2.0 gives an account of a wide range of visual search tasks. However, it is not a complete theory of visual search. On the basis of the issues and data presented here, I argue that this model or, indeed, any model of visual search, will need to include some ability to parse the visual scene into searchable items in parallel.

REFERENCES

Alkhateeb, W. F., Morland, A. B., Ruddock, K. H. & Savage, C. J. (1990). Spatial, colour, and contrast response characteristics of mechanisms which mediate discrimination of pattern orientation and magnification. *Spatial Vision, 5*(2), 143–157.

Bergen, J. R., & Julesz, B. (1983). Rapid discrimination of visual patterns. *IEEE Transaction on Systems, Man, and Cybernetics, SMC-13*, 857–863.

Bilsky, A. A., & Wolfe, J. M. (1994). Part–whole information is useful in Size × Size but not in Orientation × Orientation conjunction searches. *Perception and Psychophysics, 57*, 749–761.

Bilsky, A. A., Wolfe, J. M., & Friedman-Hill, S. R. (1994). Part–whole information is useful in Size × Size but not in Orientation × Orientation conjunction searches. *Investigative Ophthalmology and Visual Science, 35*, 1622.

Bravo, M., & Blake, R. (1990). Preattentive vision and perceptual groups. *Perception, 19*, 515–522.

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 and Performance, 16*, 479–492.

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

Chun, M. M., & Wolfe, J. M. (in press). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology.*

Cohen, A., & Ivry, R. B. (1991). Density effects in conjunction search: Evidence for coarse location mechanisms of feature integration. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 891–901.

Dehaene, S. (1989). Discriminability and dimensionality effects in visual search for featural conjunctions: A functional pop-out. *Perception and Psychophysics, 46,* 72–80.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 32–39.

Enns, J. T., & Rensink, R. A. (1990). Scene based properties influence visual search. *Science, 247,* 721–723.

Enns, J. T., & Rensink, R. A. (1992). An object completion process in early vision. *Investigative Ophthalmology and Visual Science, 33,* 1263.

Enns, J. T., & Rensink, R. A. (1993). A model for the rapid discrimination of line drawing in early vision. In D. Brogan, A. Gale, & K. Carr (Eds.), *Visual search* (Vol. 2, pp. 73–89). London: Taylor & Francis.

Eriksen, C. W., & Hoffman, J. E. (1973). The extent of processing of noise elements during selective encoding from visual displays. *Perception and Psychophysics, 14,* 155–160.

Foster, D. H., & Ward, P. A. (1991). Horizontal-vertical filters in early vision predict anomalous line-orientation frequencies. *Proceedings of the Royal Society (London B), 243,* 83–86.

Hawley, K. J., Johnston, W. A., & Farnham, J. M. (1994). Novel popout with nonsense string: Effects of predictability of string length and spatial location. *Perception and Psychophysics, 55,* 261–268.

He, J. J., & Nakayama, K. (1992). Surfaces vs. features in visual search. *Nature, 359,* 231–233.

He, J. J., & Nakayama, K. (1994). Perceiving textures: Beyond filtering. *Vision Research, 34,* 151–162.

Johnston, W. A., Hawley, K. J., & Farnham, J. M. (1993). Novel popout: Empirical boundaries and tentative theory. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 140–153.

Jonides, J., & Yantis, S. (1988). Uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43,* 346–354.

Julesz, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics, 54,* 245–251.

Kahneman, D. (1973). *Attention and effort.* Englewood Cliffs, NJ: Prentice Hall.

Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). San Diego, CA: Academic Press.

Kinchla, R. A. (1977). The role of structural redundancy in the perception of targets. *Perception and Psychophysics, 22,* 19–30.

Kwak, H., Dagenbach, D., & Egeth, H. (1991). Further evidence for a time-independent shift of the focus of attention. *Perception and Psychophysics, 49,* 473–480.

Lavie, N., & Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics, 56,* 183–197.

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

Mack, A., Tang, B., Tuma, R., & Kahn, S. (1992). Perceptual organization and attention. *Cognitive Psychology, 24,* 475–501.

McLeod, P., Driver, J., Dienes, Z., & Crisp, J. (1991). Filtering by movement in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 55–64.

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.

Nakayama, K. I. (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, England: Cambridge University Press.

Nakayama, K., & Shimojo, S. (1990). Towards a neural understanding of visual surface representation. In T. Sejnowski, E. R. Kandel, C. F. Stevens, & J. D. Watson (Eds.), *The brain* (pp. 911–924). New York: Cold Spring Harbor Laboratory.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping and the recognition of occluded objects. *Perception, 18,* 55–68.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Nothdurft, H.-C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research, 31,* 1073–1078.

Nothdurft, H.-C. (1993a). The conspicuousness of orientation and visual motion. *Spatial Vision, 7,* 341–366.

Nothdurft, H.-C. (1993b). Faces and facial expression do not pop-out. *Perception, 22,* 1287–1298.

Nothdurft, H.-C. (1993c). The role of features in preattentive vision: Comparison of orientation, motion and color cues. *Vision Research, 33,* 1937–1958.

Pashler, H. (1992). Attentional limitations in doing two things at the same time. *Current Directions in Psychological Science, 1*(2), 44–48.

Pashler, H. (1994). Dual task interference in simple tasks: Data and theory. *Psychological Bulletin, 116,* 220–244.

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 and Psychophysics, 41,* 455–472.

Sternberg, S. (1969). High-speed scanning in human memory. *Science, 153,* 652–654.

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *Quarterly Journal of Experimental Psychology, 37A,* 571–590.

Tipper, S. P. (1992). Selection for action: The role of inhibitory mechanisms. *Current Directions in Psychological Research, 1,* 105–108.

Treisman, A. (1993). The perception of features and objects. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control* (pp. 5–35). Oxford, England: Clarendon Press.

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

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

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101,* 80–102.

Wolfe, J. M. (1992a). "Effortless" texture segmentation and "parallel" visual search are not the same thing. *Vision Research, 32,* 757–763.

Wolfe, J. M. (1992b). The parallel guidance of visual attention. *Current Directions in Psychological Science, 1*(4), 125–128.

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

Wolfe, J. M. (1994b). Visual search in continuous, naturalistic stimuli. *Vision Research, 34,* 1187–1195.

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

Wolfe, J. M., & Franzel, S. L. (1988). Binocularity and visual search. *Perception and Psychophysics, 44,* 81–93.

Wolfe, J. M., & Friedman-Hill, S. R. (1992). 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 and 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 and Performance, 18,* 34–49.

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 × Color and Orientation × Orientation conjunctions. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 879–892.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology, 24,* 295–340.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2*(5), 156–161.

Yantis, S., & Jones, E. (1991). Mechanisms of attentional priority: Temporally modulated priority tags. *Perception and 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 and Performance, 16,* 121–134.