

## Perceptual Grouping and Attention in Visual Search for Features and for Objects

Anne Treisman

University of British Columbia, Vancouver, British Columbia, Canada

This article explores the effects of perceptual grouping on search for targets defined by separate features or by conjunction of features. Treisman and Gelade proposed a feature-integration theory of attention, which claims that in the absence of prior knowledge, the separable features of objects are correctly combined only when focused attention is directed to each item in turn. If items are preattentively grouped, however, attention may be directed to groups rather than to single items whenever no recombination of features within a group could generate an illusory target. This prediction is confirmed: In search for conjunctions, subjects appear to scan serially between groups rather than items. The scanning rate shows little effect of the spatial density of distractors, suggesting that it reflects serial fixations of attention rather than eye movements. Search for features, on the other hand, appears to be independent of perceptual grouping, suggesting that features are detected preattentively. A conjunction target can be camouflaged at the preattentive level by placing it at the boundary between two adjacent groups, each of which shares one of its features. This suggests that preattentive grouping creates separate feature maps within each separable dimension rather than one global configuration.

The separate existence of an object and the boundaries that define and divide it from its surroundings must be established before the object can be identified. Perceptual segregation must therefore occur early in visual processing. There has been considerable interest recently in the nature of the grouping operation, which organizes the visual field into objects and their backgrounds. Object boundaries can be signaled by differences between adjacent areas in color or brightness, or in texture. Perceptual grouping is a special case of texture segregation, in which the texture is composed of discrete elements above some minimal size. The present article is concerned with the relations between attention and perceptual grouping in determining the perception of objects.

The phenomenology of perceptual organization was first extensively studied by the Gestalt psychologists (e.g., Wertheimer, 1923). They showed that similarity, proximity, and common movement of discrete elements are strong determinants of perceptual grouping. These simple principles maximize the chance of putting together parts of the same object and separating parts of different objects; presumably, they evolved for this reason. Some organisms have also evolved means of outwitting the parsing skills of their predators, using camouflage to break up local continuities of color, brightness, line, or texture (Fox, 1978).

The manner in which similarity controls perceptual grouping was further explored by Beck (1966, 1967) and by Olson and Attneave (1970). Beck found that differences in line orientation could be as effective as differences in brightness in segregating two groups of elements. A display divided into regions containing Ts and tilted Ts segregates easily, whereas one divided between Ts and Ls does not. Beck suggests that similarity grouping is "fundamentally an acuity task involving the simultaneous discrimination of stimulus differences prior to a nar-

---

This research was supported by a grant from the National Scientific and Engineering Research Council of Canada.

I am grateful to Daniel Kahneman and William Prinzmetal for helpful criticism and discussions, and to Randy Paterson for his help in running the experiments.

Requests for reprints should be sent to Anne Treisman, Department of Psychology, University of British Columbia, 154 - 2053 Main Mall, Vancouver, British Columbia, Canada V6T 1Y7.

rowing or focusing of attention" (1972, p. 13). Any features that are easily discriminated with peripheral vision in a patterned visual field will also mediate grouping. Beck (1972) also showed that discriminability is determined differently when attention is focused and when it is spread across a display. Thus the advantage in discriminability of Ts versus tilted Ts compared to Ts and Ls disappears when single items are presented in an otherwise empty field.

Julesz (1975) tried to specify in statistical terms the properties that determine texture segregation. He proposed that the visual system automatically detects first- and second-order dependencies among local elements and segregates any areas that differ systematically at either of these two levels. Higher order dependencies are detected only slowly, with effort and "scrutiny"; they cannot mediate the spontaneous emergence of salient boundaries. Later, Caelli, Julesz, and Gilbert (1978) and Julesz and Caelli (1979) included as potential mediators of texture segregation a few specific higher order combinations, which represent the local geometric properties of collinearity, corner, and closure. Most recently, Julesz (1980) reduced to just two basic elements the local cues that he believes can distinguish textures sharing the same second-order statistics: these elements, which he calls "textons," are (a) "elongated blobs" and (b) the number of line ends or "terminators."

The theories all agree that perceptual grouping occurs automatically and in parallel, without attention or "scrutiny," as Julesz (1980) puts it. This preattentive organization should then affect all subsequent stages of processing. Treisman, Sykes, and Gelade (1977) proposed a new hypothesis about the role of attention in the perception of objects that has implications for the nature and use of perceptual grouping. The theory suggests that focused attention is necessary for the accurate combination of features or properties of complex objects whenever more than one object is presented in an unpredictable context. Separable features are registered independently and in parallel across the visual field. However, when objects must be identified and no contextual cues are available to select the expected com-

binations of features, a correct synthesis can be achieved only by focusing attention on one location at a time. Features which co-occur in a single fixation of attention are combined to form an object. If attention is diverted or overloaded, the features may be wrongly recombined, forming "illusory conjunctions" (Treisman & Schmidt, 1982).

Treisman and Gelade (1980) tested this feature-integration theory of attention in several experiments. We showed that visual search for targets defined by one or more disjunctive features (e.g., blue *or* curved) occurs in parallel across a spatial display, whereas search for a conjunction target (e.g., red *and* vertical) requires a serial, self-terminating scan through items in the display, suggesting that attention must be focused on each item in turn. The linear functions and 2:1 slope ratio of negative to positive search times were maintained across two different levels of target/distractor discriminability, which produced markedly different search rates (40 msec and 92 msec per item). The serial scan for conjunction targets was therefore not due to the overall level of difficulty. Conversely, a difficult feature search (for ellipses differing slightly in size) produced very slow search rates but neither linear functions nor 2:1 slope ratios.

In a different paradigm, we showed that feature targets can be identified without being correctly localized, whereas for conjunction targets, identity and location appear to be interdependent. The interdependence is consistent with the claim that conjunctions are correctly registered only when attention is focused on the relevant location. Finally, when attention is diverted or overloaded, illusory conjunctions of features are in fact experienced (Prinzmetal, 1981; Treisman & Schmidt, 1982). Illusory objects may recombine either dimensions, like the shape, size, and color of different real objects, or component parts, like the curves, lines, and angles of more complex shapes.

The convergence of these results from a variety of different experimental paradigms provides stronger support for the theory than any single finding on its own. Townsend (1972) has shown that some models of parallel as well as of serial search can predict

linear functions, but the fact that attention limits are implicated in the formation of illusory conjunctions and in the accurate localization of conjunction targets makes it plausible in conjunction search to attribute the linear increase in latencies with increased display size to a serial self-terminating scan with focused attention directed to each item in turn.

Some predictions about perceptual grouping and segregation also follow from the theory. The first was tested by Treisman and Gelade (1980), and the others are explored in the present article.

### 1. Determinants of Grouping

If grouping is an early, preattentive process, it should be mediated only by the discrimination of simple, separable features. Grouping should therefore not be immediately apparent when two regions of texture elements differ only in the way in which the same sets of features are combined. We showed that elements that differ either in color (red Os and Xs vs. blue Os and Xs) or in shape (red and blue Os vs. red and blue Xs) segregate easily into salient perceptual groups, whereas elements that differ only in the conjunction of colors and shapes (for example red Os mixed with blue Xs in one area vs. red Xs with blue Os in the other) do not automatically segregate into perceptual groups (Treisman & Gelade, 1980). The boundaries of groups defined only by conjunctions of features are found through slow, perhaps serial, scanning and seem to be inferred rather than directly perceived.

In general this hypothesis makes similar predictions to those of Julesz (1975) and Beck (1972); but whereas Julesz tries to define a priori the statistical criteria that result in preattentive grouping, we propose that preattentive grouping can be mediated by any independent population of feature detectors, which are present in the perceptual system. Thus we predict a correlation between the ease of segregation and the presence of any features that in other tests have demonstrated separability (Garner, 1974; Treisman & Gelade, 1980). It is also possible that new, functionally independent detectors could be set up through prolonged

practice, as in Laberge's (1973) "unitization" process or Shiffrin and Schneider's (1977) "automatization" of search with constant mapping of target and distractors. Any newly acquired feature detectors that allow automatic detection or matching unaffected by expectancy or attention, should then also be capable of mediating perceptual segregation.

### 2. Grouping Effects on Attention

A second prediction (tested in Experiment 1) is that the results of preattentive grouping should control the direction of focused attention. If the function of perceptual grouping is to set up candidate objects for further processing, and if attention is necessary for object identification, it follows that we should attend initially to groups of items when these are present rather than to single items within groups. Thus the perceptual organization of a display should have marked effects on search for conjunction targets but not for feature targets, since these are detected automatically without attention.

The relation between attention and perceptual grouping has been discussed by Kahneman and Henik (1977). They describe a phenomenon that they call "group processing." Group processing in the report of tachistoscopic arrays is characterized by similar and highly correlated probabilities of report for items within the same perceptual group and by discontinuities and negative correlations in the probabilities of report for items in different groups. Kahneman and Henik (1977) suggested that attention is allocated in a hierarchical fashion, first to a group as a whole and only subsequently to elements within a group. They also looked at selective attention in relation to the degree of spatial segregation or mixing of red and blue letters. In attempting to recall the red letters only, subjects showed a marked advantage of grouped displays compared to a checkerboard arrangement. The results suggest that attention can only (or much more effectively) be focused on relevant items if these are spatially grouped.

The metaphors chosen for attention in Kahneman's model and in feature-integra-

tion theory imply rather different functions. For Kahneman, attention plays the role of an enabling input or supply of resources that is allocated to items to facilitate their processing. In feature-integration theory, attention functions as a selective device or spotlight, selecting which features should be conjoined and how. These two analogies differ in their emphasis but are not incompatible. A spotlight can vary in intensity as well as in spatial direction and area. The result of distributing a resource more widely when the pool is limited is that each recipient gets less of it; in feature-integration theory, the result of spreading the attention spotlight over a larger number of items is to increase uncertainty in the allocation of features to objects. Attention may serve both functions: first selecting which objects will be accurately resynthesized and then facilitating any further processing they receive.

### 3. *Grouping and Illusory Conjunctions*

Another prediction also follows from the idea that attention is allocated to groups rather than to items: If an otherwise homogeneous group contains two items with contrasting features (e.g., red in a green group), the theory implies that their locations and conjunctions should remain uncertain until attention is narrowed to contain each unique item alone. We would expect, therefore, more illusory conjunctions to be formed within perceptual groups than across them. Prinzmetal (1981) has shown that illusory conjunctions are more often seen when two interchangeable features are present within a single group than when they occur in separate groups.

### 4. *Spatial Density and Search for Conjunctions or Features*

It should be possible to distinguish the effects of perceptual grouping from those of spatial distribution. The physical or retinal separation of items should affect conjunction formation only if and when it acts as a means of segregating groups of items that can then receive separate attention. Any differences in the spatial distribution of items that have no impact on their preattentive grouping

should also leave conjunction search unaffected. Retinal location may affect the accuracy of feature registration, since acuity and color discrimination decrease away from the fovea. To counteract these peripheral limits, we expect more eye movements to be made with displays in which items are distributed over a wider area whenever their features become confusable at the peripheral locations. This should not, however, result in the strictly serial checking with the mental (rather than the physical) "eyeball," which is necessary in conjunction search. These predictions are tested in Experiment 3.

### 5. *Separate Preattentive Feature Maps and Serial Scanning of Groups*

A fifth prediction is somewhat counterintuitive and contrasts with the generally accepted view first proposed by Neisser (1967). If features are registered independently of each other, they should form several separate preattentive organizations rather than one global configuration. In fact there should be one for every population of separable feature detectors. Each should offer its own candidates for object and event identification, and they should be interrelated only with the aid of focused attention.

The idea of separate feature maps may need some clarification. Features that are spatially defined, like shapes and sizes, are themselves arrangements or spatial extents of other features like brightnesses or colors, with which they are physically conjoined. We should, however, distinguish real displays from coded representations of those displays. It is logically possible to register the presence of something red without specifying what or where it is. Sizes, line orientations, curvature, intersections and so on may be abstracted from the particular brightness contrasts that physically embody them. Treisman and Schmidt (1982) showed that illusory conjunctions may exchange the shapes, sizes, colors, or solidity (outline vs. filled) of real objects and as a consequence substantially alter, for example, the amount of a certain color that is seen, or alter its spatial distribution.

Location may differ from other features, however. In order for a feature "map" to be

structured, at least two dimensions must be conjoined, one to act as medium for the other. Kubovy (1981) suggested that for visual stimuli, space and time are the "indispensable attributes" that perform the role of media to carry variation in other attributes. This would give location a special status in perceptual coding. There is in fact considerable evidence that locations are (in some sense of the word) registered early; examples are the localized receptive fields of visual cortical cells and the fine spatial resolution that mediates stereoscopic depth. It does not follow, however, that the locations of features are also automatically available to the mechanisms that discriminate and identify multidimensional objects. Location is an ambiguous term: It could refer simply to relations in space, such as "left of" or "below," or it could be defined in absolute coordinates, either on the retina or in the physical three-dimensional world. If locations were coded relationally within each feature map, without absolute coordinates or common scale, this would allow perceptual grouping and segregation of different colors, orientations, and so on, but would not be sufficient to allow correct alignment of one feature map with another. Locations may therefore be conjoined with each set of separable features at an earlier stage and by a different process from that which conjoins other features with each other.

If focused attention is, as we suggest, the means by which cross-dimensional integration takes place, it must allow access to cross-dimensional spatial coordinates so that absolute locations can be matched across the different feature maps. These absolute coordinates would presumably refer to real-world locations in three-dimensional space. Serial processing with focused attention may therefore be needed not only to conjoin features but also to interpret relative feature locations as points in a common spatial representation of the external world.

Kahneman and Henik's (1977) results suggest that focused attention can be directed to the locations of preattentively defined groups. Since we claim that there are several separate preattentive feature maps, one of them must be selected to control attention at any given time, either by its task

relevance or by the amount of organization that it contains and the economy of search that it consequently allows. A relational coding of spatial layout could be sufficient to guide the direction of attention to particular items within a single feature map. The attended features could then be localized in absolute coordinates within a representation of external space. Features at the same location but on other dimensions would also be accessed and all would be conjoined to form a structured, multidimensional object. The mechanisms that achieve the integrated spatial representation and that conjoin features to identify objects are not our present concern. This article attempts to explore the nature of preattentive processing and to specify which operations require attention and must therefore be performed serially.

What effects would this account entail for search tasks with grouped displays? The claim is that the serial scan with focused attention, which mediates search for conjunction targets, is needed to bring the different feature maps into register with each other. When the display is randomly mixed, it seems this must be done sequentially for each item in turn. However, when the display is divided into regions of homogeneous items, the exact location of each feature relative to the others within its group may be irrelevant; illusory exchanges of color and shape between two identical items would produce no detectable error. In this case, it may be sufficient to match up feature maps at a cruder resolution, focusing attention on groups to integrate these, rather than individual items. This matching at cruder resolution would be sufficient to determine, for example, that all items in a group of Xs are green, but would not decide the color of each X in turn. It should then be possible to check items within each group in parallel, even in search for a conjunction target. In looking for a green H in a background of separately grouped green Xs and red Hs, early segregation of items by color, for example, could allow attention to be focused on the homogeneous groups that share one target feature (e.g., green). A parallel search within each focused group should then allow detection of the other target feature (e.g., vertical) without requiring a serial scan to locate it

more precisely. Search times should vary with the number of groups rather than with the number of items. This prediction is tested in Experiment 2.

Our earlier results with conjunction search confirm the conclusion reached by Kahneman and Henik (1981) that there are constraints on the spatial distribution of selective attention. They show that attention cannot be distributed over a subset of items (e.g., the red ones or the curved ones) when these are spatially scattered among other items in a randomly mixed display. If subjects could attend selectively and in parallel to all the red items in a mixed display of red horizontal and green vertical lines, this should allow them to detect a conjunction target (e.g., red and vertical) within the red subset, as if it were defined by a single feature (e.g., a vertical line among horizontals). The fact that this could not be done suggests that attention can be focused only on sets of spatially grouped items. The spotlight of attention appears to be unitary (or at least not multiply divisible) and may be constrained in the possible shapes it can assume, although it may vary in size (Eriksen & Hoffman, 1972).

### 6. *Grouping and Camouflage*

A final prediction, tested in Experiment 4, also follows from the idea that a number of separate preattentive feature maps are formed, rather than one global configuration. If a conjunction target is placed at the boundary between two groups, each of which shares one of its features, it should be embedded as one of a group in both preattentive feature maps and should exist as a unique entity in neither. Thus a red X at the boundary between blue Xs and red Os should be preattentively invisible. It should become visible only when attention is narrowed sufficiently to exclude at least one group of distractor items. In the everyday world, of course, such situations would be rare; the different preattentive maps will normally give highly correlated parsings of figure and ground. Our hypothesis may help to explain some cases of natural camouflage, however. It could be that animals conceal themselves from predators by sharing, for

example, their color with some elements of their background and their shape with others.

In summary, the predictions to be tested in the present article are that (a) perceptual grouping should have little effect on feature search but large effects on conjunction search through its role in the control of attention; (b) spatial density without perceptual grouping should have no effect on the rate of conjunction search, since the serial scan involves the mental rather than the physical eyeball; (c) when homogeneous groups of items are present in the display, groups rather than single items will be serially scanned; and (d) adjacent groups of items should preattentively camouflage an item at their boundary if the item conjoins one feature from each group and has no unique distinguishing feature of its own.

### Experiment 1: Perceptual Grouping and Search for Conjunctions

Experiment 1 investigates the effect of perceptual grouping on search for objects defined either by features or by conjunctions. If feature targets are detected through the same preattentive operations that segregate groups and textures, they should be affected by the same variables that control the grouping itself—the discriminability of adjacent shapes or colors in the display. Detection of feature targets should vary with the size and number of perceptual groups only to the extent that these determine the local context of the target. Thus, a green target might be easier to detect when surrounded only by yellow distractors in its immediate vicinity than when surrounded by a mixture of blue and yellow distractors; this would follow if and because two different discriminations are harder to make than one. The effects should be small, unless the discriminability of the target from the different distractors varies substantially within the displays.

Detection of conjunction targets, on the other hand, requires attention. If items are spatially grouped by a shared feature, attention may be directed to groups rather than to items. We should then expect a large effect of the degree of spatial homogeneity on search for conjunctions of features.

Experiment 1 compared the effects of grouping and of display size on visual search. Two groups of subjects were tested: one on conjunction targets and one on feature targets. Each group was run in two conditions: (a) with display size varying from 1 to 36 items and (b) with display size fixed at 36 items, but the number of homogeneous groups within the display varying from 1 to 36.

### Method

*Subjects.* Eight subjects (five female and three male) were tested with the conjunction target and eight subjects (three female and five male) with the feature targets. All were students, aged 16-25, who volunteered to take part in the experiment and were paid \$3 an hour.

*Stimulus materials.* Two sets of stimulus cards were constructed: a feature set and a conjunction set. The distractors in both sets were red Hs and green Xs. There were two types of target: In the conjunction condition, the target was a green H, which differed from the distractors only by the conjunction of its color and shape. In the feature condition, the target was defined disjunctively: On each trial subjects were asked to search for both a blue letter and an S without knowing which, if either, would be present. The blue letter was an H or an X and the S was green or red, but these features were irrelevant to the task. Thus each distractor shared one of the target features in both feature and conjunction conditions. Subjects had to check both color and shape for the distractors in both the conjunction and the feature condition, but they had to check how color and shape were conjoined only in the conjunction condition.

Two sets of displays were made for both the feature and the conjunction search conditions. In one set, display size was varied; in the other, display size was fixed at 36 items, and the number of homogeneous groups of distractors was varied. The sets that varied display size contained examples of cards with 1, 4, 16, and 36 items; eight cards for each display size in each set contained a target (positive), and eight cards did not (negative). The items were arranged in regular rectangular matrices ( $2 \times 2$ ,  $4 \times 4$ ,  $6 \times 6$ ) with a 1.3:1 ratio of width to height. Displays containing only one item used one location from the  $2 \times 2$  matrix, taking each of the four locations equally often. The average distance of items from the fixation point was  $2.14^\circ$  for all display sizes. Each item subtended  $57^\circ$ . Target positions were chosen so that the mean distance from fixation for targets was equal to the mean distance from fixation for the distractors in each set. Targets were placed equally often in each quadrant of the matrices. If negative conjunction displays of four items had always contained equal numbers of each distractor type, it would have been possible for subjects to use an easily subitized number of features as a cue to the presence of the target. To avoid this, we made an equal number of positive displays consisting of (a) two red Hs, a green X and a green H, and (b) two green Xs, a red H and a green H. Similarly, there were equal numbers of negative displays with two red

Hs and two green Xs and with three of one color or shape and one of the other. A 3:1 ratio of green to red or of Hs to Xs could therefore not be a cue to the presence of a target. The same distractors were used in the feature target condition.

For the grouped condition, the same matrix of 36 positions was used and all displays contained 36 items. However, the red Hs and the green Xs were arranged in varying numbers of homogeneous groups: They were either randomly placed (as 36 separate items), or grouped in horizontal pairs of identical items (to make a checkerboard of 18 groups), or grouped into 9 alternating squares of  $2 \times 2$  identical items (to make 9 groups), or grouped in 4 alternating squares of  $3 \times 3$  identical items (to make 4 groups), or formed a single group of identical items, either all green Xs or all red Hs. The green H conjunction target or the blue or S disjunctive target in positive displays replaced one item, equally often in each quadrant and in each type of group (red Hs and green Xs).

The cards were shown in a two-field Cambridge tachistoscope. The onset of each display started a millisecond timer, which stopped when the subject pressed either of two response keys.

*Procedure.* One group of subjects was tested only on conjunction targets and one group only on feature targets. They were tested in two sessions consisting of one block of varied display size, followed by one block of grouped displays on the first day and then one block of grouped displays and one of varied display sizes on the second day. Each block consisted of eight trials with and eight without a target at each display size or in each grouping condition. Subjects were informed of the nature of the displays in both conditions and were shown examples of the cards. In the varied display size blocks, trials with displays of 1, 4, 16, and 36 were randomly mixed, as were trials with varying numbers of groups in the grouped display blocks.

The instructions were to decide as quickly as possible, while minimizing errors, whether the display contained a target and to press the right-hand key if it did and the left-hand key if it did not. Subjects were given a few trials in each condition for practice before beginning the experimental trials.

### Results and Discussion

The mean search times are shown in Figure 1, and details of the linear regressions are given in Table 1. The results with varied display size confirm our previous finding (Treisman & Gelade, 1980). Conjunction targets give strictly linear functions relating search time to display size, and a 2:1 ratio of negative to positive slopes. This pattern suggests a serial self-terminating scan and is consistent with the hypothesis that focused attention is directed to each item in turn to check the conjunction of color and shape. The search rate is similar to the rate we obtained earlier with a green T target, green

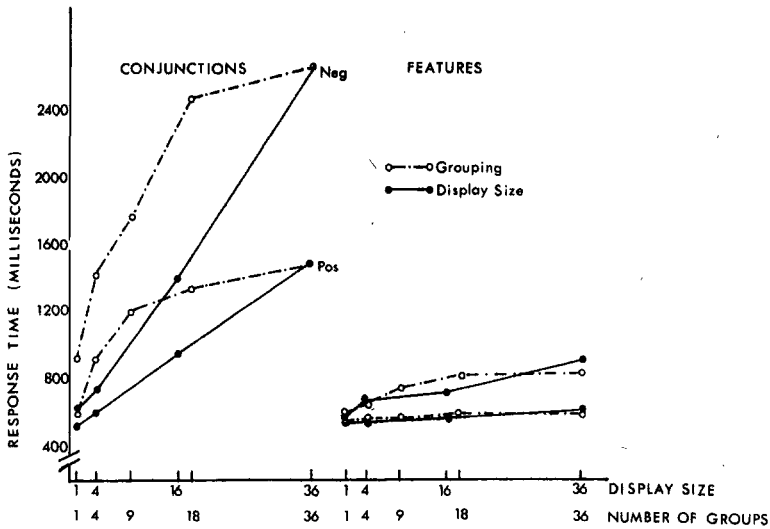


Figure 1. Mean search times for feature and conjunction targets in Experiment 1.

X, and brown T distractors spread haphazardly over the larger area of  $15^\circ$  by  $8^\circ$ .

The feature targets again showed much less effect of display size, averaging a search rate of only 2 msec per item on positive displays compared to 28 msec for conjunctions. For color targets, the effect of display size did not reach significance. The functions were also less linear (linearity accounting for only 85% and 80% of the variance due to display size for color and for shape, respectively) compared to almost 100% for conjunctions. The slope ratios for features were much larger than 2:1. Again these results replicate our earlier finding and suggest that disjunctive features, when present, can be detected in parallel across the display. Note that with single-item displays, there is no difference between feature and conjunction latencies. This makes it unlikely that the difference is due to the discriminability of the two types of target. The difficulty with conjunctions arises only when more than one item is presented and the load on attention is increased.

Individual subjects varied considerably in their search rates; the slopes of response time against display size ranged from 23 to 41 ( $SD = 5.3$ ) for conjunction positives, and from 38 to 72 ( $SD = 10.7$ ) for conjunction negatives. In every case, however, the proportion of the variance with display size that

was due to linearity was over 99%. In the feature search condition, the positive slopes ranged from  $-1.0$  to  $4.6$  ( $SD = 1.6$ ) and the negative slopes ranged from  $2.5$  to  $28.0$  ( $SD = 7.9$ ). The proportion of the variance due to linearity ranged from 6% to 99% with 85% for the median subject.

The main interest in this study was in the effect of grouping. With feature targets in positive displays, there is little scope for grouping to have an effect, since targets are detected only 60 msec to 80 msec more slowly in displays of 36 items than in displays of 1. The effect of grouping did not reach significance for the positive displays; however, it was significant for the negatives,  $F(4, 28) = 5.43$ ,  $p = .002$ . Farmer and Taylor (1980) reported a similar result: Search times for a target defined by color were unaffected by the color grouping of distractors on positive trials but showed a substantial effect of heterogeneous mixing of distractors on negative trials. It seems that on negative trials, subjects adopt the strategy of scanning the display more carefully to confirm the absence of the target. They then show some effect both of display size and of grouping, although these are an order of magnitude smaller than with conjunction targets.

In conjunction search, on the other hand, the number of groups has a striking effect. The results show a major decrease in search



Table 1  
*Analysis of Mean Search Times in Experiment 1*

Condition	$p <$	Slope	%V <sup>a</sup>	Slope ratio	Int.	% errors
Display size						
Conjunction						
Positive	.001	28.2	> 99	.49	486	6.7
Negative	.001	58.0	> 99		511	4.5
Feature						
Shape positive	.001	2.4	85	.28	521	8.7
Color positive	—	1.7	80	.20	539	7.2
Negative	.001	8.6	96		589	4.5
Grouping						
Conjunction						
Positive	.001	21.5	75 <sup>b</sup>	.47	809	8.6
Negative	.001	45.8	82 <sup>b</sup>		1,202	2.7
Feature						
Shape positive	—	.6	15	.10	527	4.2
Color positive	—	2.2	99	.37	552	8.7
Negative	.002	5.9	77		640	2.9

Note. Int. = Intercept.

<sup>a</sup> Percentage of variance with display size due to linearity.

<sup>b</sup> Departure from linearity significant ( $p < .01$ ).

time as the groups become larger and fewer, until with a homogenous display of 36 items the positive search time is only slightly longer than with a single target item. It is clear that subjects do not serially check each of the 36 items when these are perceptually grouped. However, the search time does not decrease linearly with the number of groups. It decreases much more slowly at first, as the number of groups drops to 18 and then to 9 and then more steeply as it drops to 4 and then to 1. Subjects benefit less from small groups (of 2 or 4 items) than they should if they were serially checking groups of items. One reason might be that focusing attention accurately enough to include more than one homogeneous item but to exclude any heterogeneous items is itself a time-consuming operation. The smaller the groups the less perceptually salient they become.

Another possibility is that in randomly mixed displays, subjects also perceptually group items into clusters averaging 2 or 3 homogeneous items that happen to be spatially adjacent to each other. This would

change the interpretation of the slope in the varied display size condition; it should then be understood not as 58 msec per serial check of each item in turn, but, perhaps, as 116 msec per serial check of each group of 2 items in turn, or 174 msec per group of 3 homogeneous items. If this hypothesis is correct, the grouped condition would show no difference between search times for 18 groups and for 36 randomly mixed items, simply because the 36 items are in fact treated as approximately 18 groups. As the groups in the grouped condition grow larger than the randomly occurring groups in heterogeneous displays, the search times start to reflect their presence.

These results differ from those of Kahneman and Henik (1981), who found no effect of grouping on latencies for target detection. There are many differences between the experiments: the displays in Kahneman and Henik's experiment were presented for a fixed, short interval (200 msec), whereas ours remained available until the subject responded. Perhaps more important is the fact

that their targets were digits in digit distractors, and the basis they used for organizing groups was a spatial discontinuity, a property that was irrelevant to the definition of the target. In terms of feature-integration theory, there would be no gain in Kahneman and Henik's task from focusing attention serially on groups: If the target digit happened to have a unique feature, it would be picked up in parallel across the whole display; otherwise it would require serial checking, item by item, regardless of whether the items were spatially grouped.

### Experiment 2: Visual Search Between and Within Perceptual Groups

Experiment 1 explored the effect of perceptual grouping on displays of constant size. The number of groups therefore varied inversely with the number of items per group. The next experiment varied these two factors orthogonally to compare directly their effects on search time. The theory predicts that when groups are homogeneous, subjects carry out a self-terminating series (one per group) of feature checks, which are parallel within each group of homogeneous items. The total search time should therefore reflect the number of groups to be checked multiplied by the time taken for a feature check of a group of that size. The search of groups should be self-terminating, while the search of items within groups should be exhaustive up to the last group searched.

### Method

*Subjects.* Two groups of nine subjects were tested, four men and five women in each. They were students, aged 16–30, who volunteered to take part in the experiment and were paid \$3 an hour. They had not taken part in Experiment 1.

*Stimulus materials.* We used the same colored letters as in the conjunction condition of the previous experiment. They were spread over a larger area however, and displays of up to 81 letters were included. The groups in this experiment were spatially separated by empty space as well as segregated by the color and shape of their items. There were nine types of display, comprising all the combinations of one, four, and nine groups and one, four, and nine letters per group. Examples of four types are shown in Figure 2. The groups were arranged in a  $3 \times 3$  square matrix with the outer corners at  $6.9^\circ$  from the center. Each group of nine letters subtended  $2.35^\circ$ , and the total display of 81 letters subtended  $9.67^\circ \times 9.67^\circ$ . Letters within a group were ar-

ranged in a  $3 \times 3$  matrix, or in a smaller  $2 \times 2$  matrix keeping the same interletter spacing as the  $3 \times 3$  matrix, or a single letter was placed at the center position within a group. For displays containing only one group, each of the nine group positions was used equally often; for displays containing four groups, the positions of the groups were chosen randomly from the nine possible, except that each position was used equally often and no combination of four positions was used more than once.

The distractors within each group were always homogeneous, comprising either green Xs or red Hs. There were equal numbers of green X and of red H groups across the complete set of displays. Within the displays of four groups, there were two of each kind, and within the displays of nine groups, there were four of one kind and five of the other. Their positions were randomly mixed in the matrix of nine group positions. Nine displays of each type contained no target. Nine other displays had one item replaced by a green H. In five displays of each type it replaced a green X and in four it replaced a red H. It appeared once in each group and once in each position within the groups of nine, twice or three times in each position within the groups of four. The stimuli were again presented in a Cambridge two-field tachistoscope, and subjects' responses were timed with a millisecond timer.

*Procedure.* Two groups of nine subjects each were used in this experiment. One group was given all the display types randomly mixed together. The other group was given separate blocks containing the displays with one letter per group, four letters per group, and nine letters per group to see whether different strategies would be adopted. The order of these three conditions was counterbalanced across subjects and reversed for the second session of each subject.

The procedure was otherwise the same as in Experiment 1. The task was to search for the conjunction target (green H) and to press the right-hand key if it was present and the left-hand key if it was not.

### Results and Discussion

There were no significant differences between the results of the two groups of subjects (those who had all the trial types randomly mixed and those who had displays with different numbers of letters per group in separate blocks). It seems that any strategy differences were minor and insignificant.

The theory predicts serial, self-terminating search of groups and feature search within each group. The feature search time within groups would be expected to show a small effect of group size. In Experiment 1, search times on feature trials increased slightly with display size. In the present experiment all groups except the last one checked on positive trials would give negative results; extrapolating from the negative latencies in the feature trials of Experiment

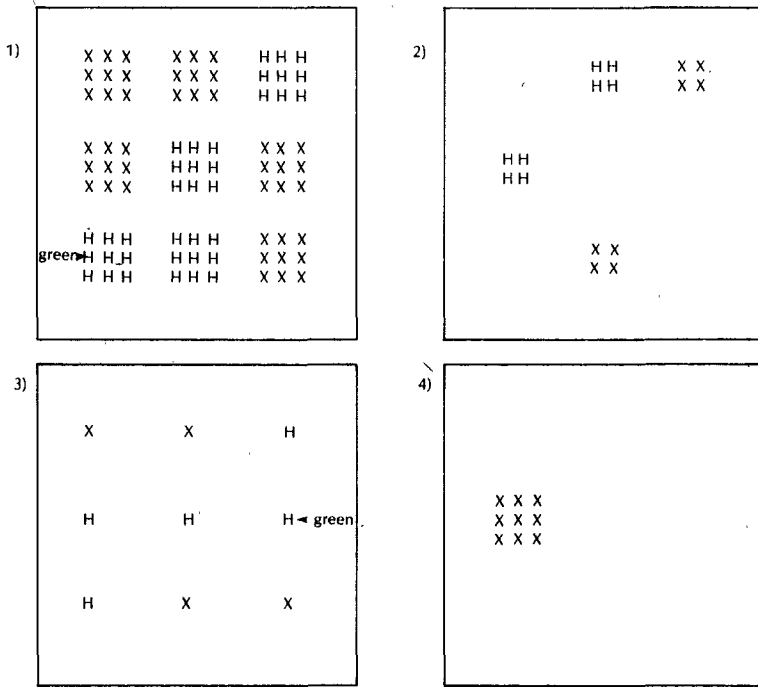


Figure 2. Examples of displays used in Experiment 2: (1) Nine groups of nine letters per group with target present; (2) four groups of four letters per group with target absent; (3) nine groups of one letter per group with target present; (4) one group of nine letters with target absent. (All Xs were green and all Hs were red with the exception of the green H targets, labeled in the figure, examples [1] and [3].)

1, we expect the effect of the numbers of letters per group in Experiment 2 to average 8.6 msec per extra letter. There could also be an effect of group size on the intercept, for example, through changes in confidence or in lateral masking. This makes the number of parameters to be fitted rather large. However, three predictions can be tested against the data: (a) for each group size, the effect of the number of groups on search times should be linear; (b) for each group size, the ratio of positive to negative slopes for groups should be 1:2; (c) the effect of the number of letters per group should be shown in the differences between the slopes for groups of one, four, and nine letters; this effect should be small and nonlinear, approximating 8.6 msec per letter. If, on the other hand, search were serial and self-terminating across items, regardless of grouping, the slope for groups of nine letters should be nine times the slope for groups of one.

The results, shown in Figure 3, clearly

reject the prediction of serial search of each item in turn. The ratio of the slope for nine-letter groups to the slope for one-letter groups is closer to two than to nine. On the other hand, the results are consistent with the prediction that search of groups is serial, whereas search within groups is parallel. Table 2 gives the best-fitting slopes and intercepts and the proportion of the variance with number of groups that is due to linearity. The ratios of positive to negative slopes approximate .5, and in each condition at least 96% of the variance due to display size is accounted for by a linear function. Departures from linearity were significant only in one case, positive displays with four letters per group,  $F(1, 34) = 12.1, p < .01$ . When present, these departures all take the same form, displays of four groups being a little slower than expected. A likely reason for this disparity is that more eye movements *per group* were required in these displays; the four groups were randomly dispersed over the nine possible locations, whereas

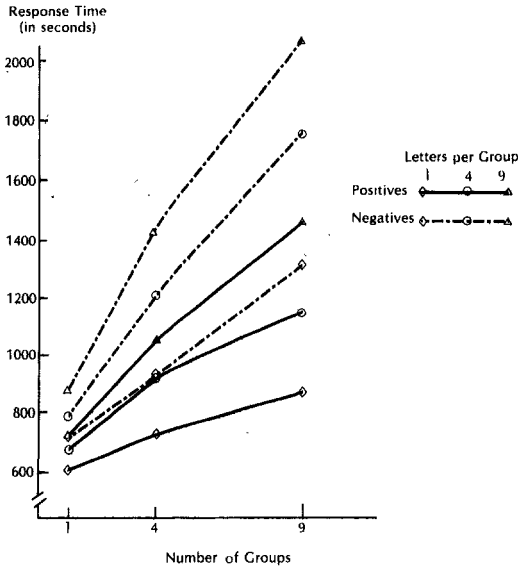


Figure 3. Mean search times in Experiment 2.

with nine groups all locations were filled. Analyzing separately the negative displays in which the four groups were located in two adjacent rows or columns, we found a mean reduction of 50 msec for the four-letter groups and 90 msec for the nine-letter groups. This brings the mean reaction times much closer to the predicted values on a linear function.

The difference between the negative slopes for groups of different sizes indicates a difference of 72 msec between search times through groups of nine homogeneous letters and search times for a single letter. This is very close to the 69 msec difference predicted from the feature search results of Experiment 1 and supports the claim that search within each group consists of checking for single features.

The search rate through groups, on the other hand, is similar to but slightly slower than the rate for conjunctions in the ungrouped displays of Experiment 1, averaging 69 msec per group compared to 58 msec per item in the more densely packed displays used earlier. Thus the present data appear to combine a serial conjunction search of groups with a parallel feature check within groups.

Individual subjects again varied considerably in their search rates, but the general

pattern of results was similar for all. Positive slopes for individual subjects ranged from 10 to 48 ( $SD = 13$ ) for displays with one letter per group, from 36 to 76 ( $SD = 16$ ) for four-letter groups and from 48 to 205 ( $SD = 45$ ) for nine-letter groups; negative slopes ranged from 33 to 186 ( $SD = 35$ ) for one-letter groups, from 44 to 219 ( $SD = 44$ ) for four-letter groups, and from 42 to 259 ( $SD = 55$ ) for nine-letter groups. Linearity accounted for more than 98% of the variance with display size in 60 of the 108 cases and for less than 90% in only 10 cases.

### Experiment 3: Spatial Density and Search for Conjunction and Feature Targets

A possible confounding variable in Experiment 2 is the spatial density of the items. The displays with nine groups of nine letters used the whole area of  $9.7^\circ \times 9.7^\circ$ , whereas the displays with one group occupied only  $2.35^\circ \times 2.35^\circ$ . Eye movements could therefore have contributed to the effect of groups but much less to the effect of items within groups. The question of how spatial density affects search for conjunctions and for features is also of some general interest. We claim that serial search for objects defined by conjunctions of features is centrally rather than peripherally determined; it represents

Table 2  
Linear Regressions of Mean Search Times on Number of Groups in Experiment 2

Letters per group	Slope	Intercept	%V <sup>a</sup>
Positive			
1	32	583	98
4	57	654	96 <sup>b</sup>
9	89	653	99
Negative			
1	73	638	98
4	119	696	99
9	145	771	98

Note. Slope ratio for 1-letter groups = .44; for 4-letter groups = .48; for 9-letter groups = .61 ( $M = .51$ ).

<sup>a</sup> % Variance with display size due to linearity.

<sup>b</sup> Departure from linearity significant ( $p < .01$ ).

successive fixations of attention rather than of the eyes. We expect an effect of spatial spread only if and when it affects the discriminability of features.

There are alternative accounts that would link the difficulty of forming conjunctions to the characteristics of peripheral vision, however. It is likely, for example, that localization becomes less precise the more peripherally the items are presented. Receptive fields are much larger in the periphery than in the fovea.

Some findings by Beck (1972), Beck and Ambler (1973), and Ambler and Finklea (1976) are relevant here. Beck (1967) and Beck and Ambler (1973) found much worse performance with distributed than with focused attention in a task requiring discrimination of line arrangements (e.g., T vs. L) but not in a task involving line orientation (T vs. X). However, Ambler and Finklea (1976) found that the difference between focused and distributed attention for line arrangements was absent or much reduced when the stimuli were presented foveally, suggesting that the difficulty is characteristic only of peripheral vision. More recently, Ambler, Keel, and Phelps (1978) obtained the opposite results in a better-controlled experiment. With line arrangements, performance deteriorated as the number of items increased, even with foveal presentation, whereas with line orientation, parallel processing was possible in both cases. It seemed worth investigating with color-shape conjunctions whether retinal area was relevant or whether search depends, as we predict, on a more central attentional scan.

We also compared the effects of the spatial density of the display on search for feature targets. Here search should be parallel rather than serial in both central and peripheral vision; the only differences should be due to poorer acuity with peripheral relative to central vision. Eye movements might become necessary when features are hard to discriminate and could introduce an effect of display size on sparsely scattered items that would be absent with foveal displays.

### Method

*Subjects.* The eight subjects (three male and five female) were students at high school or university aged

between 16 and 25. They volunteered for the experiment and were paid \$3 an hour.

*Stimulus materials.* Two sets of stimulus cards were used; a sparse set and a dense set. The dense set was the cards used in Experiment 1 for conjunction and for feature search in the varied display size condition; the sparse set was newly constructed to match in everything but spatial location. On both sets of cards, the items were arranged in regular rectangular matrices ( $2 \times 2$ ,  $4 \times 4$ ,  $6 \times 6$ ) with a 1.3:1 ratio of width to height. The items in the sparse set for each display size were on average  $4.28^\circ$  from the fixation point; the items in the dense set for each display size were on average  $2.14^\circ$  from the fixation point. All the letters subtended  $.57^\circ$ . Target positions were chosen so that the mean distance from fixation for targets was equal to the mean distance from fixation for the distractors in each set. Targets were placed equally often in each quadrant of the matrices. The apparatus and task were the same as those in Experiment 1.

*Procedure.* The four conditions, dense and sparse conjunction search and dense and sparse feature search, were run in four separate blocks with the order counterbalanced across subjects as follows: Half the subjects started with conjunctions and were tested for two sessions, one starting with a block of 64 dense trials, then a block of 64 sparse trials, and the other starting with 64 sparse trials followed by 64 dense trials. These subjects were then tested in the same way on two further sessions with feature targets. The other half of the subjects started with feature targets and then did the conjunction targets. The different display sizes, 1, 4, 16, and 36, were randomly mixed within blocks, as were positive and negative trials.

### Results

The mean correct response times in each of the four conditions are shown in Figure 4. Table 3 shows the slopes and the intercepts relating response times to display size in each condition. The slopes for individual subjects ranged from 22.8 to 45.6 for the conjunction positives ( $SD = 5.5$ ) and from 44.5 to 86.9 ( $SD = 14.4$ ) for the conjunction negatives. In all but three of the 32 cases the proportion of the variance with display size that was due to linearity was over 98%. For the feature targets, individual slopes ranged from 1.8 to 7.3 ( $SD = 1.8$ ) for the positives and from  $-.6$  to 37.6 ( $SD = 10.1$ ) for the negatives. Twenty-one out of 32 slopes had less than 98% variance attributable to linearity, the lowest being 15%.

Analysis of variance (ANOVA) on the conjunction positive trials showed a small but significant effect of density,  $F(1, 7) = 6.72$ ,  $p = .04$ , but no interaction between display size and density  $F(3, 21) = 1.0$ . Thus, for

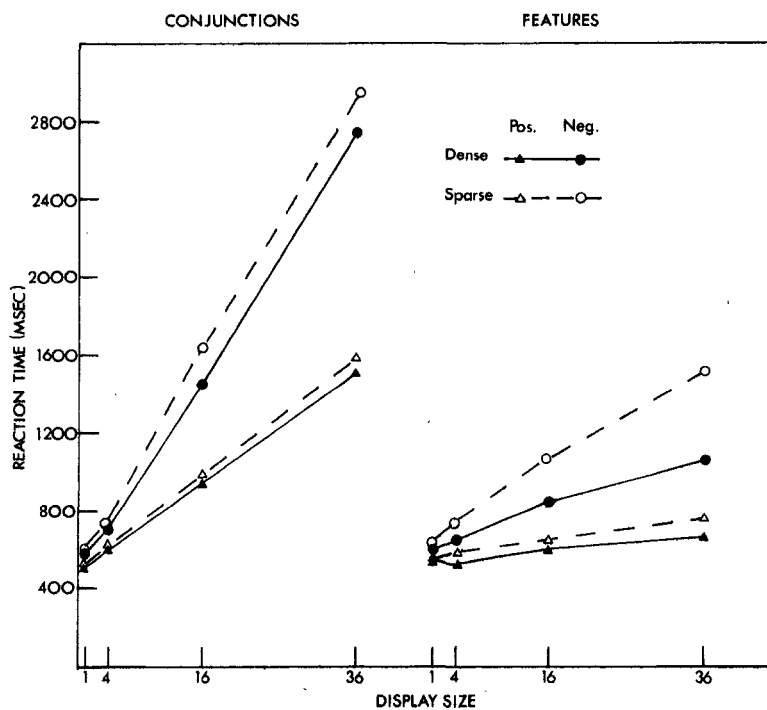


Figure 4. Mean search times in Experiment 3.

positive displays, doubling the spatial density slightly speeded up the decision times but had no effect on the rate of serial search for conjunction targets. On negative trials, density did interact with display size,  $F(3, 21) = 4.78, p = .01$ , as well as affecting the overall decision time,  $F(1, 7) = 11.86, p = .01$ . Subjects took an extra 5 msec per item on average, perhaps reflecting some additional rechecking before the decision that there was no target.

For feature targets there was a small overall effect of density on search for positives,  $F(1, 7) = 14.5, p < .01$ , and a larger effect on negatives,  $F(1, 7) = 28.8, p = .001$ , and in both cases a significant interaction with display size,  $F(3, 21) = 3.79, p < .05$ , and  $F(3, 21) = 26.06, p < .001$ , for positives and negatives, respectively.

### Discussion

The results with conjunction targets again confirm our previous findings of linear functions relating search time to display size and a 2:1 ratio of negative to positive slopes. The aim of the present experiment was to see

whether search rates were affected by the spatial spread of items in the display. The fact that doubling the area to be searched had little or no effect on search rates is con-

Table 3: Linear Regressions of Mean Search Times on Display Size in Experiment 3

Display type	Slope	Int.	Slope ratio	%V <sup>a</sup>
Conjunction				
Dense				
Pos.	28.5	482	.46	>99.9
Neg.	62.3	486		>99.9
Sparse				
Pos.	30.6	495	.45	>99.9
Neg.	67.7	523		>99.9
Feature				
Dense				
Pos.	3.4	539	.26	96.4
Neg.	13.0	603		99.3
Sparse				
Pos.	5.9	559	.24	99.5
Neg.	24.7	638		99.8

Note. Int. = Intercept. Pos. = positive. Neg. = negative.  
<sup>a</sup> Variance with display size due to linearity.

sistent with the hypothesis that the serial scan is centrally controlled and represents successive fixations of attention rather than peripheral fixations of the eyes. Eye movements were certainly unnecessary with the dense targets, which were all within  $2.2^\circ$  of the fovea and well within the acuity limits for discrimination, as shown by the speed of detecting the feature targets.

The only large effect of density was shown on feature negative trials, where the decision that no target was present was made much more slowly with the sparse than with the dense displays. Subjects may have been less confident that the target was absent when the distractors were peripheral than when they were centrally located. The increased time in this condition could reflect increased eye movements and fixations. The fact that feature targets, when they were present in the sparse displays, were still detected easily and with little effect of display size shows that the slow checking of negative displays was actually unnecessary. However, it appears to be a typical strategy with feature search under difficult or confusable conditions (cf. Treisman & Gelade, 1980, Experiments 3 and 4).

The absence of much effect of density on conjunction search contrasts with the substantial effect of discriminability that we obtained in an earlier experiment (Treisman & Gelade, 1980). The speed of identifying the features that define a conjunction target should affect each serial check in the conjunction search and should therefore be reflected in slope differences. The results confirmed this prediction: The rate for conjunctions of easy features (red, green, O, and N) was more than double the rate for conjunctions of more difficult features (blue, green, X, and T). On the other hand, we have little reason for the theory to expect slope differences in the present experiment where only spatial density was varied, since this should affect the time taken for each serial check only if the shapes (X and H) or the colors (red and green) were harder to distinguish in the outside positions of sparse displays.

The results are consistent with Ambler et al.'s (1978) most recent findings, and not with the earlier ones of Ambler and Finklea

(1976), and extend their conclusion from line arrangement to conjunctions of values on the different dimensions of shape and color. The conclusions may also be relevant to an alternative account of errors in object perception proposed by Wolford (1975) in his feature perturbation model. Wolford suggested that features become locally displaced over time and produce errors of identification among neighboring items. Since these perturbations are local, they should affect dense displays more than the sparse ones, making them more prone to error and therefore inducing slower, more careful search. The fact that search rates hardly differed in our experiment suggests that spatial distance between features may be irrelevant. However, one would need to vary density and distance from the fovea independently of each other before drawing any firm conclusion, since the two variables may have opposite effects.

#### Experiment 4: The Group Boundary Effect on Visual Search for Conjunctions and Features

The final prediction we test in this article is the most direct test of the idea that preattentive organization exists only *within* dimensions—within a color map, within a shape map, within a map of movements or orientations—and that these maps are related to each other where and when attention is focused.

If the generally accepted view were correct, the early level of preattentive organization should result in a single, global configuration to which all the different dimensions contribute according to their salience. A target should emerge spontaneously whenever it is locally salient as figure against ground. On this view, the conjunction targets were difficult to find in our earlier experiments (Treisman & Gelade, 1980) because the preattentive organization was too complex; the colors and shapes were randomly intermingled. If, however, we introduce locally homogeneous groups within which a conjunction target contrasts on a highly discriminable feature, it should "pop out" by automatic preattentive segregation. The re-

sults so far described refute this prediction. The conjunction target did not emerge automatically from grouped displays, even when the groups were few and the target very salient in its local context. Feature-integration theory, on the other hand, precludes conjunction "pop out." If texture-segregation occurs only *within* each separate feature map, focal attention will be required before the different feature maps (e.g., the color map and the orientation map) can be integrated or interrelated to locate the unique conjunction of color and shape.

The further possibility then arises that we might effectively camouflage an object by placing it at a boundary between two groups, each of which shares one of its features. We can choose an object that, within either group alone, would be quite salient and see if adding the second group makes it harder to see. For example, we can place a red X at the boundary between a group of red Os and a group of blue Xs. Figure 5 shows an example. In order to detect the presence of the red X at the boundary, we predict that attention must be narrowed down to exclude the adjacent red Os and blue Xs and to focus on the item itself, although in either group alone it would be salient and easy to detect. When the target is in the center of a group, attention need be narrowed only to the group which contains it. Thus detection should be faster for a conjunction target placed in the center of a group than at the boundary between groups, although it should be slower in both cases than with one group only. On the other hand, if the target is defined by a unique feature (the color green or vertical lines), we would expect its detection to be independent of its position relative to the boundary.

The displays in this experiment consisted of just six items arranged in a row. There were always two, three, or four blue Xs and four, three, or two red Os homogeneously grouped. In half the displays one item in a group of three or four was replaced either by a conjunction target (a red X or blue O) or by a feature target (a red or blue H or a green X or O). The task was to detect whether the display contained an odd item that was not a blue X or a red O. The target was placed either at the boundary between

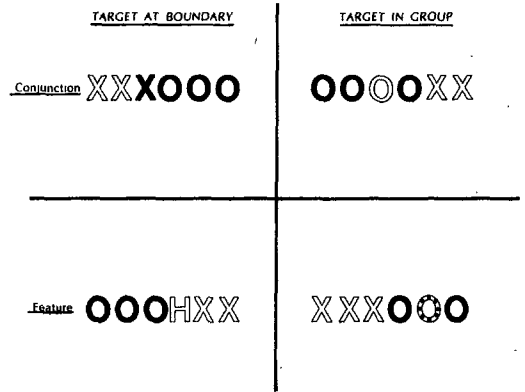


Figure 5. Examples of displays used in Experiment 4. (The solid letters represent red, the outline letters blue and the dotted letter green.)

the blue Xs and red Os or one position away from the boundary.

### Method

**Subjects.** There were 10 subjects (6 women and 4 men), all students at the University of British Columbia, who volunteered for the experiment and were paid \$3 an hour.

**Stimulus material.** Each display contained a row of six colored letters, each subtending  $1.18^\circ$ , with the whole row subtending  $8.5^\circ$ . There were 30 negative displays, equal numbers containing two, three, and four blue Xs on one side of a boundary and four, three, and two red Os on the other. Half of each type had the red Os on the left and half on the right. The positive displays were divided into four categories, depending on whether the target was a conjunction target (red X or blue O) or a feature target (green X, green O, blue H or red H) and on whether the target was next to the boundary or one item away from it. The target occurred equally often in positions 2, 3, 4, and 5, and never in positions 1 and 6. It never replaced one of a group of two distractors, since this would have resulted in two unique items for that display. These variations produced 8 displays for each of the four categories of positive display. Thus the complete pack, with the 30 negative displays, consisted of 62 displays.

**Procedure.** The task was to decide, as quickly as possible without making errors, whether the display contained a unique item (i.e., one which was neither a blue X nor a red O). Subjects pressed the right-hand key if it did and the left-hand key if it did not. They were told what the six unique items could be and were shown examples of the positive and negative displays. They were given a few trials for practice before beginning the experiment proper. Each subject was tested on four blocks of 62 trials.

### Results and Discussion

The mean reaction times and the mean percentage of errors are shown in Table 4.



Reaction times that were more than 3 standard deviations from the mean for a given subject in a given condition were discarded; these averaged .5%. An ANOVA showed significant effects of target type: conjunction versus feature,  $F(1, 9) = 34.28, p = .0002$ ; of target position (at the boundary or centered in one group),  $F(1, 9) = 14.97, p = .004$ , and a significant interaction between target type and target position,  $F(1, 9) = 11.72, p = .0076$ . Newman-Keuls tests were carried out to compare specific pairs of means and are reported in context below.

There are two main findings: first, a camouflaging effect of placing a conjunction target at the boundary between groups; second, the fact that even within a homogeneous distractor group, the conjunction target is detected more slowly than the feature targets.

Conjunction targets at the boundary were detected 135 msec more slowly than feature targets at the boundary ( $p < .01$ ). Conjunction targets were also more slowly detected at the boundary than in the center of a group ( $p < .01$ ), whereas for feature targets there was no difference. Thus a unique feature, whether color or shape, can be detected equally well at a boundary between two dif-

ferent groups and when placed between identical items. A conjunction target, on the other hand, is harder to detect when placed at a boundary between items that on one side share its color and on the other its shape. Subjects missed it altogether on 9.3% trials (compared to a mean of 2.6% for all other conditions) and took considerably longer (104 msec) to find it when they were successful. Phenomenologically, the conjunction target is effectively camouflaged at the boundary unless attention is directed specifically to its location. There appear to be two competing ways of preattentively grouping this display—one within the color map and one within the shape map. The conjunction target exists in neither, although the feature target is always unique in one of the two. In order to see the conjunction target, the two automatic groupings must be broken. This involves relating shapes to colors and appears to require an extra operation that takes time and focuses attention specifically on the location of the unique conjunction.

A second inference can be made by comparing detection of a conjunction target within a group with detection of a feature target within a group. Even when the target is removed from the group boundary, it is

Table 4  
*Mean Reaction Times, Standard Deviations, and Percent Errors in Experiment 4*

Display	Reaction times			% Errors		
	At boundary	In group	Difference	At boundary	In group	Difference
Conjunction						
<i>M</i>	769	665	104	9.3	3.3	6.0
<i>SD</i>	262	195				
Feature						
<i>M</i>	634	623	11	2.1	2.4	-3
<i>SD</i>	206	203				
Feature shape						
<i>M</i>	625	616	9	1.8	1.2	.6
<i>SD</i>	198	179				
Feature color						
<i>M</i>	643	630	13	2.4	3.6	-1.2
<i>SD</i>	216	230				
Negatives						
<i>M</i>	681			3.3		
<i>SD</i>	265					

detected more slowly when defined by a conjunction of features than by a unique feature ( $p < .05$ ). This is unlikely to result from a difference in discriminability at the feature level: The conjunction features (red and blue and X and O) are less confusable with each other on average than they are with green or with H. If discriminability at the feature level had been matched, we would expect, if anything, a larger advantage of the feature over the conjunction targets.

Thus, the presence of distractors elsewhere in the display, which share the locally distinctive feature of the conjunction target, forces attention to narrow down at least sufficiently to exclude them. This conclusion was also implicit in the results of the earlier experiments. In Experiment 2, for example, the time to find a green H in an otherwise homogeneous group of eight red Hs was more than 800 msec longer when there were four other groups of red Hs and four of green Xs present in the display than when the single group of eight red Hs with its green H target was presented alone (in an unpredictable location). Here again the presence, even spatially segregated some distance away from the target, of distractor letters that shared the target color acted as potent camouflage, destroying the salience of the local color contrast, which would otherwise have made the target impossible to miss. The conjunction target apparently does not "pop out" of the display as a whole, however homogeneous its local context.

Although grouping variables had little or no effect on feature targets over the range of conditions we tested here, it would probably be possible to devise displays in which they did: For example, in a display with 60% red and 20% each of blue and of yellow distractors, we could probably make a green target more salient by embedding it in a homogeneous group of red distractors than by embedding it in a mixed group of blue and yellow distractors. However, local contrast and discriminability interactions might account for all or part of this difference. At present, the claim should simply be that the effects of grouping are both greater and also differently mediated when conjunctions must be specified than when features will suffice.

## General Discussion

We began with three premises: (a) Texture segregation and perceptual grouping are early preattentive processes, which are designed to specify candidates for object identification (Neisser, 1967); (b) perceptual groups are organized separately within each dimension, and the different configurations formed within the color and the shape maps are interrelated only when attention is narrowed onto an item or group; (c) focused attention is necessary to identify objects whenever these must be specified by conjunctions of separable features (Treisman & Gelade, 1980). This article describes four experiments that relate these premises by exploring the effects of spatial configurations on search for objects and for features.

We infer a link between grouping and the distribution of attention from the pattern of search times produced by different spatial arrangements of targets and distractors. Variations that changed the predicted risk of illusory conjunctions affected search times, whereas those that did not had no effect. Detection of feature targets was, in general, independent of the spatial arrangement of distractors, although negative trials were slower when distractors were more distant from the fovea. Search for conjunction targets was serial across items when items were randomly mixed and was hardly affected by their spatial density. It is unlikely, therefore, to be causally determined by eye movements, although it is certainly the case that subjects moved their eyes during the 1,200 and 2,600 msec they spent searching the displays of 16 or 36 items and presumably moved them more often during the longer search latencies than during the shorter ones. Perceptual grouping, on the other hand, had a very substantial effect on conjunction search. Search appeared to be serial across groups of items whenever the display contained spatially homogeneous groups. We predicted and explain this result by the claim that when the structure of the display allows attention to be spread over several homogeneous items without risking an illusory target, the features of those items will be checked in parallel. Prinzmetal (1981) has obtained

converging evidence consistent with this hypothesis. He found that subjects were more likely to interchange features (e.g., moving a diameter from one circle to another) within a perceptual group than between perceptual groups, even when the physical distance was greater within than between groups. His displays were too brief to allow serial attention to each item in turn. If attention were attracted, at least on some proportion of trials, to one or other perceptual group, we would expect illusory conjunctions to form more frequently within than across the group boundaries.

The contrast between attending to single items and attending to groups is clearly related to the difference between local and global processing. When we attend to the global configuration, each local element loses its separate identity, except insofar as it contributes to the identity of the whole. Thus, for example, the location of each dot in a row is specified at the global level only by the fact that it is colinear with the others. A change of processing level entails a change in the encoding of the same physical stimulus. Thus a change in the direction of curvature of the mouth becomes a change from a sad to a smiling face; at this global level the identity of the local cue may be lost. However, the existence of emergent properties (Pomerantz, Sager, & Stoeber, 1977) at the global level should place constraints on the acceptability of illusory conjunctions between local features within an attended group of items. Feature exchanges that would alter global as well as local units should more often be rejected than those that leave the global encoding unchanged. Thus a red X might be easier to detect and less likely to exchange its color or shape when it is at a uniquely specified point in a global configuration.

The claim that preattentive organization occurs only within and not between dimensions explains why conjunction targets are invisible until attention is narrowed down to the local group within which they possess a unique, distinctive feature. It also predicts that embedding a conjunction target between two competing groups should make it more difficult to detect. Experiment 4 de-

monstrated this conjunction camouflage. Although in either group alone the target would be salient (a red H with red Os or a blue O with red Os), adding the second group effectively concealed it at the preattentive level. The results suggest that we have many separate preattentive worlds, each organized along its own lines and each offering its own particular candidates for objecthood. In normal life these worlds will tend to coincide, since physical objects have highly correlated feature boundaries. If ever they disagree, we may discover the truth only when we turn our attention to the misleading cues.

It would be interesting to know whether the different feature maps are simultaneously accessible. One way to test this is to see whether subjects detect a feature target more quickly when they know which map to consult than when either could reveal the target. In a supplementary experiment, eight new subjects were tested on the feature cards from Experiment 4 in two different conditions, one with the H and the green targets randomly mixed and one with the H and the green targets separated in different blocks. When subjects knew which feature to look for (green or H) they found it 68 msec faster than when they were looking for either,  $F(1, 7) = 18.49, p = .005$ . However, there was no interaction between known versus unknown feature and the position of the target, suggesting that attention in the spatially selective sense was not involved. Although the features within both the shape and the color map are registered in parallel, the two dimensions may compete or alternate with each other in their access to decision and response.

What is the relation between the preattentive grouping that segregates potential objects for focused attention and the phenomenal experience of a perceptually organized world? It follows from the present account that preattentive parsing cannot be available to consciousness. We do not experience shapeless color or colorless shapes but always some combination of the two. Neisser (1967) appeared to equate preattentive organization with the representation that we consciously experience of regions

outside the current focus of attention. Thus he suggested that preattentive processing guides our movements and creates a global impression of our surroundings. Other writers have also operationally equated preattentive with phenomenological organization. For example, Banks and Prinzmetal (1976) asked subjects to indicate the phenomenal grouping that they experienced in a display and correlated this with the efficiency of search for target items in the same display. It seems empirically plausible, although not logically necessary, that preattentive feature segregation should also determine experienced organization. However, our claim is that the separate preattentive feature maps are not themselves directly accessible to conscious awareness. Conscious access must follow a level of resynthesis at which colors are allocated to shapes and movements to objects, however imprecisely defined these features may be. Whether the combinations are correct or not will depend (a) on whether attention was focused on the item or groups in question and (b) whether, in the absence of attention, prior expectations and knowledge of the context can select the possible or plausible features to conjoin.

A final question concerns the nature of the attention spotlight. What are the constraints on the shape it can adopt? Can it focus on a concave or indented shape like a cross, or must the edges of the spotlight be convex? Can it lock onto a group of moving stimuli or continuously vary its shape to track a changing configuration of elements over time? Can it in fact be spatially split when its target is a moving group interspersed with stationary stimuli? Does it conform to all the Gestalt grouping principles or is it constrained by rules of its own? If the function of both preattentive grouping and focused attention is the same—to specify objects—we might expect the same laws to determine both. The paradigm of conjunction search could perhaps be used to throw more light on these issues.

### References

- Ambler, B., & Finklea, D. L. The influence of selective attention in peripheral and foveal vision. *Perception & Psychophysics*, 1976, 19, 518-526.
- Ambler, B. A., Keel, R., & Phelps, E. A foveal discriminability difference for one vs. four letters. *Bulletin of the Psychonomic Science*, 1978, 11, 317-320.
- Banks, W., & Prinzmetal, W. Configurational effects in visual information processing. *Perception & Psychophysics*, 1976, 19, 361-367.
- Beck, J. Effect of orientation and of shape similarity on perceptual grouping. *Perception & Psychophysics*, 1966, 1, 300-302.
- Beck, J. Perceptual grouping produced by line figures. *Perception & Psychophysics*, 1967, 2, 491-495.
- Beck, J. Similarity grouping and perceptual discriminability under uncertainty. *American Journal of Psychology*, 1972, 85, 1-19.
- Beck, J., & Ambler, B. The effects of concentrated and distributed attention on perceptual acuity. *Perception & Psychophysics*, 1973, 12, 482-486.
- Caelli, T. M., Julesz B., & Gilbert, E. N. On perceptual analyzers underlying visual texture discrimination: Part II. *Biological Cybernetics*, 1978, 29, 201-214.
- Eriksen, C., & Hoffman, J. Temporal and spatial characteristics of selective coding from visual displays. *Perception & Psychophysics*, 1972, 12, 201-294.
- Farmer, E. W., & Taylor, R. M. Visual search through color displays: Effects of target-background similarity and background uniformity. *Perception & Psychophysics*, 1980, 27, 267-272.
- Fox, J. R. Continuity, concealment and visual attention. In G. Underwood (Ed.), *Strategies of information processing*. New York: Academic Press, 1978.
- Garner, W. R. *The processing of information and structure*. Potomac, Md. Erlbaum, 1974.
- Julesz, B. Experiments in the visual perception of texture. *Scientific American*, 1975, 232(4), 34-43.
- Julesz, B. Spatial nonlinearities in the instantaneous perception of textures with identical power spectra. In C. Longuet-Higgins & N. S. Sutherland (Eds.), *The psychology of vision. Philosophical Transactions of the Royal Society, London*, 1980, 290, 83-94.
- Julesz, B. & Caelli, T. On the limits of Fourier decomposition in visual texture perception. *Perception*, 1979, 8, 69-73.
- Kahneman, D., & Henik, A. Effects of visual grouping on immediate recall and selective attention. In S. Dornic (Ed.), *Attention and performance VI*. Hillsdale, N.J.: Erlbaum, 1977.
- Kahneman, D., & Henik, A. Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981.
- Kubovy, M. Concurrent-pitch segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981.
- Laberge, D. Attention and the measurement of perceptual learning. *Memory & Cognition*, 1973, 1, 268-276.
- Neisser, U. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.
- Olson, R., & Attneave, F. What variables produce similarity grouping? *American Journal of Psychology*, 1970, 83, 1-21.
- Pomerantz, J. R., Sager, L. C., & Stoever, R. J. Per-

- ception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3, 422-435.
- Prinzmetal, W. Principles of feature integration in visual perception. *Perception & Psychophysics*, 1981, 30, 330-340.
- Shiffrin, R., & Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 1977, 84, 127-190.
- Townsend, J. T. Some results on the identifiability of parallel and serial processes. *British Journal of Mathematical and Statistical Psychology*, 1972, 25, 168-199.
- Treisman, A., & Gelade, G. A feature integration theory of attention. *Cognitive Psychology*, 1980, 12, 97-136.
- Treisman, A., & Schmidt, H. Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 1982, 14, 107-141.
- Treisman, A., Sykes, M., & Gelade, G. Selective attention and stimulus integration. In S. Dornic (Ed.), *Attention and performance, VI*. Hillsdale, N.J.: Erlbaum, 1977.
- Wertheimer, M. Untersuchungen zur lehre von der Gestalt. *Psychologische Forschung*, 1923, 4, 301-350.
- Wolford, G. Perturbation model for letter identification. *Psychological Review*, 1975, 82, 184-199.

Received June 26, 1981 ■