Conjunction Search Revisited

Anne Treisman and Sharon Sato
University of California, Berkeley

Search for conjunctions of highly discriminable features can be rapid or even parallel. This article explores three possible accounts based on (a) perceptual segregation, (b) conjunction detectors, and (c) inhibition controlled separately by two or more distractor features. Search rates for conjunctions of color, size, orientation, and direction of motion correlated closely with an independent measure of perceptual segregation. However, they appeared unrelated to the physiology of single-unit responses. Each dimension contributed additively to conjunction search rates, suggesting that each was checked independently of the others. Unknown targets appear to be found only by serial search for each in turn. Searching through 4 sets of distractors was slower than searching through 2. The results suggest a modification of feature integration theory, in which attention is controlled not only by a unitary "window" but also by a form of feature-based inhibition.

Objects in the real world vary in a large number of properties, at least some of which appear to be coded by specialized, independent channels or modules in the perceptual system (see Braddick, Campbell, & Atkinson, 1978; Graham, 1985; Livingstone & Hubel, 1987; Treisman, 1986; Treisman & Gormican, 1988, for some reviews of the evidence). To perceive and identify the many thousands of objects one encounters each day, one must specify not only their separate features[1] but also how these features are combined in the correct structural relations. If every possible conjunction had to be directly sensed by its own specialized detectors, there would quickly be a combinatorial explosion. Three general solutions seem possible: (a) A first solution would be to index the separate features present at any time by the locations they occupy and to scan those locations serially, conjoining the features currently attended (Milner, 1974; Minsky, 1961; Treisman, 1977; Treisman & Gelade, 1980). (b) A second solution would use differences in the latency of the neural information coming from different objects as they appear, disappear, move, or change, and would conjoin features whose onsets coincide in time (Von der Malsburg, 1985). (c) A third solution (Pomerantz, Sager, & Stoever, 1977; Treisman & Paterson, 1984) is to code at least some subset of possible conjunctions by directly sensing emergent features of their structure (e.g., closure for the three lines of a triangle; shape

or area for the length and width of a rectangle). In addition, further special strategies may be used to conjoin features in particular perceptual tasks. In this article, we discuss two such strategies that may play a part in visual search.

Treisman and Gelade (1980) and Treisman and Schmidt (1982) reported a variety of results consistent with the first hypothesis, invoking spatial attention. Search for targets defined only by a conjunction of features gave linear functions relating latency to the number of items in the display, suggesting a serial check of each distractor in turn. When attention was divided, subjects reported many illusory conjunctions, recombining features from different objects present in the display. Perceptual segregation and boundary detection appear to be mediated by differences in separate features but not by conjunctions of features. Identification of conjunction targets was totally dependent on correct localization, whereas identification of feature targets could be correct even when they were mislocated in the display. Finally, more recently, Grabowecky and Treisman (see Treisman, 1988, pp. 213–214) found that the probability of correct report of conjunctions of features could be quite accurately predicted from the product of the probabilities of correctly reporting each of their component features. This was true even at zero delay between the display and the cue indicating which item should be reported. Thus, there was no evidence for an initial holistic perception followed by rapid decay of the conjunction information.

Most of these results were obtained with conjunctions of color with aspects of shape (curved vs. straight edges or vertical–horizontal vs. diagonal), but serial search, illusory conjunctions, and failures of texture segregation have been shown also for parts of shapes (Julesz, 1986; Prinzmetal, 1981; Treisman & Gormican, 1988; Treisman & Paterson, 1984)

[1] We will use the term "feature" to refer to a value on a dimension (e.g., "red" on the color dimension; "vertical" on the orientation dimension). A dimension is a complete set of mutually exclusive values, at least one of which must characterize any stimulus to which the dimension applies.

and illusory conjunctions have been found for color, size, and outline versus filled shape (Treisman & Schmidt, 1982).

The second account of the conjoining process—the temporal coincidence hypothesis—was recently tested by Keele, Cohen, Ivry, Liotti, and Yee (1988), who found no indication that illusory conjunctions occur any more frequently for features whose presentation times coincide than for those whose presentations appear sequentially within 166 ms. Further evidence against the temporal coincidence account is the finding that features do appear to migrate between successive temporal intervals (Intraub, 1985; Lawrence, 1971), provided that they appear in the same location (McLean, Broadbent, & Broadbent, 1982).

The third hypothesis, that some conjunctions are directly sensed by specialized detectors, is consistent with physiological evidence that single units in most visual areas respond selectively on more than one physical dimension. Most cells in area V1, for example, are tuned both for spatial frequency and for orientation (De Valois, Yund, & Hepler, 1982); many cells here and in prestriate areas are tuned both to a particular direction of motion and to a particular orientation, or both to a color and to an orientation (e.g., Desimone, Schein, Moran, & Ungerleider, 1985; Maunsell & Van Essen, 1983, Thorell, De Valois, & Albrecht, 1984). However, one cannot assume that the organism can directly access the specialized sensitivities of any individual cells, and even if it could, the message from any one cell is inherently ambiguous because of the principle of univariance. The effective perceptual codes are likely, therefore, to consist of distributed patterns of activity across large populations of cells, and these could reflect separate dimensions rather than conjunctions of features.

There is behavioral evidence for a very limited number of emergent features. Closure (a triangle among separate lines and angles) can mediate parallel search and seems also to prevent the formation of illusory conjunctions (Treisman & Paterson, 1984). A few three-dimensional features, such as the orientation of a cube (Enns, in press), the direction of lighting (Enns & Rensink, 1990), and convexity conveyed by gradients of shading (Ramachandran, 1988), can mediate grouping or apparent motion as well as rapid or parallel search, offering some support for the third hypothesis as well as the first. However, the number of emergent features directly sensed by the visual system must be limited in order to avoid the combinatorial problem. Treisman and Gormican (1988) looked for parallel processing of simple emergent features produced by relating pairs of oriented lines (e.g., potential features such as intersection, juncture, and convergence). By the parallel search criterion, we found no evidence that any of these was directly available at preattentive levels.

The spatial attention hypothesis seemed, then, to offer the best general account of the data available. In the past 4 years, however, a number of investigators have reported exceptions to the claim that search for conjunction targets must be serial. Nakayama and Silverman (1986a) found that targets defined by conjunctions of binocular disparity with color and with motion gave flat search functions relating latency to the number of elements. Conjunctions of color and motion, on the other hand, gave steeply increasing linear slopes. The

parallel conjunction of disparity with color or with motion could be explained by extending the spatial attention hypothesis to allow selection of a plane in depth (cf. Downing & Pinker, 1985). The odd color or direction of motion would then "pop out" of the selected plane because of its unique value on that single dimension.

However, some further exceptions have since been discovered: Nakayama and Silverman (1986b) found parallel (or close to parallel) search functions for a different version of color–motion conjunctions and for every pairing of binocular disparity, spatial frequency, size, color, and direction of contrast, provided that the two values on each dimension were highly discriminable (e.g., bright red and green patches, motion oscillating vertically vs. horizontally, black vs. white on a gray background). McLeod, Driver, and Crisp (1988) found almost flat slopes for conjunctions of shape with direction of motion; Steinman (1987) found the same for conjunctions of binocular disparity with orientation and with Vernier offsets, and, after extended practice, for conjunctions of Vernier offset with orientation and lateral separation; Wolfe, Cave, and Franzel (1989) reported completely flat functions for conjunctions of highly discriminable sizes, orientations (horizontal and vertical bars), shapes (plus and circle), and colors (red and green).

In addition, a finding by Pashler (1987) cast some doubt on the claim that search was serial and self-terminating when displays of fewer than eight items were used. Even though search latencies increased linearly with display size in his experiments, the slopes for negative and for positive trials were parallel rather than in the two-to-one ratio that we had previously found with larger displays. Pashler suggested that subjects might search groups of up to eight items in parallel and that search became serial and self-terminating only across separate groups of about eight items at a time. The parallel slopes with small display sizes are not a universal finding: Parallel functions were found also by Houck and Hoffman (1985), but in other experiments (size–shape conjunctions in Quinlan & Humphreys, 1987; shape–color in Treisman & Gelade, 1980) there is little sign of a break in the search function around display sizes of eight. It is not yet clear under what conditions one finds parallel slopes, but it will be important to clarify the controlling factors.

The finding of parallel search for conjunction targets appears inconsistent not only with feature integration theory (Treisman & Gelade, 1980) but also with the data from the other experimental paradigms that had initially prompted the theory. It therefore seems worth exploring carefully both the conditions that allow parallel detection of conjunction targets and any accounts that could reconcile that result with the other findings described above. Prompted by the initial reports by Nakayama and Silverman (1986b), we began a series of experiments to replicate their results and to explore some possible interpretations with further experimental tests. In particular, we considered whether special strategies to control attention might be available in the search task but not more generally in other perceptual tasks.

We tested three possible strategies for conjunction search, each of which could be consistent with the previous, more general account of spatial attention and feature integration.

The first is that special grouping mechanisms might be invoked to segregate the two sets of distractors, allowing selective attention to one set and single feature search within the selected set (Dehaene, 1989; Nakayama, 1990; Steinman, 1987; Treisman, 1988), as previously shown for spatially grouped distractors (Treisman, 1982). The second is that subjects might use a small number of conjunction detectors for certain pairs of dimensions, available at preattentive levels of processing and activated by highly discriminable pairs of features. Likely candidates would be the feature pairs that activate single cells at early stages of visual coding. The third is that some preselection might be achieved by reducing the activation of distractor locations containing features that are inconsistent with the target.

Two of these hypotheses suggest new ways in which selective attention may modulate visual processing to allow the correct conjunctions of features to be formed. In feature integration theory, as it was previously formulated, the sequential processing of objects was achieved by a spatial scan of one location at a time. Figure 1 (from Treisman, 1988) illustrates how attention could be used to ensure the correct conjunctions of features. The selection is controlled extrinsically by a spatial aperture or "window"[2] that can be narrowly focused or more widely opened (cf. the "zoom lens" analogy used by Eriksen & Hoffman, 1972, and the spotlight or searchlight analogy used by Crick, 1984; Treisman, 1982). There is some evidence suggesting that the attention window is unitary and cannot normally be opened onto two spatially separated locations at once (Posner, Snyder, & Davidson, 1980), although other results have qualified this claim (Bashinski & Bacharach, 1980). Feature integration theory suggests that attention selects one area at a time within a "master map" of locations,

thereby retrieving the features linked to the corresponding locations in a number of separable feature maps (Treisman, 1985). The alternative segregation and feature inhibition strategies that we consider in this article control selection through the same master map of locations, but do so by reducing the activation from one or more of the feature maps instead of through an externally controlled scan. The segregation hypothesis assumes that one set of stimuli is selectively inhibited, leaving the other set available for attentional processing. The feature inhibition hypothesis assumes that inhibition can be controlled through more than one feature map, reducing the interference from all distractor locations rather than from a single subset. A similar account has been proposed by Wolfe et al. (1989); we discuss their results and a possible way of distinguishing two versions of the model later in this article. The third hypothesis, based on conjunction detectors, is tested in Experiments 2 and 3.

The Segregation Hypothesis

We begin by considering the possibility that parallel detection of conjunction targets in visual search depends on perceptual segregation between the two sets of distractor items. Many of the conjunctions that Nakayama and Silverman (1986b) tested include features related to phenomenological separation in depth. Binocular disparity is the most obvious example, but stimuli differing in the direction of motion and stimuli differing in size or spatial frequency also often appear to segregate into different planes. Both motion parallax and size gradients are useful cues to depth. If such perceptual segregation appeared salient, subjects might attend selectively to one of the two planes and do a parallel feature search within that plane for the other target-defining feature. For example, in a display of color–motion conjunctions, the items oscillating horizontally might segregate from those oscillating vertically. Within either plane, a target differing in color from the distractors should then pop out without any need for focused attention to each item in turn.

The feature integration model can be modified to allow this optional strategy when the two sets of distractors differ in some highly discriminable feature. The suggestion is that spatial selection can be achieved not only by an externally controlled window acting directly on the master map but also by changing the relative activation produced in the master map by one or other of the distractor feature maps (Treisman, 1988, see Figure 2). If attention could control the level of activation of some subset of master-map locations through their links to one or more feature maps reducing the activity in locations that contained distractors with a salient nontarget feature, a parallel feature search across the remaining locations might be sufficient to detect the target. Whereas the selection that is extrinsically controlled by an attention window seems to be restricted to a single area at a time (Posner et al., 1980),
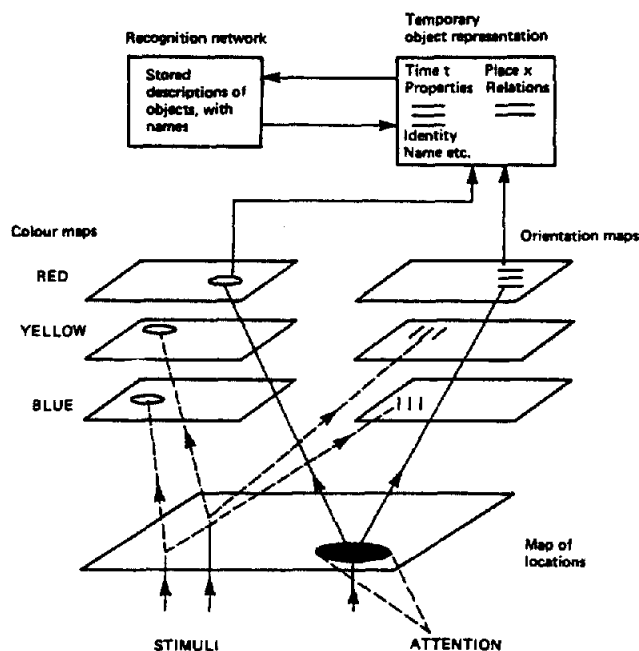

Figure 1. Model for the role of attention in feature integration (from Treisman, 1988).

the inhibition controlled through a feature map could affect locations that are spatially interspersed with other, noninhibited locations. The effect of selection would otherwise be the same in both cases: It would limit the set of features that are passed on together to be conjoined as parts or properties of the same perceptual object.

Thus, for dimensions on which two sets of distractors differ sufficiently to produce nonoverlapping distributions of activity in feature space, the constraints imposed by a unitary spatial attention window would become irrelevant. As the target and distractor features become more similar, the feature-based inhibition would have progressively less effect on the signal-to-noise ratio, and the external scan of locations with the attention window would become more important. The display would be scanned with more and more narrowly focused attention, giving increasingly steep search functions (Treisman & Gormican, 1988).

The concept of feature inhibition developed here differs from that proposed by Bjork and Murray (1977). In our account, feature inhibition is an optional strategy used to facilitate selective attention rather than an automatic form of mutant lateral suppression generated between neighboring identical features. The feature inhibition we envisage is not a local interaction, and it is reversible when the target of attention is changed. In addition to facilitating rapid search for conjunction targets, it provides a mechanism for figure-ground segregation, which is an essential task for early vision.

To avoid circularity, however, we need some independent measure of the extent to which particular displays allow perceptual segregation and selective attention to a subset of

Figure 2. Model for modulating attention by inhibition from a feature map as well as from an attention window (from Treisman, 1988).

spatially interspersed elements. Otherwise we are reduced to inferring salient segregation from conjunction pop-out, while at the same time using it to explain conjunction pop-out, like explaining the effect of opium by its "dormitive power." In Experiment 1, we describe an attempt to find converging evidence that perceptual segregation allows selective access to all the elements of one type, using displays in which we also test conjunction search.

Experiment 1: Perceptual Grouping and Global Shape Recognition

The index of segregation that we chose measures the ease of access to the global shape of a perceptual group. The claim is that if subjects can selectively attend to a whole subset of distractors, boundaries of the selected set should be simultaneously available to mediate recognition of their global shape. Julesz (1971) showed that boundaries between areas can be defined by differences in the binocular disparity of the dots they contain. These boundaries create global shapes in the same way as boundaries defined by discontinuities in color or luminance. Similarly, Cavanagh (1987) has shown that boundaries defined by differences in the spatial frequency, motion, or disparity of random dots can support many aspects of shape discrimination. In these demonstrations, however, the elements to be grouped were spatially contiguous or formed a good Gestalt (e.g., an annulus or outline cube). In the present experiment, we ask whether salient differences in features can mediate segregation and shape recognition for a randomly distributed set of elements interspersed with other irrelevant distractor elements.

The subjects' task was to decide whether one set of four or five identical elements (in a 3 × 3 matrix) defined an area that matched a global gray pattern presented simultaneously beside them, or whether they differed from the gray pattern by one square, either added or deleted. We tested six types of displays generated by conjoining pairs of values on each of four dimensions: color, size, orientation, and direction of motion. One obvious difference between Nakayama and Silverman's (1986b) results and our earlier tests of conjunction search was that the features they used were highly discriminable. We therefore used rectangular bars differing in pairs of values that were maximally discriminable on each of the four dimensions. Figure 3 shows examples of these displays with elements differing in size and orientation, together with the corresponding gray shapes for the "same–different" discrimination. The prediction from the perceptual segregation account of conjunction search is that there should be a correlation between the speed of matching the gray shape to the area containing a particular set of distractors and the speed of search for a conjunction target among the same sets of distractors.

Method

a)

b)
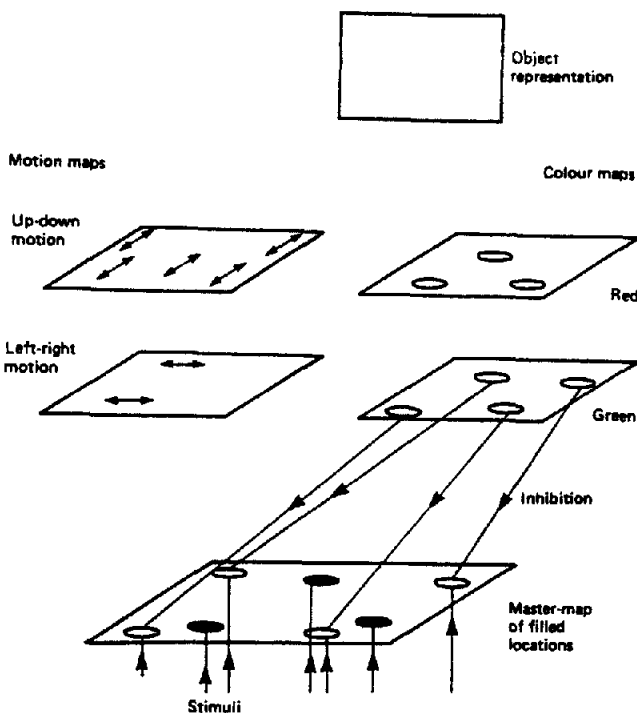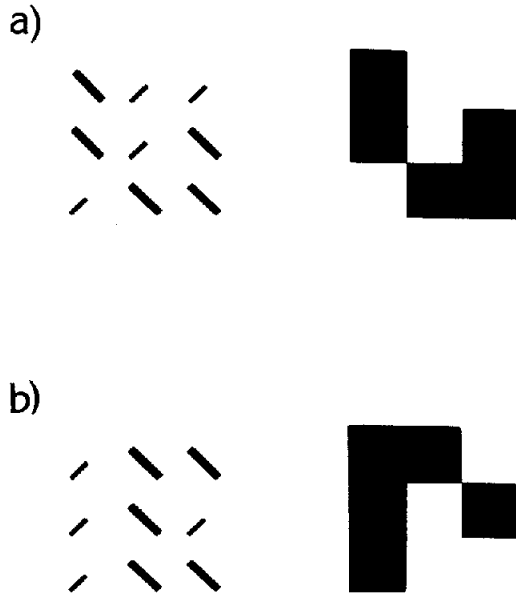
Figure 3. Examples of size–orientation displays used in same–different matching task of Experiment 1; (a) "same" pair; (b) "different" pair.

one of two directions. They could differ along each of four dimensions: The colors we tested were pink (luminance 15.9 cd/m²; CIE [Commission Internationale d'Éclairage] coordinates .363, .220) and green (luminance 18.7 cd/m², CIE coordinates .228, .429) against a white background (luminance 48.8 cd/m², CIE coordinates .254, .250). The sizes were 0.5° × 1.6° and 0.3° × 0.9°; the orientations were 45° and 135°; and the directions of motion were vertical and horizontal oscillation (over a distance of 0.4°).

The four dimensions were paired to form six conjunction conditions: motion–color (MC), motion–size (MS), motion–orientation (MO), color–orientation (CO), size–orientation (SO), and color–size (CS). In the motion–color condition, for example, a display would contain either pink bars moving vertically mixed with green bars moving horizontally, or pink bars moving horizontally mixed with green bars moving vertically. The two dimensions not tested in any given condition were set at neutral values: These were gray for color (luminance 13.6 cd/m², CIE coordinates .246, .234), vertical for orientation, 0.4° × 1.1° for size, and stationary for motion. The six conditions were run in separate blocks.

In the conjunction search task, displays of 4, 9, and 16 bars were used, randomly mixed within blocks. The two types of distractors in each condition were randomly placed in the cells of a square matrix, keeping density fixed and equating as closely as possible the number of distractors of each type. The largest matrix (4 × 4) subtended 8.5° at a viewing distance of 50 cm. The centers of the 2 × 2 and 3 × 3 matrices were randomly located within the larger 4 × 4 matrix to equate the average distance of the stimuli from the initial central fixation point. In half of the displays, one distractor was replaced by a target, which always shared one feature with each of the distractor elements. For example, with small green bars and large pink bars, the target for half the subjects would be a small pink bar, and for the other half it would be a large green bar. The location of the target was selected randomly on each trial. For each condition, all four combinations of the values on the two dimensions were tested with different subjects.

In the same–different matching task, two displays were presented side by side. On the left was a display of bars, identical to the 3 × 3

condition used in search on nontarget trials. This display always contained nine elements (five of one type of distractor and four of the other), randomly positioned in a square matrix. On the right was a display consisting of connected solid gray squares (11 × 10 mm each), filling the squares in the matrix corresponding to those that contained a selected set of distractors. On "same" trials, the gray squares matched the area occupied by the selected set; on "different" trials, one gray square was randomly added or deleted so that the areas differed by either 20% or 25% from the corresponding perfect match. Each display was 66 × 66 mm. The two displays were approximately 42 mm apart. Thus the total display was 174 × 66 mm (visual angle = 19.7° × 7.6°).

Procedure. Subjects were tested on the conjunction search tasks in the first two sessions, and then in the same–different matching task in a third session. In the conjunction search tasks, on each trial, we asked the subject to determine whether the display contained a target and to press one key if the target was present and another key if it was not. The assignment of right and left hand keys was counterbalanced across subjects. A consistent mapping procedure was used: Each subject was assigned one feature from each of the four dimensions that would define his or her conjunction targets in all the different conditions. For instance, the first subject was assigned pink color, vertical motion, large size, and 45° orientation. The targets for this subject were pink bars with vertical motion, bars tilted 45° with vertical motion, large bars with vertical motion, large pink bars, pink bars tilted 45°, and large bars tilted 45°. Each value on each dimension was used equally often across the 8 subjects and all combinations of values were tested, except that orientation was perfectly correlated with color. (Sixteen subjects would have been needed to test all possible combinations, and these two dimensions seemed least likely to generate an emergent feature.) The order in which the 6 conditions were tested was counterbalanced across subjects.

A fixation point appeared 1.75 s before moving displays and subjects were given 8 s to respond. With stationary displays, the fixation appeared 1 s before, and subjects had 5 s to respond. The subject's response triggered the next trial with a 500-ms delay. Feedback was given on incorrect responses. Subjects were tested in two sessions with three blocks of 48 trials in each condition in each session. The first of the three blocks was discarded each time as practice.

In the same–different matching task, subjects were tested with the same six sets of stimuli. They were told that on each trial they would see two displays. On the left would be a display identical to the 3 × 3 displays they had seen in the previous experiment, except that there would be no targets. This display would be divided into two groups of elements. Subjects were instructed to attend to one group of elements and to try to notice the global shape formed by the boundaries that divided them from the other elements. They were to decide whether the global shape containing the selected elements exactly matched the formation of gray squares displayed on the right. If the two formations were identical, subjects pressed one key; otherwise, they pressed the other key. They were instructed to respond as quickly as possible without making any errors. Subjects were shown examples of each of the six conditions and were asked to select which type of distractors they wanted to match to the gray squares in each condition (e.g., pink moving vertically or green moving horizontally in the color–motion condition). The free choice allowed subjects to benefit from any learned segregation strategy they might have developed during the conjunction search sessions.

A fixation point appeared 1 s before each display. Subjects were given up to 8 s to respond and received feedback on incorrect responses.

Subjects. Eight subjects (4 men and 4 women) were tested. They were students at the University of California, Berkeley, who volunteered for the experiment and were paid $5 an hour for participation.

Results and Discussion

Table 1 gives the mean search times and error rates and the same–different matching times and error rates for each display type. We first discuss the search data.

Conjunction search. All the search functions have slopes that are significantly greater than zero ($p < .001$ in every case). The slopes are all quite linear, with nearly all the variance due to display size accounted for by the linear functions. There were consistent differences in the conjunction slopes for the different targets, $F(5, 35) = 2.82$, $p < .05$, in an analysis of variance (ANOVA) on the slopes for each subject with each of the six conjunction targets. Some conjunctions were clearly more difficult to detect than others: Size–color conjunctions were easiest and motion–orientation the most difficult.

The results are consistent with other recent findings that suggest that search for conjunction targets can be very rapid. Although the mean slopes for conjunction targets in all six conditions were significantly greater than zero, the effects of display size were appreciably less than in the early experiments in which they ranged from about 40 to 100 ms per item on negative trials (Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977). Three of the present subjects had mean slopes for target present that averaged 10 ms or less across the six conditions, and 7 of the 8 subjects had at least one target for which the slope on target present trials was under 10 ms. For the 8 subjects tested, the mean number of conditions with positive slopes under 10 ms was 2.6 out of 6.

We had previously shown that the discriminability of the component features can have a substantial effect on conjunction search rates (Treisman & Gelade, 1980, Experiment 2). In that experiment, we compared search for a more discriminable target (red O among green Os and red Ns) with search for a less discriminable target (green T among green Xs and blue Ts). The slopes for the more discriminable targets were less than half those for the less discriminable ones (40 ms compared with 100 ms per item on negative trials). We attributed the reduction to a faster but still serial check of the more discriminable distractors. In that experiment, we used tachistoscopic presentation of line drawings in colored ink on white cards. The present results, with brighter, filled bars in computer-generated displays, take the discriminability effect further and cast some doubt on the idea that search remains

serial for each separate item in the display. We discuss the conjunction search results further in the context of Experiment 2.

Same–different matching. Subjects were able to match with fairly high accuracy the shape of the gray area to the global area containing one set of bar stimuli; errors were 6% or less in all conditions. The latencies varied significantly with the stimuli, $F(5, 35) = 4.63$, $p < .01$, with size–color again the easiest and motion–orientation the hardest.

The main question raised in this experiment concerned the correlation between same–different matching times and the speed of conjunction search with the same displays. There was a strong positive correlation (.92) across display types between the group slopes (mean of positive and negative slopes) and the group mean latencies in the same–different matching task. The correlations within individuals were, however, in most cases much lower than the group mean. The 8 individual subjects' correlations were as follows: −26, .08, .10, .18, .30, .36, .70, and .73, giving a mean r of .31. Although seven of the eight are positive, none of these correlations individually reached significance. However, practice effects and the choice of targets could not be counterbalanced within each individual subject and clearly added noise to the data. There does, for the group as a whole, appear to be some shared factor determining the efficiency of segregation and global matching and the speed of search for a conjunction target. Because the individual data are less compelling, however, it is worth exploring the other possibilities outlined in the introduction before accepting the segregation account as a complete explanation.

The Conjunction Detector Hypothesis

The second hypothesis we proposed to account for cases in which conjunction search is fast or even independent of display size was that, for certain pairs of dimensions, there might after all exist a number of specialized detectors coding conjunctions of values as integral perceptual units. Likely candidates are those pairings that signal important variables in the real three-dimensional world, for example "looming" (pairs of diverging parallel edges) or "shape from shading" (luminance or texture gradients created by changing illumination on solid objects). Regan, Beverley, and Cynader (1979) have in fact found single units that appear selectively to code

Table 1
Mean Slopes of Search Functions and r^2 Measure of Linearity; Mean Matching Times and Errors in the Same–Different Matching Task

	Search functions					Matching task	
	Positive		Negative				
Stimulus displays	Slope	r^2	Slope	r^2	Errors %	Reaction time	Errors %
Size–color	10.2	.995	17.3	.961	1.9	1243	6.6
Size–motion	11.7	.989	30.2	1.000	2.6	1362	4.3
Size–orientation	17.4	.968	35.2	1.000	2.9	1374	5.4
Color–motion	11.0	.930	23.6	.999	1.6	1332	4.3
Color–orientation	18.4	.998	37.0	.998	3.0	1336	5.2
Motion–orientation	20.5	.936	61.4	.998	4.8	1496	5.0

looming. However, the search experiments that have shown parallel coding of conjunctions used apparently arbitrary, though highly discriminable, pairs of values (e.g., red vs. green with circle vs. cross or with horizontal vs. vertical motion). If these conjunctions of color with shape or direction of motion can be directly sensed, we might expect the same direct coding to be available for all arbitrary pairings of equally discriminable values on the same dimensions.

Can we predict from the physiological evidence which pairs of dimensions are most likely to be coded as conjunctions? Single units in many visual areas (V1, V2, V4) do appear to be tuned to different combinations of particular orientations with particular spatial frequencies or particular directions of motion (Desimone et al. 1985; DeValois, Albrecht, & Thorell, 1982). There seem to be fewer units responding to combinations of color with orientation, at least at high spatial frequencies (Hubel & Livingstone, 1987), and even fewer responding selectively to conjunctions of color and motion. Color and motion appear to be largely segregated between the parvo- and the magno-cellular pathways and between areas V4 and MT (see reviews by De Yoe & Van Essen, 1988; Livingstone & Hubel, 1987; Maunsell & Newsome, 1987), although the separation of different attributes on different pathways is far from absolute (see, for example, Albright, 1984; Van Essen, 1985). There is also psychophysical evidence of separation in the coding of motion and color, for instance the marked reduction or disappearance of perceived motion in isoluminant displays (e.g., Cavanagh, Tyler, & Favreau, 1984). The prediction for search, then, is that color–motion conjunctions should be less likely to give parallel search functions than conjunctions involving orientation with size or motion.

A second test of the conjunction detector hypothesis looks for *separable* effects of the component features on conjunction search latencies. If the relevant conjunctions are directly sensed, there is no reason to expect consistent effects of the different component features on search rates. In Experiment 2 we compare search latencies with the same four pairs of values in each of the six possible combinations of dimensions. We used the same displays as in Experiment 1, but collected more extensive data. This allowed us to ask whether the ease or difficulty of conjunction search depends on the particular *conjunctions* tested, or whether it can be predicted simply from one or both of the two features whose conjunction defines the target.

Experiment 2: Conjunction and Feature Search With Targets Defined by Color, Size, Orientation, and Motion

In this experiment, we replicated the search conditions of Experiment 1, testing more subjects in more sessions. We also compared performance in search for conjunction targets with performance in search for targets defined by each of the same features on its own.

In addition, as a further test of the perceptual segregation hypothesis, we looked at the effects of two different spatial distributions of the distractors in one of the conjunction search tasks. Subjects search much more rapidly for conjunc-

tion targets when the distractors are spatially grouped into homogeneous clusters (Treisman, 1982); the search rates suggested a serial check of groups rather than of individual items. In random arrays, there will also be homogeneous groups of various sizes. If the rapid search rates in Experiment 1 depended on parallel search through any spatial pairs, triplets, or larger clusters that happened to contain homogeneous items, we might expect substantial increases in the slopes of the search functions when spatial clusters are eliminated (e.g., with a regular checkerboard) and perhaps also when the displays are spaced out over a larger area, making homogeneous clusters less salient.

Method

Stimuli. In the six different conjunction search conditions, the displays were identical to those used in Experiment 1, except that for 8 of the 10 subjects, the viewing distance was accidentally increased to 56 cm with a corresponding decrease of 12% in all the visual angles. For the feature search conditions, the distractors were homogeneous and differed from the target only on one of the four dimensions, with the other three dimensions set at the neutral value.

To test the effect of different spatial distributions, we used displays with motion–color conjunctions. In one condition, the display was spread over a larger area, with the distances between the bars increased by a factor of 1.5, keeping the bar size constant. In another condition, the two types of distractors were presented in the usual area, but they were arranged in a regular alternating checkerboard pattern. There were therefore no homogeneous clusters of adjacent items.

Procedure. The procedure was the same as in the search tasks of Experiment 1, with the exception that subjects were tested on the six conjunction search conditions in three 1-hr sessions on separate days. Each session consisted of four blocks of 48 trials for each condition. The subjects were therefore tested in a total of 3,456 trials—twice as many as the subjects in Experiment 1. The first session and the first block in each condition of the other two sessions were discarded as practice, leaving 288 trials to be analyzed for each subject with each type of conjunction target. The order of conditions was counterbalanced across subjects.

Ten subjects were tested, of whom 8 were also tested in two further sessions with the targets defined by a single feature on each of the four dimensions and with the two spatial variants of the motion–color conjunctions. The order of conditions within the single feature search and within the different spatial tests with color–motion conjunctions was counterbalanced across subjects.

Subjects. The 10 subjects (4 men and 6 women) were students at the University of California, Berkeley, who volunteered and were paid $5.00 an hour for participation.

Results and Discussion

The search functions in each condition are summarized in Table 2, which gives the mean slopes relating search time to display size, the intercepts, and the proportion of the variance with display size that can be accounted for by a linear function. Error rates averaged 6% or less at all display sizes in all conditions. Figure 4 shows the mean latencies in the conjunction and in the feature search conditions. The feature search conditions all gave very short latencies and flat search functions, none of which departed significantly from zero slope with display size. Thus, the different values on each of the

Table 2
Mean Search Functions in Experiment 2 and r² Measures of Linearity

| Condition	Target present			Target absent		
	Slope	Intercept	r^2	Slope	Intercept	r^2
Conjunction						
Size–color	6.8	432	.993	8.0	487	.960
Size–motion	9.0	580	1.000	16.1	607	.993
Size–orientation	10.9	536	.994	21.0	517	1.000
Color–motion	11.8	566	.969	17.7	576	.984
Color–orientation	11.7	533	.995	23.5	480	.994
Motion–orientation	9.5	884	.955	29.5	839	.999
Simple features						
Size alone	1.6	400	.826	0.2	411	.172
Color alone	0.8	356	.417	−1.3	397	.991
Motion alone	−1.4	508	.458	0.2	498	.110
Orientation alone	0.2	393	.267	0.3	420	.085
Spatial variants of color–motion displays						
Standard (session 3)	9.9	521	.999	13.4	548	.999
Larger area	11.7	491	1.000	13.3	508	.973
Checkerboard	9.9	478	.954	11.0	513	.965
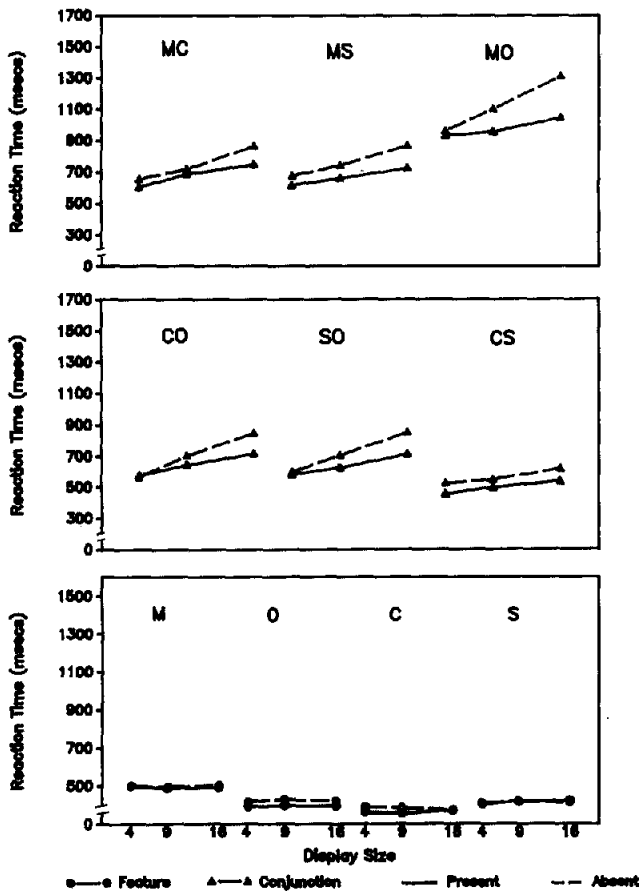
Figure 4. Mean search times in Experiment 2. (M = movement, C = color, S = size, and O = orientation.)

four dimensions were sufficiently discriminable to give parallel detection with the targets popping out of the displays.

As in Experiment 1, the slopes in the conjunction conditions were all significantly greater than zero ($p < .001$ in every case), and again linearity accounted for nearly all the variance due to display size. There were significant differences between the different conjunctions, both in mean search times, $F(5, 45) = 25.10$, $p < .0001$, and in the interaction with display size, $F(10, 90) = 2.85$, $p < .01$. The mean search times did not differ significantly for the subjects in the two experiments, but the search rates (given by the slopes) were faster in Experiment 2 than in Experiment 1, probably because these subjects had twice as much practice. An ANOVA showed significant interactions of experiment with display size, $F(2, 32) = 5.08$, $p < .05$, with target type, $F(5, 80) = 62.11$, $p < .0001$, and a three-way interaction with display size and positive versus negative trials, $F(2, 32) = 5.56$, $p < .01$. Within Experiment 2, there were also large individual differences in the mean search rates averaged across all six conditions (with a range from 7.1 to 31.0 ms per item, and there were differences between individuals in which conditions were most difficult.

The slope ratios relating positive to negative trials averaged .57 across all display sizes for all conditions. Overall, we do not replicate Pashler's (1987) finding of parallel slopes up to display size 8. However, the color–size and the color–motion conditions did produce Pashler's result in this experiment, and color–size also did so in Experiment 1. What might the controlling variables be? There is a strong negative correlation of −.85 between the slope ratio from displays of 4 to 9 and the overall mean of the positive and negative slopes for displays of 4 to 16. The easier the search overall, the more parallel the slopes became for small display sizes. The slope

ratios for small displays were also closer to 1.0 for the more practiced subjects of Experiment 2 than for those of Experiment 1. It seems that as search becomes easier, it is increasingly likely to be exhaustive, either because subjects process groups of items in parallel (as proposed by Pashler) or because it is easier, when the search rate is very fast, to scan the display exhaustively than to decide after each item whether to terminate the search (cf. Sternberg, 1966, in the context of memory scanning experiments).

The conditions in which the spatial layout was varied showed almost no effects of spatial spread or of spatial grouping. The effect of layout was not significant, and neither was the interaction between layout and display size. Performance on these motion–color displays seemed to be independent of the regularity, density, and mean size of homogeneous groups. If the flatter search functions depend on segregation of one set of distractors from the other, the segregation in these displays must be controlled by some mechanism that is independent of spatial proximity.

Our main aim was to get more reliable data to explore the conjunction detector hypothesis. Two sources of evidence seem relevant. The first compares the relative difficulty of the different conjunctions with the suggested availability of single unit detectors. The predictions from physiological "conjunction detectors" are not fulfilled. In fact, if anything, the reverse is the case: Conjunctions including orientation as one of the relevant features gave slower search than conjunctions of color and motion. Nakayama and Silverman (1986b) obtained a similar ordering of difficulty, although they did not use identical features in each of their different conditions.

The second test of conjunction detectors looked for interactions as opposed to independence in the contributions of each component feature to the overall search rates. Unlike earlier experiments, this one kept the individual features constant and varied only their conjunctions. We could, therefore, observe the effect of each feature in each possible combination to see whether there was any consistency. This analysis is also relevant to the segregation hypothesis. A simple version would be that subjects use one feature to segregate the display, allowing attention to one subset of distractors from which the conjunction target pops out by virtue of its remaining unique feature. (The feature search results show that each of the four features could be detected in parallel when it uniquely defined the target.) The prediction would be that any condition involving a feature that allowed effective segregation would give equally flat slopes, regardless of the other features with which it was paired. For example, if segregation by size is salient so that targets defined by size and color give color pop-out within the subset defined by the relevant target size, so also should targets defined by size and motion give motion pop-out and targets defined by size and orientation give orientation pop-out.

The data do not conform to this prediction. No single feature within the different conjunctions consistently determined the ease or difficulty of search. Each feature in a conjunction target was associated with a range of different search rates depending on which other feature it was paired with. Because the most salient feature could differ between individuals, we took each subject's flattest slope on trials with

the target present. (We analyzed both the 8 subjects of Experiment 1 and the 10 from Experiment 2.) The mean flattest slope was indeed very shallow (3.5 ms) and the target varied across individuals. If the same feature that allowed parallel detection in the fastest search condition could also determine the search rate when conjoined with each of the other two features, the slopes obtained with conjunctions including the favored feature should be flatter than the others and also more similar to each other than to the remaining conditions. For example, if CS were the fastest condition for one individual, then the slopes for either MS and SO or MC and CO should also be shallower than those for MO. Moreover, the mean difference between SO and MS or between CO and MC should be smaller than the mean difference between SO and MC and between MS and CO.

The mean slopes for the conjunctions that shared a feature with the fastest search condition were 10.7 ms/item for the faster pair and 17.6 ms/item for the slower pair. For the condition with no shared features, the slope was 13.5 ms/item. The difference between 10.5 and 13.5 ms/item was just significant, $t(17) = 2.197$, $p < .05$. However, the mean differences between the two pairs of slopes with shared features and the two pairs with no shared features were almost identical (8.1 and 8.6 ms). There is little evidence, then, that any one feature determined the possibility of parallel coding, as it might have done if the segregation hypothesis could fully account for the rate of search.

Additivity of feature effects on conjunction slopes. We can go beyond this conclusion and try to throw some light on the conjunction detector hypothesis as well. The results suggest that the two features in each conjunction made independent and additive contributions to the slopes of the search functions. The mean search rates for all 18 subjects (i.e., the negative slope plus twice the positive slope, divided by two) can be predicted to within less than 1 ms on the assumption that when each was a target feature, size contributed 6.5 ms, color 7.5 ms, motion 15.0 ms, and orientation 21.5 ms to the average time required to process each conjunction (see Table 3). Thus, for example, the difference between CO and MO was about the same as the difference between CS and MS, suggesting that motion always contributed about 7.0 ms more than color to the search rates.[3]

The apparent additivity of the contributions from each relevant dimension to the conjunction search rates puts constraints on possible interpretations. First, it seems inconsistent

---

[3] A similar additivity holds for the results of Experiment 2 alone (with estimated contributions of 4.5 ms for size, 7.0 ms for color, 13.0 for motion, and 16.0 ms for orientation), except that the observed slope on movement–orientation conjunctions was too low by 4.6 ms. This was due entirely to 1 subject who gave aberrant results on positive trials in this condition, with a very high intercept (1,200 ms) and a slope of −13 ms per item. For latencies on negative trials only, the additivity was almost perfect in Experiment 2, as it was for the mean of negative and positive trials excluding the one aberrant subject. The smaller estimates of the time contributed for each feature can probably be explained by the extra practice the subjects had in Experiment 2.

Table 3

Mean Slopes for Each Conjunction Target, Together With Predicted Slopes (Pred.), Assuming Additive Contributions From Each Feature Separately

Variable	Color		Size		Motion	
	M	Pred.	M	Pred.	M	Pred.
Size	14.4	14.0				
Motion	21.7	22.5	21.4	21.5		
Orientation	29.5	29.0	27.5	28.0	36.4	36.5

Hypothesized additive contributions to slopes	
Size	6.5
Color	7.5
Motion	15.0
Orientation	21.5

Note. The slopes are taken from the subjects in both Experiment 1 and Experiment 2. They are the mean slopes taken from positive and negative trials assuming a serial, self-terminating scan (0.5 [2 Pos. + Neg.]).

with the hypothesis that holistic conjunction detectors directly code each combination of features. It suggests instead that each dimension is separately processed before the target is found and that each plays an independent role in locating the target or in determining its absence. The simplest account is that subjects check all the elements in every display, either individually or in groups, rejecting each on the basis of whichever feature differentiates it from the target.[4]

The result is surprising in the light of an experiment by Egeth, Virzi, and Garbart (1984). They showed that when only three items shared the target color in displays of 5, 15, or 25 items, search times were about the same for all three display sizes. They inferred that in conjunction search tasks, subjects check serially through only one set of distractors (presumably the smaller set when the numbers are unequal) and not through the whole display. Our apparently additive effects conflict with this conclusion. However, it is possible that the strategy that subjects use for displays like ours, which had equal numbers of each distractor type, differs from the strategy they use for displays with small subsets of distinctive elements. We discuss these results further after a final test of the conjunction detector hypothesis.

Experiment 3: Search for Unknown Conjunction Targets

Experiment 2 explored the idea that a conjunction *target* might be detected by the presence of activity in specialized detectors for its particular conjunction of features. We now test the converse prediction that if the conjunction *distractors* are familiar and highly discriminable, attention might filter them out as such, allowing the target to emerge as the only remaining element. If search depends on identifying and inhibiting known, nontarget conjunctions, it should not be affected by the number of possible targets. Each would be identified simply as differing from both the distractors. The search rate should therefore be the same for any given conjunction target whether its identity was known in advance or whether it could have been any of *n* possible conjunctions.

To explore this possibility, we chose as the distractors for each subject a pair of maximally contrasting three-dimen-

sional stimuli varying in size, color, and orientation. For example, a subject might be given as distractors throughout the experiment the large pink left-tilted bars and the small green right-tilted bars. For each of these pairs of distractor types, there are six possible conjunction targets that recombine the same set of features in different ways. Subjects were asked to search for these six targets, either without knowing which would appear on any given trial or after they were told what the target would be for a given block of trials.

Method

Stimuli. The stimuli in this experiment were the usual rectangular bars, varying in size, color, and orientation, with two values on each dimension. This experiment was run on a different display and the colors and luminances were slightly different from before: the pink had luminance 27.2 cd/m², CIE coordinates .409, .241; the green had luminance 39.4 cd/m², CIE coordinates .246, .518; and the white background had luminance 101.0 cd/m², CIE coordinates .263, .284. The sizes were 0.5° × 1.8° and 0.2° × 1.0° when viewed at a distance of 50 cm. The orientations were 45° and 135°. The displays contained 4, 9, or 16 items, with density controlled, in matrices of 8.7° × 8.7° for the 16 elements and 3.4° × 3.4° for the 4 elements.

With three dimensions defining the stimuli, there are four possible pairs of distractors such that each member of the pair differs from the other on all three dimensions. For each of these pairs there are six possible conjunction targets. Each target differs in one feature from one distractor type and in two features from the other. These four sets of eight stimuli (two distractors with their six possible conjunction targets) were used in different combinations to generate the search displays described below.

Procedure. Each of the four possible pairs of distractors was allocated to 2 of the 8 subjects. The distractors for any subject remained fixed throughout the two sessions of the experiment. In one condition, subjects looked for any of the six conjunction targets—in other words, for any element that differed from both the distractor conjunctions. In each of two sessions, they were given one block of 72 trials for practice, followed by four test blocks of 72 trials each. In six other conditions, they were tested on each known target in turn. They were shown one of the six targets before each block began and were then tested for 72 trials with that target. The first 12 trials were discarded each time as practice. Subjects were tested in both conditions in counterbalanced order in each of two separate 1-hr sessions.

Subjects. We tested 7 women and 1 man from the subject pool; they volunteered and were paid $5 an hour.

Results and Discussion

Figure 5 shows the mean search times in the known and unknown target conditions. There were large and significant

---

[4] An alternative possibility would be that on any given trial, subjects used just one dimension to control the search but that they varied which they used from trial to trial. However, this account requires a somewhat unlikely proviso, namely that the proportion of trials on which the subjects as a group used any given dimension was the same whatever the other dimension with which it was conjoined. Without this assumption, the overall additivity of effects could not hold across the six different pairings of dimensions. This account seems more far-fetched than the alternative, that in general both sets of distractors, each with its own relevant feature, contributed to the slopes on every trial.
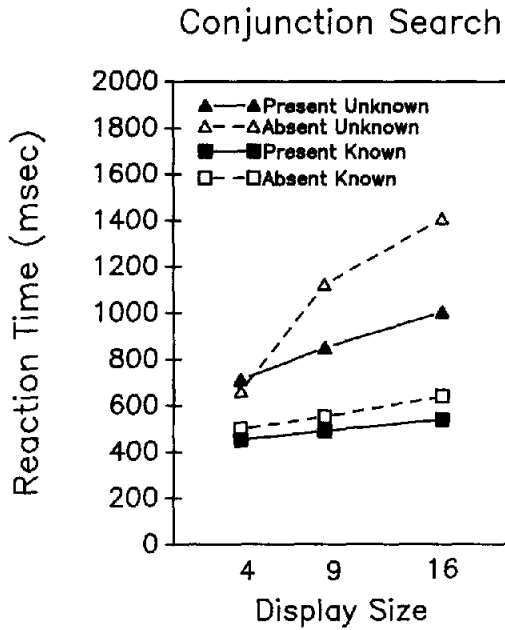
Conjunction Search

Figure 5. Mean search times for known and unknown conjunction targets in Experiment 3.

Table 4
Mean Slopes, Intercepts, and r^2 Measures of Linearity for the Various Conditions of Experiment 3

Feature differentiating target from most similar distractor	Slope	Intercept	r^2	% errors
Known target				
Large size				
Pos.	2.9	394	.979	2
Neg.	3.9	461	.922	2
Small size				
Pos.	8.4	387	.980	4
Neg.	13.6	415	.990	1
Color for large target				
Pos.	3.8	374	.986	2
Neg.	4.0	452	.991	2
Color for small target				
Pos.	6.9	403	.993	2
Neg.	12.3	428	.994	2
Orientation for large target				
Pos.	5.5	487	.999	4
Neg.	11.8	484	.983	4
Orientation for small target				
Pos.	16.6	494	.990	5
Neg.	25.1	463	.998	4
Mean				
Pos.	7.4	423	.999	3
Neg.	11.8	451	.998	3
Unknown target[a]				
Large Size				
Pos.	10.5	607	.996	3
Small size				
Pos.	12.5	609	.789	6
Color for large target				
Pos.	11.2	638	.991	4
Color for small target				
Pos.	22.7	533	.951	4
Orientation for large target				
Pos.	34.6	695	.995	17
Orientation for small target				
Pos.	53.6	648	.986	26
Mean				
Pos.	24.2	622	.996	10
Neg.	60.9	475	.948	4

[a] The negative trials could not be separately assigned to the different targets because these were randomly mixed within blocks. We therefore give only the mean for all negative trials.

differences in the means, $F(1, 7) = 50.44$, $p < .001$, and in their interaction with display size, $F(2, 14) = 15.95$, $p < .001$. Looking separately at the six different targets, we see that there were also large differences between the six target types, both in mean latencies, $F(5, 35) = 43.68$, $p < .001$, and in the search rates (the Display Size × Target Type interaction), $F(10, 70) = 10.20$, $p < .001$. Finally, the three-way interaction between target type, display size, and known vs. unknown target was also significant, $F(10, 70) = 5.06$, $p < .001$, indicating that the differences between the search rates for different targets were much larger when the nature of the target was unknown.

We classified the targets for each subject according to the feature that differentiated each one from the more similar of the two distractor types. Table 4 shows for each type of target the mean slopes, intercepts, error rates, and proportions of the variance with display size that were due to linearity.

We look first at search for each single, known target. Targets that differed only in size from the small distractors (i.e., they were large) and targets that differed only in color from the large distractors were found essentially in parallel. Although the slopes differed significantly from zero, the means were under 6 ms on both positive and negative trials. In this case, no serial check seems to have been required. The targets differing only in orientation from the small distractors were clearly harder to find than the others, and the other three conditions gave intermediate slopes.

When the target was unknown, there was a considerable increase on average both in the intercept of the search function and in the mean slopes and the error rates. However, the increases were much greater for some targets than for others; the fastest targets showed little increase in slope when their identity was unknown, although all showed an increase in intercept.

The results rule out the hypothesis that rapid or parallel conjunction search can be mediated by parallel rejection of distractor *conjunctions*. It would certainly have been a good strategy, if it had been possible, to reject the distractors in parallel by directly sensing their specific conjunctions of features. The two distractor types for any subject were highly discriminable, were known to the subject, and were constant throughout the two sessions of the experiment. However, subjects were apparently unable to filter out the distractor conjunctions as such. We can also rule out a simple serial scanning strategy in which each distractor conjunction is rejected in turn. If subjects had found the target by identifying each distractor as one of the two nontarget conjunctions, there should have been no difference in search rates for the different

unknown targets. At most, there might have been an intercept effect, reflecting differences in the salience of the target and in the subject's confidence once it was found.

A Feature Inhibition Hypothesis

Can the additivity found in Experiment 2 and the range of different search rates observed in Experiment 3 be reconciled with the framework proposed to account for feature integration phenomena in the wider range of tasks outlined in the introduction? The most obvious move is to extend the segregation hypothesis that attention can operate through the feature maps to modulate the activity in the master map of locations. Different strategies may be available, depending on the requirements of the task. When the aim is to segregate and attend to one subgroup of items as a whole, as in the global matching task of Experiment 1 or in a conjunction search task with very unequal proportions of the different distractor types, the strategy may be to inhibit the irrelevant or the larger subset through the feature map that most clearly distinguishes the figure from the ground or the small set of distractors from the large. In natural scenes, small figures are

typically salient relative to their more extended backgrounds. However, when the proportions of the different distractor elements are more equal and only a single known target is relevant, a more efficient strategy may be to inhibit both sets of distractors in relation to the target, controlling their salience through each of the two feature maps on which they differ from the target (see Figure 6). If there is some additional cost to controlling inhibition from each additional feature map, there could be a trade-off between the number of items to be searched when only one set is inhibited and the number of feature maps controlling the inhibition when more than one set is inhibited. Consider, for example, a display containing a vertical red target among vertical green and horizontal red distractors. When there are very few red items (as in Egeth et al., 1984), the best strategy may be to inhibit only locations containing green items and to search the remainder. But when there are equal numbers of red and green items, it may be more efficient to inhibit all the distractor locations as effectively as possible by using both color (green) and orientation (horizontal) to define the unwanted elements. The red vertical target will then be the sole survivor on the battlefield, or the least affected if the inhibition is only partially effective.
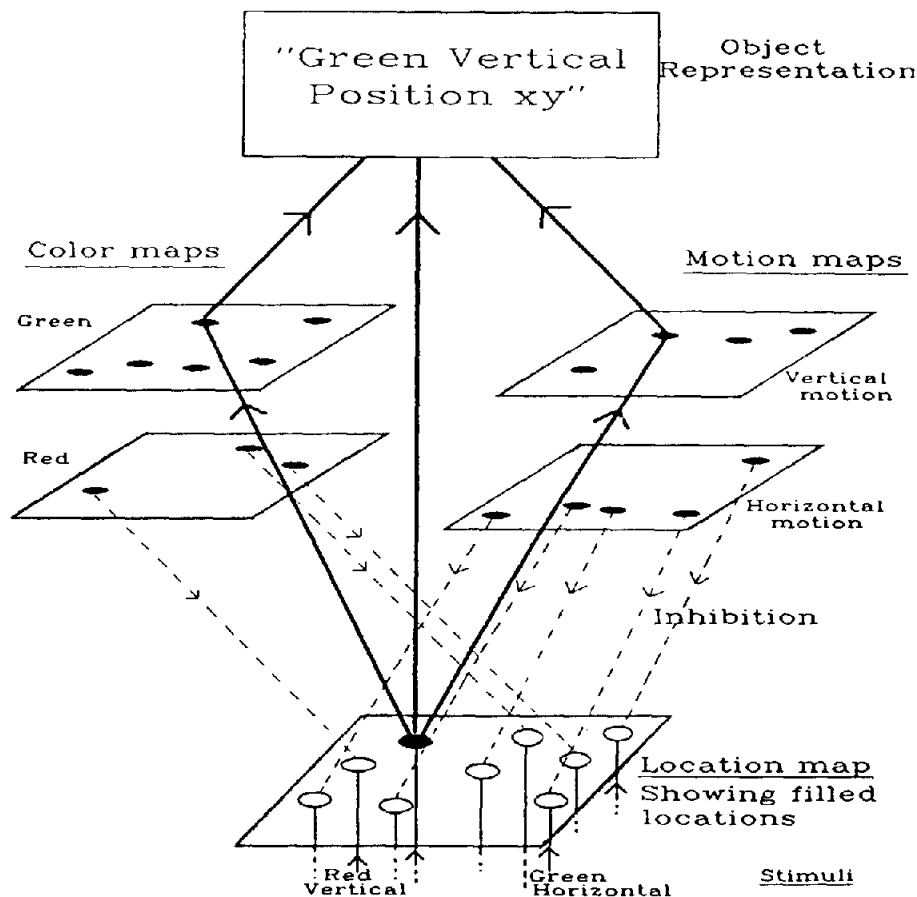

Figure 6. Modified model showing inhibition from both distractor feature maps. To simplify the figure, we show complete inhibition of each distractor location, with activation passed on only from the target. If inhibition were only partial, some partial activation would be passed back from the distractor locations through the feature maps to the object level.

What happens in the case when the target is not known in advance? Experiment 3 showed that the slope for the fastest target was unaffected by advance knowledge of what the target would be, whereas the slopes for the slowest were dramatically steeper and the error rates extremely high. This difference rules out a single serial scan through the distractors and suggests instead that subjects also used the feature-based inhibition strategy to control search for the unknown targets. However, in this condition no single pair of feature maps could reliably make the target salient. Instead, the strategy would be to apply feature inhibition successively with different pairs of distractor feature maps and to see, for each pair, if any location remained unaffected. Any uninhibited but filled location should contain the target. Thus, a series of "fishing expeditions" could be tried: For example, if the distractors were large pink 45° bars and small green 135° bars, the first pass might filter out pink elements and 135° elements and test if anything remained; if so, it must be a target (a large or small green 45° bar). If this attempt failed, the second pass might reject green elements and large elements; if anything remained, it must be a target (a 45° or 135° small pink bar). A minimum of three such passes would be necessary and sufficient to exhaust the possible targets. (The third in this example would reject small elements and 45° elements; if anything remained, it must be a target—a pink or green large 135° bar.)

If the fishing expeditions are tried in a consistent order, this strategy predicts that some targets should be found much more rapidly than others. The first should be found as quickly as in the known condition, but the last very much more slowly. Indeed, this is what we found. The targets that differed only in size from the small distractors and those that differed only in color from the large distractors were found almost as fast in the unknown as in the known target condition. On the other hand, the two conditions involving orientation were very much harder in the unknown than in the known condition, consistent with the difficulty of orientation conjunctions found in Experiment 2.[5] The asymmetry between large and small items throughout the data suggests that it may be harder to inhibit locations containing items that are naturally more salient, in this case large rather than small. (Similar asymmetries were reported by Treisman & Gormican, 1988.)

The strategy of inhibiting different feature values successively and checking for each target in turn allows more detailed predictions relating the slopes in the unknown target condition to those in the known condition. The simplest hypothesis is that a serial self-terminating check is made for each of the six targets at the same rate as in the known target condition. The slope for the first unknown target would then be the same as for the same target in the known condition; the slope for the second target would be the sum of the negative slope for the first target in the known condition and the positive slope for the second; the slope for the third target would be the sum of the negative slopes for the first two targets and the positive slope for the third, and so on.

Different subjects might search for the targets in different orders, so we calculated the expected rates separately for each individual; we took the order of search in the unknown condition from a rank-ordering of the observed slopes for that subject in the unknown condition. The first two columns in Table 5 show the mean observed slopes in the unknown target condition, rank-ordered for each subject separately, and the mean predicted slope, estimated by summing the corresponding slopes in his or her known target conditions.

The observed slopes match the predicted slopes quite well for the four targets that gave the fastest search rates, but are lower than predicted for the two slowest targets and for the negatives. However, this undershoot can be explained by the large numbers of missed targets in these conditions. With displays of 16 items, subjects missed 28% of targets that differed only in orientation from the most similar distractor. It is likely that subjects cut short the search on the later passes through the display before they had checked every item. Consistent with this hypothesis, the mean error reaction times were about the same as the mean correct times within the same conditions (averaging 1,121 ms compared with 1,105 ms).

If the search times do reflect premature termination on the more difficult trials, we can correct the observed slopes as follows: We assume that the proportion of the display that is actually searched is given by the proportion of detected targets on positive trials. The proportion is smaller on unknown than on known target trials, indicating that subjects search fewer items on each pass through the display with an unknown target than they do in the single pass for a known target. We used the observed errors in the known and unknown conditions for each target type to correct the slopes so as to reflect the number of items actually searched. The percent detections at each display size were used to estimate the effective display size (the actual number searched in each condition). The ratio of the slopes for items actually searched in the two conditions was then calculated. For example, if subjects detected all the 4, 9, or 16 targets in the known condition but only 90% of the targets on displays of 4, 80% on displays of 9, and 70% on displays of 16 in the unknown target condition, we assume that they searched $.9 \times 4 = 3.6$, $.8 \times 9 = 7.2$, and $.7 \times 16 = 11.2$ items in the unknown condition, and all 4, 9, and 16 in the known condition. The ratio of the slopes would then be .63 (i.e., the regression of the number of items searched in the unknown condition on the number searched in the known condition). We can then predict the slopes in the unknown condition, taking the calculated ratio of unknown to known slopes for each type of target. The corrected predictions are shown in the third column of Table 5. The fit is now quite close, with a correlation of .995 and with no systematic differences between the observed and the predicted slopes.

The results of this last experiment give strong support to the feature inhibition account. They confirm that conjunctions cannot be preattentively identified as such, even though the displays consistently contained the same two highly discriminable and familiar conjunctions. Parallel coding of these distractor conjunctions should generate the same search functions for the unknown as for the known targets. They also

---

[5] It is possible that the canonical orientations, horizontal and vertical, would be more easily processed than the 45° orientation used here. This should be checked in further research.

Table 5
Mean Observed and Predicted Slopes in the Unknown Target Condition

Trial	Observed	Predicted	Predictions corrected for errors
Positive	5.3	4.5	4.1
	9.7	10.3	9.7
	15.9	19.1	18.3
	23.0	25.7	24.4
	35.5	42.2	37.8
	55.6	61.1	51.5
Negative	60.8	70.4	58.1

Note. The observed slopes are rank-ordered for each subject separately.

rule out the original simpler version of feature integration theory, which predicted a single, serial scan of the distractors. If this strategy had been used, the search rates for the different unknown targets should all have been the same. Instead, subjects appear to have organized their search through a series of restructurings of the display, inhibiting different pairs of features and scanning the remaining active locations until the target emerged.

Wolfe et al. (1989) recently proposed a similar model to account for the flat functions they obtained in search for conjunctions of highly discriminable features. Wolfe et al. retain the idea of feature maps that are functionally separated from a master map of locations. They suggest that each of the target feature maps can activate the locations that contain their particular feature, producing double excitation at the location of the target. The serial scan with the attention spotlight is then directed to locations in order of their current level of activation. The higher the level of background "noise" in the signal from the feature maps to the location map, the longer it will take to locate the target. When the activation is strong enough, either only a few or no distractors compete for attention in the serial scan that specifies conjunctions of features. Wolfe et al. point out that their model could equally well be stated in terms of inhibition of distractor feature locations and suggest that the two alternative accounts are very hard to distinguish from each other.

In their article, they report some additional empirical data that are consistent with both accounts. (a) They replicated the finding by Bergen and Julesz (1983) that search for a T among Ls (in four different orientations) is serial, and confirmed that this was the case even for subjects who showed parallel search with conjunctions of color, size, and shape. This is consistent with the claim that parallel search depends on control from separate feature maps. Ts and Ls differ only in the spatial arrangement of the same two oriented lines; thus, neither orientation on its own can be used to distinguish the target from any distractor. (b) They also confirmed an earlier finding by Quinlan and Humphreys (1987) that search for triple conjunction targets differing from each distractor in two of their features was faster than search for the same targets among distractors that differed from them only in a single feature (see also Dehaene, 1989). Either activation or inhibition from each of two feature maps may combine to increase

the relative activation of the target location more effectively than activation or inhibition from a single feature map.

We had independently also compared conjunction search with targets differing in either one or two features from the distractors, using the same motion, color, and orientation features as in Experiments 1 and 2. We obtained a similar large improvement in search when the target differed in two features from each distractor rather than one. The mean slopes were 12.4 and 27.0 ms per item for positive and negative trials with two features different, compared with 46.0 and 83.4 ms per item with one feature different. Interestingly, the search rates for the triple conjunction with one feature different were significantly slower than the slopes for each of the corresponding double conjunctions (MC, MO, and CO) from Experiments 1 and 2, although these also differed in only one feature from the distractors. Our results suggest a cost due either to the increased heterogeneity of the distractors or to the increase in the number of target features to be conjoined. If feature inhibition is used to help the search, it may be more difficult to implement when three separate feature maps are involved than when only two are.

The search rates we obtained for the triple conjunction with two features different were about the same as those for the fastest relevant double conjunction of Experiment 1 (MC). Note that the double conjunction stimuli (MC, MO, and CO) are contained within the triple conjunction stimuli (MCO). Subjects could use inhibition from the motion and color feature maps to help them detect the triple conjunction targets, just as they did for the MC double conjunction targets. Inhibition from two features would be enough to remove all the distractor activity when the triple conjunction target differed in two features from each distractor. Hence, there was no reason to expect a difference in search rates.

Experiment 4: Inhibition or Activation?

Our account so far has made one arbitrary choice where Wolfe et al. (1989) made another: We have couched the theory in terms of inhibition from feature maps characterizing the distractors, whereas they talked primarily of activation from feature maps characterizing the target. The data reported so far do not differentiate between these two versions of the model, and, of course, it is quite possible that both are used. Is there any way of distinguishing the two accounts? If inhibition is used, the task might become more difficult and costly the more different features are involved. This could explain the greater difficulty of triple than of double conjunctions when both differ in one feature from the distractors. However, there are other possible explanations for that result. A better test would pit a prediction from feature-based inhibition directly against the prediction from activation. This could be achieved by increasing the number of distractor features while at the same time making the added features more discriminable from the target. The more similar the features of the target are to those of the distractors, the more one would expect any activation of target features to spread to the distractor features, thus reducing the efficiency of the activa-

tion strategy. On the other hand, the larger the number of different distractor types present in the display, the less efficient the inhibition strategy might become, even when the added distractors are more discriminable from the target.

The next experiment compared search with one and search with two values defining sets of distractors on each relevant dimension. We compared search for the same conjunction target (a green bar at 27°) in either the standard condition with two distractor types (green bars at 63° and gray bars at 27°) or a new condition in which the same two distractor types were mixed with two others with less similar features (*pink* bars at 27° and green bars at 90°). We chose features (gray and 63°) for the standard two-distractor conditions that would be clearly more similar to the 27° green target than the added features in the four-distractor condition (90° and pink).

Method

Stimuli. The displays were the same as the color-orientation displays in Experiment 3, with the following exceptions: Two new orientations, 27° and 63° from the horizontal, were used instead of 45° and 135°; the largest displays subtended 8.7° and the smallest 3.4°, with stimulus bars subtending 0.3° × 1.5°; and the stimuli were presented on a black rather than a white background with the following luminances and colors: pink luminance 29.6 cd/m², CIE coordinates .436, .282; green luminance 65.1 cd/m², CIE coordinates .292, .523; gray luminance 51.5 cd/m², CIE coordinates .260, .306.

Procedure. The target was always a green bar oriented at 27° from the horizontal. In the two-distractor condition, the distractors were green bars at 63° and gray bars at 27°. In the four-distractor condition, half of these distractors were replaced by equal numbers of green bars at 90° (vertical) and pink bars at 27°. There were, as usual, three display sizes, 4, 9, and 16, with density controlled. Subjects were given seven blocks of 48 trials in each of the two conditions in counterbalanced order in a single 1-hr session. The first block in each condition was practice and was not included in the analysis.

Subjects. We tested 4 men and 6 women from the volunteer subject pool.

Results and Discussion

The mean search latencies in each condition are shown in Figure 7. An ANOVA showed significant effects of condition, $F(1, 9) = 22.31, p < .01$, of display size, $F(2, 18) = 41.60, p < .001$, of positive versus negative trials, $F(1, 9) = 19.89, p < .01$, and significant interactions of Condition × Display Size, $F(2, 18) = 17.65, p < .001$, and of Positive vs. Negative Trials × Display Size, $F(2, 18) = 13.06, p < .001$. The slopes were 17.1 ms and 30.9 ms per item on positive and negative trials with displays containing two distractor types, and 23.3 ms and 40.7 ms per item with displays containing four distractor types. Error rates were correlated with mean reaction times: Subjects missed 2%, 3%, and 5% targets in displays of 4, 9, and 16 items respectively when there were two distractor types, and 3%, 7%, and 10% targets when there were four distractor types.

A similar detrimental effect of distractor heterogeneity has also been shown by Farmer and Taylor (1980), Bundesen and Pedersen (1983), and McIntyre, Fox, and Neale (1970).
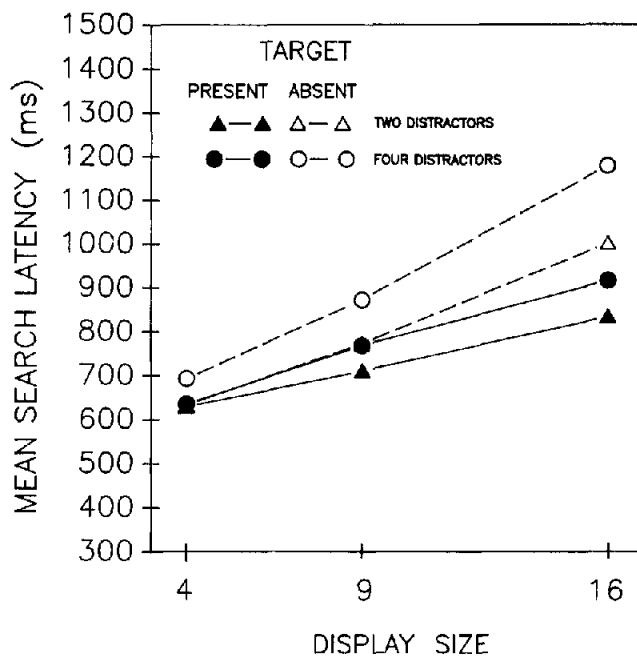

Figure 7. Mean search times with two and with four types of distractors in Experiment 4.

Farmer and Taylor (1980) and Bundesen and Pedersen (1983) varied the number of different distractor colors presented in search for a color target. However, they did not compare the effect of replacing some similar color distractors with some that differed more from the target in the same direction. Increasing heterogeneity was therefore confounded either with an increase in the number of potentially confusable colors or with an increase in the directions in color space in which a discrimination had to be made. McIntyre et al. (1970) varied the similarity of distractor letters to target letters. Comparing across two of their experiments, it seems that increasing heterogeneity by adding less similar letters (e.g., Os or Us to a display containing a target F or T among distractor Is) led to a decrease in accuracy. The results are consistent with ours, and suggest that rejecting varied distractors is more difficult than rejecting homogeneous distractors, even when the latter are on average more similar to the target.

Our aim in Experiment 4 was to test whether search is more likely to be facilitated by activation of locations containing target features or by inhibition of locations containing distractor features. If selection depended solely on activating target features, the four feature displays should be searched at least as fast as the two-feature displays, because the extra two features were clearly less similar to the target than the first two. If anything, performance should have improved when half the original distractors were replaced by more discriminable ones, because their locations would receive less spreading activation from the target feature maps. In fact, performance was significantly worse with the four-feature displays, suggesting a process of active inhibition that was more difficult to implement when more different features were involved.

An alternative account is that distractor heterogeneity interferes with search simply because it creates additional boundaries or gradients that attract attention (Julesz, 1984). In this case, variation on both relevant and irrelevant dimensions should be detrimental. On the other hand, if heterogeneity impairs search by making feature-controlled inhibition more costly, it should only do so on dimensions that distinguish the distractors from the target. Treisman (1988) found that variation on irrelevant dimensions had little or no effect on search for feature-defined targets. Again, this result is consistent with the idea that distractor heterogeneity is detrimental primarily when it makes it more difficult or more costly to filter out nontarget features.

Duncan and Humphreys (1989) have recently proposed that a combination of dissimilarity between distractors and similarity between the target and the distractors can account for all the variance in search performance. Distractor differences, on their account, impair search by reducing subjects' ability to group the distractors and to reject them at a more global level. Our account of distractor heterogeneity is consistent with theirs. However, some recent data from experiments that control both forms of similarity suggest that the need to conjoin features does add a further component to the difficulty of search (Treisman, 1990a).

General Discussion

Summary of Results

The main findings in this series of experiments were as follows: (a) We confirmed the results of Nakayama and Silverman (1986b), Wolfe et al. (1989), and others, which showed that search for conjunction targets can be fast, and in some cases parallel, when the features are highly discriminable. In our data, conjunctions involving size gave the fastest search rates, those involving color were next, motion third, and those involving orientation were typically quite slow. The rank order could, of course, change if the discriminability on any dimension were reduced. (b) There was a strong correlation between the ease of conjunction search and the ease of segregating the same displays to allow the perception of global boundaries. (c) Each feature appeared to make an additive contribution to the time required to scan the display, suggesting that the search process operates at the level of separate features rather than conjunctions. The additivity also implies that when the display contains equal numbers of each type of distractor, both sets may be checked. When one set is much smaller than the other, as in Egeth et al. (1984), a more selective strategy may be followed, segregating the smaller set and scanning only that. (d) Known targets were found more quickly on average than unknown targets. When the targets were unknown, some showed little change in search rate (slope), whereas others showed a substantial increase, both in slope and in errors (missed targets). The search rates for the unknown targets could be predicted by summing a sequence of rates for the known targets, as if they were found through a serial check for each possible target in turn. (e) Finally, it was more difficult to find a conjunction target among four

different types of distractors than among two, even when the extra two distractors were more discriminable from the target than those they replaced. Thus, distractor heterogeneity on the target-defining dimensions makes selection more difficult. We also found slower search for a triple compared with a double conjunction target when both differed only in one feature from each type of distractor. The increased latencies here could also be due to distractor heterogeneity, because there were three distractor types for the triple conjunction and only two for the double conjunction.

The Conjunction Detector Account

We considered three possible accounts of the data. The simplest was that certain conjunctions are directly coded in parallel by specialized detectors tuned to respond to particular combinations of values on different perceptual dimensions. Our data raise some difficulties for this account, however. The first is that the correspondence with physiological evidence is weak: The conjunctions of orientation with size and with motion, for which the physiological evidence of direct neural coding is strongest, are those that are hardest to detect in search tasks. This objection is not conclusive because the functional interpretation of neural single cell activity is still unknown.

A second problem for an account based on conjunction detectors is the fact that subjects seem unable to find an unknown target by coding and rejecting two known distractor conjunctions in parallel, even when these distractors are highly discriminable and constant throughout all the conditions in two sessions of search (as in Experiment 3). On the other hand, when two different distractor types differ from the target in four different *features*, they can be easily filtered out in parallel (e.g., the brown Ts and green Xs with targets "blue" or "S" in Treisman & Gelade, 1980, Experiment 1). The present Experiment 3 shows that the same efficient selection process is impossible for two conjunction distractors, suggesting that they are coded differently from the separate features.

Thirdly, the search rates for known conjunctions can be predicted by assuming additive contributions from each component feature (see Experiment 2). The natural inference is that each is separately processed, even in conjunction search.

A final consideration is that the conjunction detector account conflicts with the various findings in other paradigms that originally prompted the development of the feature integration model. We need an account that is consistent with (a) the large advantage of precuing the target location when conjunctions are involved, (b) the occurrence of illusory conjunctions (recently confirmed with the present highly discriminable features, Treisman, 1990b), and (c) the dependence of conjunction identification on accurate localization. Finally, the hypothesis of direct conjunction coding leaves unexplained the observed continuum of difficulty, both for conjunctions on different dimensions and for conjunctions that differ in the discriminability of the relevant values on any given pair of dimensions. There seems to be no clear dichotomy between conjunctions coded in parallel and conjunctions coded serially; instead we find a range of search rates, depend-

ing both on which dimensions are paired and on the discriminability of the values tested on each of those dimensions.

None of these objections rules out a direct coding hypothesis for some conjunctions of features. However, taken together with the constraints imposed by the potential combinatorial explosion, they suggest that it is worth considering alternative special strategies for visual search tasks with conjunction targets, strategies that could be compatible with the original feature integration hypothesis.

The Segregation Hypothesis

We explored two such strategies—the segregation strategy and the feature inhibition strategy. Both share the assumption of the original feature integration theory that perceived conjunctions are formed by sequentially linking separate features through a serial scan of a shared map of locations. Both suggest additional ways in which a known conjunction target could be found in visual search without any parallel coding of the other conjunctions in the display. Like the original theory, both link feature integration to spatial attention. They differ from it and from each other only in the additional mechanism for controlling the spatial selection of potential targets.

The segregation account combines the idea proposed by Egeth et al. (1984) that attention can be narrowed on the basis of one feature to exclude one set of distractors, with the idea that a parallel feature search within the remaining subset might then become possible (Treisman, 1982). We suggested that the attentional segregation could be achieved by inhibition from the feature map coding a salient nontarget feature, resulting in reduced activation in the locations in the master map that currently contain that feature. The remaining distractors are then scanned in parallel for the unique feature that characterizes the target.

We tested whether segregation was controlled by the same variables that allowed parallel search for conjunctions and found a similar ordering of difficulty across different displays, as if the two tasks did depend on some shared mechanism. However, the more extensive testing in Experiment 2 revealed an apparent additivity of feature effects on the conjunction search functions, suggesting that at least with equally divided displays, both features contribute independently to the search latencies.

The Feature Inhibition Hypothesis

The additivity is more consistent with a third possible strategy for search, the feature inhibition strategy. This differs from the segregation strategy only in allowing inhibition from more than one separate feature map. Rather than removing just one set of distractors from the search process and searching the other set in parallel, feature inhibition could be generated in two or more feature maps coding nontarget features, thus reducing the activity in all distractor locations. At the extreme, with sufficiently distinct and separable features, it might eliminate the activity generated in the master

map by distractor elements, allowing the target to pop out equally well whatever the display size. When the inhibition is incomplete however, we assume that a serial scan is made through the master map of locations, in which locations differ only in their level of activation. The order in which the locations are scanned (although not their size) must be independent of the features they contain in order to give linear slopes with a two-to-one ratio of target absent to target present trials. We suggest that the order is determined by spatial adjacency either of groups or of individual items.

Is feature inhibition sufficient to explain all the results without also postulating a serial scan? We think the results are best explained by a combination of the two. Feature inhibition alone does not account for (a) the range of slopes, from shallow to very steep, that vary continuously with feature discriminability but remain linear throughout; (b) the two-to-one slope ratios that are generally obtained; and (c) the elimination of the slopes when attention is cued in advance to the location of the target (Treisman, 1988). Certainly other models are possible to explain the linear functions (See Townsend, 1971). Further research using other converging operations will be needed to settle the issue; for the present, our hypothesis is simply a hypothesis, one attempt to account for all the data presently available.

How then do we explain the range of slopes we and others have obtained in conjunction search tasks and the additivity of feature effects found in the present experiments? In relating search rates to feature discriminability, Treisman and Gormican (1988) suggested that shallower slopes may reflect search through subgroups, checking items within groups in parallel. Instead of attending to one item at a time, we attend to pairs, triplets, or even larger groups. According to the theory, the level at which features are assembled to form object representations receives a pooled response from each feature map, reflecting the activation produced by whatever stimuli are currently within the attention window, together with their location. The pooled response from each map allows an assessment of the likelihood that the particular feature coded by that map is present in the attended area. It is higher the more instances of the feature are included in the area, and lower the more inhibited their master-map locations have been. Inhibition from nontarget feature maps reduces the response not only from their own nontarget features but also from any target features that share the same master-map locations. The more distinctive the target feature, (i.e., the less its feature map responds also to the distractor features), the more diagnostic of the target a given pooled response will be.

In applying the group-processing model to search for conjunction targets, we face the additional constraint of avoiding illusory conjunctions. If the attention window encompasses examples of both distractor types, then both target features will be passed on to the object level at which conjunction targets are identified. To avoid illusory conjunctions, we assume that some criterion level of response must be simultaneously reached for each of the target features before the subject decides that the conjunction target is present. The more effective the feature-based inhibition on a particular dimension, the larger the number of elements sharing the

nontarget feature that can be attended together without the pooled response to their target feature exceeding the criterion for a positive response. For example, suppose that color is an effective dimension for feature inhibition and orientation is not. If a subject is looking for a pink 45° target, master-map locations containing green 45° distractors will be strongly inhibited, whereas locations containing pink 135° distractors will only be slightly inhibited. This might produce a pooled orientation response to *two* inhibited green 45° distractors that is nevertheless below the response to a *single* uninhibited pink 45° target. On the other hand, the pooled color response to one inhibited pink 135° distractor might be only a little below the response to the single uninhibited pink 45° target.

The strategy then might be to adjust the attention window on-line to take in groups of elements whose summed activation on each target feature was below that expected for the target by some criterion amount. A systematic scan through master-map locations would take in varying numbers of adjacent elements, adjusting the size of the aperture until the pooled feature activation summed to some fixed criterion level. If a local area happened to contain only strongly inhibited green 45° targets, the attention window would pool the response to several at a time; if it contained only pink 135° distractors, it would be narrowed to take only one or two at a time, and if it contained both types of distractors, the attention window would typically include at most one pink 135° element with one green 45° element.

This strategy would explain the additive effect of each separate dimension on the slopes of search latencies. The more discriminable the feature, the more effective the inhibition and the greater the number of distractors sharing that feature that could be rejected in parallel. Our results suggest a contribution to conjunction slopes of 7.5 ms for the color dimension; because half the display shared the target color, this is equivalent to 15 ms for each item that differed in color from the target. If pairs were checked in parallel, the rate would be equivalent to 30 ms per pair; if triplets, the rate would be equivalent to 45 ms for each. Similar inferences can be made for the other three dimensions.

The feature inhibition hypothesis is similar to one proposed by Wolfe et al. (1989) and, in more general terms, to the two-stage model of Hoffman (1979). It is consistent with the evidence from Bergen and Julesz (1983) and from Wolfe et al. that search is serial for a conjunction of the same two features in different spatial arrangements (e.g., Ts among Ls in four randomly varying orientations). If we assume that Ts and Ls are both composed of one horizontal and one vertical line, then neither has any unique feature through which inhibition could be controlled, so that item-by-item search is required. The hypothesis is also consistent with the finding by Quinlan and Humphreys (1987, replicated by Wolfe et al., Dehaene, 1989, and by us) that distractors differing in two of their features from a triple conjunction target are rejected more efficiently than distractors differing only in one.

Wolfe et al. (1989) attribute the rapid search rates they obtained with conjunctions of highly discriminable features to a reduction in the number of distractors that are checked before the target is found. In their model, the distractors are checked in an order determined by their level of activation,

starting with the most active location, which is presumably the most likely to contain the target. If the background noise is high relative to the top-down feature-based control of activation, several distractors may be checked before the target is located. The model with this second assumption does not naturally predict two-to-one slope ratios: The target with its high level of activation should on average be found earlier than half way through the display. Yet the data suggest that when conjunction slope ratios deviate from the two-to-one pattern, they are more likely to approximate equal slopes than ratios larger than two to one, at least for small display sizes (Pashler, 1987).

The feature inhibition hypothesis makes another prediction that may differentiate it from accounts based solely on top-down feature activation. By keeping the target features constant, we showed that search was impaired rather than helped when we replaced half the distractors by others that differed from the target on the same two dimensions but to a greater degree. If the search strategy had been to preactivate the features characterizing the target, the greater discriminability of the new distractors should, if anything, have reduced the interference they caused. Any model in which search is guided only by top-down activation of target features should have difficulties with this result.

Extensions to Other Experimental Paradigms

Can we apply the feature inhibition account to search for targets defined only by the presence or absence of a single feature (Treisman & Gormican, 1988; Treisman & Souther, 1985)? The results from these tasks may in fact help us to select between two possible versions of the feature inhibition model. In one experiment, we found that whereas search for a circle with a slash among circles without slashes was parallel, search for a circle without a slash among circles with slashes showed strong effects of display size. If inhibition could be directed to master-map locations that contained a slash, the one circle without a slash should be detected as the only element remaining unscathed.

To explain the difficulty of search tasks in which the target lacks a feature that is present in all the distractors, we must assume that the locations that get inhibited are not the global areas in which patterned elements are located, but rather the specific points occupied by the inhibited features. For separate dimensions like color, size, and orientation, exactly the same set of points can be occupied by each different feature; for different parts of a shape, this is not the case. The slash that intersects a circle occupies a different set of points from the circle itself. Inhibiting the locations of the distractor slashes would then leave the distractor circles intact and indistinguishable from the target circle without a slash. On the other hand, when the target is the one circle with a slash among distractor circles without slashes, inhibiting the circles would eliminate the distractors completely, leaving the target slash to signal the presence of the target.

Finally, we may note that the debate over whether attention is controlled by inhibiting or filtering out unwanted signals or by activating attended signals goes back a long way. Within the early selective-listening paradigm, a related result was

obtained with auditory speech messages (Treisman, 1964). Selective listening to a message on the right ear was impaired more by a message in the left ear and one in both ears than by two messages in both ears. Attention seemed to "filter out" unwanted stimuli (Broadbent, 1958) or to "attenuate" their effects (Treisman, 1960) rather than to move one or more auditory windows ("mental microphones"?) to selected items. The filter analogy suggests that in the absence of attention, all the features present in the scene are automatically registered and perhaps tend to form all their possible perceptual conjunctions. Attention, according to this view, is needed to exclude irrelevant features from the level at which the representations of objects are assembled.

We have presented an account based on inhibition rather than activation. Admittedly, the evidence distinguishing the two is still quite scanty, and an activation account may do equally well with most of the data. It is also quite possible that both play a role. Cave and Wolfe (1990) propose a second factor—variations in bottom-up activation that depend on interdistractor differences—that could also account for our result. Both accounts are consistent with the general hypotheses about feature integration that emerged from converging results in a variety of other experimental paradigms. Finally, either could subserve some more generally useful functions in everyday perception: They could guide search for predetermined targets, group the separate parts of partially occluded objects, and allow figure-ground segregation with the concomitant emergence of boundaries to global groups of elements sharing common values on different perceptual dimensions.

References

Albright, T. D. (1984). Directions and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology, 52,* 1106–1130.

Bashinski, H. S., & Bacharach, V. R. (1980). Enhancement of perceptual sensitivity as the result of selectively attending to spatial locations. *Perception & Psychophysics, 28,* 241–248.

Bergen, J. R., & Julesz, B. (1983). Parallel versus serial processing in rapid pattern discrimination. *Nature, 303,* 696–698.

Bjork, E. L., & Murray, J. T. (1977). On the nature of input channels in visual processing. *Psychological Review, 84,* 472–484.

Braddick, O., Campbell, F. W., & Atkinson, J. (1978). Channels in vision: Basic aspects. In R. Held, H. W. Leibowitz, & H. L. Teuber (Eds.), *Handbook of sensory physiology* (Vol. 8, pp. 3–38). New York: Springer Publishing.

Broadbent, D. E. (1958). *Perception and communication.* New York: Pergamon Press.

Bundesen, C., & Pedersen, L. F. (1983). Color segregation and visual search. *Perception & Psychophysics, 33,* 487–493.

Cavanagh, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity, and shape. *Computer Vision, Graphics and Image Processing, 37,* 171–195.

Cavanagh, P., Tyler, C. W., & Favreau, O. E. (1984). Perceived velocity of moving chromatic gratings. *Journal of the Optical Society of America, 1,* 893–899.

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

Crick, F. (1984). Function of the Thalamic reticular complex: The searchlight hypothesis. *Proceedings of the National Academy of Sciences, 81,* 4586–4590.

Dehaene, S. (1989). Discriminability and dimensionality effects in visual search for featural conjunctions: A functional pop-out. *Perception & Psychophysics, 45,* 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.

De Valois, R. L., Albrecht, D. G., & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research, 22,* 545–559.

De Valois, R. L., Yund, E. W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research, 22,* 531–544.

De Yoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences, 11,* 219–226.

Downing, C. J., & Pinker, S. (1985). The spatial structure of visual attention. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance XI* (pp. 171–187), Hillsdale, NJ: Erlbaum.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Egeth, H., 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. (in press). Three-dimensional features that pop out in visual search. In D. Brogan (Ed.), *Visual search.*

Enns, J. T., & Rensink, R. A. (1990). Influence of scene-based properties on visual search. *Science, 247,* 721–723.

Eriksen, C. W. W., & Hoffman, J. E. (1972). Temporal and spatial characteristics of selective encoding from visual displays. *Perception & Psychophysics, 12,* 201–204.

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.

Graham, N. (1985). Detection and identification of near-threshold visual patterns. *Journal of the Optical Society of America, 2,* 1468–1482.

Hoffman, J. E. (1979). A two-stage model of visual search. *Perception & Psychophysics, 25,* 319–327.

Houck, M. R., & Hoffman, J. E. (1985). Conjunction of color and form without attention: Evidence from an orientation-contingent color after effect. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 186–199.

Hubel, D. H., & Livingstone, M. S. (1987). Segregation of form, color and stereopsis in primate area 18. *Journal of Neuroscience, 7,* 3378–3415.

Intraub, H. (1985). Visual dissociation: An illusory conjunction of pictures and forms. *Journal of Experimental Psychology: Human Perception and Performance, 11,* 431–442.

Julesz, B. (1971). *Foundations of cyclopean perception.* Chicago: University of Chicago Press.

Julesz, B. (1984). Toward an axiomatic theory of preattentive vision. In G. M. Edelman, W. E. Gall, & W. M. Cowan (Eds.), *Dynamic aspects of neocortical function* (pp. 558–612). New York: Wiley.

Julesz, B. (1986). Texton gradients: The texton theory revisited. *Biological Cybernetics, 54,* 464–469.

Keele, S. W., Cohen, A., Ivry, R., Liotti, M., & Yee, P. (1988). Tests of a temporal theory of attentional binding. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 444–452.

Lawrence, D. H. (1971). Two studies of visual search for word targets with controlled rates of presentation. *Perception & Psychophysics, 10,* 85–89.

Livingstone, M. S., & Hubel, D. H. (1987). Psychological evidence for separate channels for the perception of form, color, movement

and depth. *Journal of Neuroscience, 7,* 3416-3468.

Maunsell, J. H. R., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10,* 363-401.

Maunsell, J. H. R., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey: I. Selectivity for stimulus directions, speed and orientation. *Journal of Neurophysiology, 49,* 1127-1147.

McIntyre, C., Fox, R., & Neale, J. (1970). Effects of noise similarity and redundancy on the information processed from brief visual displays. *Perception & Psychophysics, 7,* 328-332.

McLean, J. P., Broadbent, D. E., & Broadbent, M. H. P. (1982). Combining attributes in rapid serial visual presentation. *Quarterly Journal of Experimental Psychology, 35A,* 171-186.

McLeod, P., Driver, J., & Crisp, J. (1988). Visual search for a conjunction of movement and form is parallel. *Nature (London), 332,* 154-155.

Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review, 81,* 521-535.

Minsky, M. (1961). Steps towards artificial intelligence. *Proceedings of the Institute of Radio Engineers, 49,* 8-30.

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). New York: Cambridge University Press.

Nakayama, K. & Silverman, G. H. (1986a). Serial and parallel encoding of visual feature conjunctions. *Investigative Ophthalmology and Visual Science, 27*(Suppl. 182).

Nakayama, K., & Silverman, G. H. (1986b). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264-265.

Pashler, H. (1987). Detecting conjunctions of color and form: Reassessing the serial search hypothesis. *Perception & Psychophysics, 41,* 191-201.

Pomerantz, J. R., Sager, L. C., & Stoever, R. G. (1977). Perception of wholes and their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance, 3,* 422-435.

Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109,* 160-174.

Prinzmetal, W. (1981). Principles of feature integration in visual perception. *Perception & Psychophysics, 30,* 330-340.

Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combination of color, shape and size: An examination of the task constraints on feature and conjunction searches. *Perception & Psychophysics, 41,* 455-472.

Ramachandran, V. (1988). Perceiving shape from shading. *Scientific American, 259,* 76-83.

Regan, D., Beverley, K. I., & Cynader, M. (1979). The visual perception of motion in depth. *Scientific American, 241,* 136-151.

Steinman, S. B. (1987). Serial and parallel search in pattern vision. *Perception, 16,* 389-399.

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

Thorell, L. G., De Valois, R., & Albrecht, D. G. (1984). Spatial mapping of monkey V1 cells with pure color and luminance stimuli. *Vision Research, 24,* 751-769.

Townsend, J. T. (1971). A note on the identifiability of parallel and serial processes. *Perception & Psychophysics, 10,* 161-163.

Treisman, A. M. (1960). Contextual cues in selective listening. *Quarterly Journal of Experimental Psychology, 12,* 242-248.

Treisman, A. M. (1964). The effect of irrelevant material on the efficiency of selective listening. *American Journal of Psychology, 77*(4), 533-546.

Treisman, A. (1977). Focused attention in the perception and retrieval of multidimensional stimuli. *Perception & Psychophysics, 22,* 1-11.

Treisman, A. M. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 194-214.

Treisman, A. (1985). Preattentive processing in vision. *Computer Vision, Graphics, and Image Processing, 31,* 156-177.

Treisman, A. (1986). Properties, parts and objects. In K. Boff, L. Kaufman, and J. Thomas (Eds.), *Handbook of Perception and Human Performance: Vol. 2* (chap. 35, pp. 1-70). New York: Wiley.

Treisman, A. (1988). Features and objects: The Fourteenth Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology, 40A,* 201-237.

Treisman, A. (1990a). *There's more to search than similarity: Conjoining features can take time.* Unpublished manuscript.

Treisman, A. (1990b). [Types or tokens in illusory conjunction]. Unpublished raw data.

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., & Paterson, R. (1984). Emergent features, attention and object perception. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 12-31.

Treisman, A., & Schmidt, N. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology, 14,* 107-141.

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

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Dornic (Ed.), *Attention and performance, VI* (pp. 331-361). Hillsdale, NJ: Erlbaum.

Van Essen, D. C. (1985). Functional organization of primate visual cortex. *Cerebral Cortex, 3,* 259-329.

Von der Malsburg, C. (1985, July). *A mechanism for invariant pattern recognition.* Paper presented at a conference on visual attention and action, Bielefeld, Germany.

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