Density Effects in Conjunction Search: Evidence for a Coarse Location Mechanism of Feature Integration

Asher Cohen
Indiana University

Richard B. Ivry
University of California, Berkeley

Four experiments used the visual search paradigm to examine feature integration mechanisms. Reaction time to determine the presence or absence of a conjunctive target is relatively fast and exhaustive for low-density displays. Search rate is slow and self-terminating for high-density displays. Density effects do not arise when the target is defined by a unique feature. Two mechanisms are proposed for feature integration. A fast mechanism integrates features on the basis of coarse location information coded with the initial registration of the features. This coarse location mechanism requires that display items be spaced apart. A second, slower mechanism is used when objects are clumped together. The 2-mechanism hypothesis provides a resolution to conflicting findings in the visual search and illusory-conjunction literature. A possible interpretation of the findings with a single guided search mechanism for feature integration is also discussed.

Multiple lines of research suggest that processing of visual objects is done in at least two processing stages (for comprehensive reviews, see Livingstone & Hubel, 1987; Treisman, 1986). First, features such as color, line orientation, and direction of motion are processed in parallel across the visual scene. Second, the particular visual features of each object in the scene are conjoined.

The analysis of the visual features at the early stage of processing creates a problem for the visual system, often called the *binding problem* (Crick, 1984; Hinton, McClelland, & Rumelhart, 1986; Keele, Cohen, Ivry, Liotti, & Yee, 1988). Given that features at the first stage are processed without regard to the objects from which these features originate, how does the visual system correctly bind the features of an object rather than bind features of different objects? What is the mechanism that prevents the formation of illusory conjunctions between features of different objects (Cohen & Ivry, 1989; Treisman & Schmidt, 1982)?

Treisman and her colleagues (e.g., Treisman & Gelade, 1980; Treisman & Gormican, 1988) have suggested that binding involves a focal attention mechanism that operates as follows: The mechanism marks a specific spatial region. All of the features within that area are conjoined. Given sufficient time, the focal attention mechanism can correctly conjoin features of a given object even if this object is located near other objects by marking the precise area in which the target object is located. If the focus of the binding mechanism is not sufficiently constricted to span a single object, however, fea-

tures of separate objects will not be differentiated inside the marked area, and illusory conjunctions may occur. Cohen and Ivry (1989) provided evidence for the existence of such a mechanism.
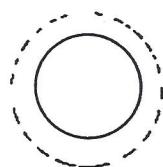
In this article we present evidence in support of a second mechanism for feature integration. The essence of this mechanism is that coarse location information is coded with the initial registration of visual features (see Cohen & Ivry, 1989, for review of the literature on this topic). The visual system can use this information to bind features. Because the location information is imprecise, it can only be used in situations in which objects are not crowded together in the visual field. The process for conjoining color and shape is illustrated in Figure 1. Assume that the dotted circle represents the coarse location information perceived with a shape feature (e.g., line orientation), and assume that the solid circle represents the perceived location information of a color feature. Figure 1a shows the coarse location representation of the shape and color features when only one object is present in the visual field. Because the color and shape features are located at the same position, there is an overlap between the location information of the two features. Figure 1b shows a situation in which two spatially distant objects are present in the scene. The coarse location information of the color and shape features is sufficient to separate the features of one object from the features of the other object. Thus, binding can occur on the basis of the overlap in location information. The notion is that following the registration of the features with their coarse location information, the coarse location mechanism detects features with overlapping location information and binds them. Figure 1c shows a situation in which two spatially adjacent objects are present in the scene. Here, the coarse location information is not sufficient to separate the features of one object from the features of the other object. Under these conditions, the slower focal attention mechanism will be used to separate the features of one object from the features of its neighbor.

The idea of the coarse location mechanism has useful implications. Conjoining features that are based on coarse
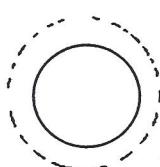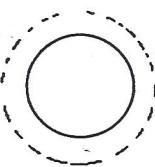
Correspondence concerning this article should be addressed to Asher Cohen, Department of Psychology, Indiana University, Bloomington, Indiana 47405. (Electronic mail: ascohen @ ucs.indiana.edu)

A: A SINGLE OBJECT

B: TWO DISTANT OBJECTS

C: TWO ADJACENT OBJECTS
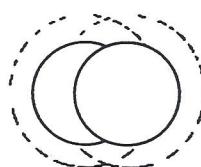
Figure 1. An illustration of the operation of the coarse location mechanism (see text for details). (The dotted circle is larger only for expository reasons.)

location information is a fast method for feature binding because the information is assumed to be available during feature registration. The disadvantage of this mechanism is that it cannot be used when objects are spatially adjacent (Figure 1c). In these situations, the slower focal attention mechanism is required to ensure correct binding.

Evidence for the existence of the coarse location mechanism comes from research on illusory conjunctions. In this method, two or more objects are simultaneously presented for a brief duration. When subjects make mistakes in these studies, they tend to combine features from different objects in the display rather than report a feature that was absent from the display. When the objects are presented inside the focus of attention, illusory conjunctions between features of different objects occur regardless of the actual distance between the objects (Cohen & Ivry, 1989; Treisman & Schmidt, 1982). This finding is one of the main sources of evidence for the existence of the focal attention mechanism of feature integration mentioned earlier. When the objects are presented outside the focus of attention, however, illusory conjunctions occur only between adjacent objects (Cohen & Ivry, 1989; Keele et al., 1988; Snyder, 1972). The hypothesis of a coarse location mechanism provides an explanation for how features are integrated into objects when the objects are located outside the focus of attention. The mechanism will produce correct

conjunctions when objects are sufficiently distant from each other. When objects are adjacent to each other, illusory conjunctions will occur.

Our main goal in the present research is to provide stronger evidence for the existence of the coarse location mechanism. Our claim is that the visual system can conjoin the features of an object when this object is not located near other objects. To test the generality of this hypothesis, we use the visual search paradigm. Moreover, we show that this hypothesis can account for a number of discrepant results in this literature. We first briefly review previous findings in visual search.

In the visual search method, subjects have to detect whether a target is present among a varying number of distractors. Numerous experiments have shown that when the target is defined by a unique feature, the reaction time for its detection is minimally affected by the number of distractors in the field. When the target is composed of a conjunction of two features (e.g., a red vertical line in a display containing red horizontal lines and green vertical lines), however, the search time increases linearly with the number of distractors in the field. Furthermore, the ratio of the slope for target-absent trials (i.e., when there is no target in the visual field) is twice as large as the slope for target-present response (e.g., Egeth, Virzi, & Garbart, 1984; Ivry & Cohen, 1990; Nakayama & Silverman, 1986; Treisman & Gelade, 1980; Treisman & Gormican, 1988). These findings demonstrate that detecting a conjunction of features is done by a limited-capacity system.[1] The 2:1 ratio of negative to positive slopes suggests that the search is self-terminating. On target-present trials, search ends when the target is detected. Thus, only half of the items will be processed on average. When the target is absent, the entire array must be examined before a decision is made.

A number of recent results, however, challenge the generality of a slow, self-terminating search process with a conjunctively defined target (Houck & Hoffman, 1986; McLeod, Driver, & Crisp, 1988; Nakayama & Silverman, 1986; Pashler, 1987; Wolfe, Cave, & Franzel, 1989). Interestingly, these findings, though at odds with previous findings, are not always in accord with one another. Pashler (1987) reported a slope ratio of close to 1:1 in a conjunction search of color and shape with displays up to eight items. A 2:1 ratio was obtained when there were more distractors in the display. Furthermore, the search time per item was faster in the small-display-size condition than in the large-display-size condition. Pashler suggested that up to about eight objects can be conjoined in parallel and without capacity constraints. Wolfe et al. (1989) found similar display-size effects when using a small number of stimuli; however, they reported that when there were more distractors in the field (up to 32) the reaction time to the target was even less affected by the distractors. In addition to the inconsistencies among these results, alternative models proposed by Pashler and Wolfe et al. do not deal with the

---

[1] The typical interpretation of these results is that detecting a conjunction of features is done serially, one object at a time. As was shown repeatedly by Townsend (e.g., Townsend, 1976, 1990), however, this method cannot distinguish between a limited-capacity parallel system and a serial system.

findings of illusory conjunctions reviewed earlier (Cohen & Ivry, 1989; Treisman & Schmidt, 1982).

The coarse location hypothesis can account for these apparently discrepant results. Wolfe et al. (1989) presented the stimuli in such a way that the distance between items was always greater than 1°. With this spacing, the coarse location information is sufficient for correctly conjoining the features of each object (Cohen & Ivry, 1989). In Pashler's (1987) research, the distance between the objects was greater than 1° only when there were eight or fewer items in the field. In these cases, the search was only minimally affected by the number of distractors. When the number of distractors was increased, however, the distance between neighboring items was decreased, rendering the coarse location information insufficient for accurate binding.[2] The visual system had to use the attentional binding mechanism, which is limited in capacity, and this resulted in slow search times in which the reaction time was a function of the number of distractors. The experiments reported in this article test this hypothesis.

Experiment 1

The first experiment was designed to show that a relatively flat search function and a slope ratio of 1:1 between positive and negative trials with small array sizes are obtained only with displays in which the objects are spread apart. Subjects were asked to detect a target among a variable number of distractors. The array sizes were 2, 4, and 8. There were two main conditions. In one condition, the items were spatially spread in such a way that the distance between any two items was greater than 1° of visual angle. We refer to this as the *spread condition*. In the other condition, the items were arranged so that the distance between neighboring items was 0.62° of visual angle. We refer to this as the *clump condition*. According to our hypothesis, the coarse location mechanism is sufficient for correct binding in the spread condition, and thus the search functions will be relatively flat. The resolution of location information is not sufficient for correct binding in the clump condition. We therefore expect slower search-rate functions in this condition.

The items were positioned in an imaginary circle around a fixation point. Although this method keeps eccentricity constant, it creates another difference between the spread and the clump conditions. Because the stimuli in the clump condition are adjacent to each other, the total space occupied by the stimuli in this condition is smaller for any given display size than the total area occupied by the stimuli in the spread condition. This difference in total area ought to make the task harder in the spread condition, however. Thus, it ought to work against the coarse location hypothesis, which states that search will be harder in the clump condition.

Recently, Klein (1988) showed that following a response in a visual search task, subjects are slow to perform a luminance-detection task when the luminance target appears in a location previously occupied in the search task. This inhibition was strongest in trials in which no target was present in the visual search task. Subjects did not display this form of inhibition when the visual search task was easy and not affected by the number of distractors. Our hypothesis is that the coarse location mechanism, which is quick and results in a fast search rate, will be used in the spread condition and that the focal attention mechanism, which is slow and results in a slow search rate, will be used in the clump condition. Thus, it is possible that subjects will display inhibition in the clump condition but not in the spread condition. We wished to examine this possibility. After most of the trials in the visual search task, we used a simple luminance-detection task to examine this question.

Method

Subjects. Fourteen undergraduate psychology subjects from the University of Oregon participated as part of their course requirements. All of the subjects had normal or corrected-to-normal vision and no known color-perception deficiency.

Apparatus and stimuli. The stimuli were presented on an Amdek color monitor controlled by an Apple IIe computer. Subjects viewed the display from a distance of 116 cm. A chin rest was used to fix this distance. The stimuli consisted of an asterisk and the letters X and O. An individual character in the display subtended approximately 0.62° in height and 0.46° in width. The two letters were presented in one of two colors: yellow or blue.

Design. The subjects performed a visual search task followed in 80% of the trials by a simple luminance-detection task. The target in the visual search task was a yellow X, and the distractors were yellow Os and blue Xs. The subjects' task was to determine as quickly as possible whether a yellow X was present. Each display consisted of either two, four, or eight letters. In half of the trials (target-present trials), a single yellow X was presented with the distractor or distractors. In the target-absent trials, the items were all distractors. The letters were positioned on an imaginary circle that had a radius approximately 2.5° from the center. There were 24 possible locations. The items in the clump condition were positioned in adjacent locations. The distance between the center of one item and the center of a neighboring item was approximately 0.62°. In the spread condition, the positions of the items were selected in such a way that the distance between adjacent letters was at least 1.24°. The display items were selected with the constraint that there could not be runs of three identical colors or letters in spatially adjacent locations.

Spread and clump trials were presented in mixed blocks. Each subject participated in 12 conditions (2 [clump, spread] × 3 [display size] × 2 [present, absent]). The subjects first received a short practice block with 24 trials and then performed two experimental blocks consisting of 288 trials, or 24 trials per condition.

On 80% of the trials, the response to the visual search task was followed by a luminance task. The display for this task contained a single asterisk presented at one of the 24 locations on the imaginary circle. On 50% of these trials, the asterisk appeared in a location occupied by one of the items in the preceding visual search trial. For the remaining 50% of the luminance trials, the asterisk appeared in a location that was empty in the visual search trial. The luminance task required subjects to respond as quickly as possible after the onset

---

[2] Pashler (1987) ran an additional experiment with large array sizes (Experiment 3) in which the stimuli could be spread apart on some of the trials. The array sizes in this experiment included 2, 4, 8, 16, or 24 items. The items were placed randomly on a grid with 36 possible positions. Note that as the array size increases there is a higher probability that at least some items will be adjacent to each other because there was no constraint on the possible locations of the letters. In this respect, then, this experiment was similar to Pashler's other experiments.

of the stimulus. No asterisk was presented on 20% of the trials: These served as catch trials.

Procedure. At the beginning of each trial, an asterisk, serving as a fixation point, was presented on the center of the visual screen. After 1,000 ms, the asterisk was replaced by the visual search display. The display remained visible until the subject responded. Subjects were instructed to respond as fast as they could while minimizing their mistakes. Subjects responded by pushing a two-key device in which a right keypress indicated target present and a left keypress indicated target absent. Incorrect responses were followed by a short tone. Subjects were instructed to expect occasional errors because of the emphasis on speeded responses. The tones provided a means for self-monitoring error rates. The screen went blank immediately following the subject's response in the visual search task. The asterisk for the luminance-detection task appeared 500 ms later on 80% of the trials. The subjects pressed a single key with the index finger of their left hand as soon as they detected the luminance stimulus. They were told that the asterisk would not be presented on some trials so they ought to avoid anticipating the response. The intertrial interval was 2,000 ms.

Results and Discussion

Visual search. Figure 2 presents the reaction time data for correct responses on the visual search task. The data were entered into a 2 (spatial position: clump vs. spread) × 2 (target: present vs. absent) × 3 (array size: 2, 4, or 8) analysis of variance. The analysis revealed significant main effects for all three variables. Responses were slower in the clump condition than in the spread condition, $F(1, 13) = 139.04$, $p < .05$. Responses were also slower when the target was absent in comparison to target-present trials, $F(1, 13) = 74.5$, $p < .05$, and responses got slower as the array sizes increased, $F(2, 26) = 74.24$, $p < .05$. All of the two-way interactions and the three-way interaction were significant as well: Spatial Position × Target, $F(1, 13) = 6.24$, $p < .05$; Target × Array Size, $F(2, $

$26) = 18.86$, $p < .05$; Spatial Position × Array Size, $F(2, 26) = 15.78$, $p < .05$; Target × Array Size × Spatial Position, $F(2, 26) = 10.5$, $p < .05$.

These results are in accord with our predictions: Search for a conjunctively defined target was slower in the clump condition. This effect was qualified by the array size. The difference between the clump and the spread conditions increased with an increase in the array sizes. In other words, the slope of the function that related array size to search time was steeper in the clump condition. Finally, the effects of both spatial position and array size were also qualified by the target condition. The ratio of the slope of the target absent to target present was approximately 2:1 in the clump condition (47 ms in the target-absent trials and 25 ms in the target-present trials). The ratio in the spread condition was lower (24.4 ms in the target-absent trials and 15.8 ms in the target-present trials). Note, however, that this value is higher than the 1:1 ratio obtained by Pashler (1987). This issue is discussed further in Experiment 3.

In each array size, the percentage of errors in the clump condition was larger than the percentage of errors in the spread condition. Therefore, the difference in reaction time between the clump and the spread conditions cannot be attributed to a speed–accuracy trade-off.

Luminance-detection task. The reaction time to the detection of the asterisk following the visual search task is presented in Table 1. The results are presented as a function of the preceding visual search trial and whether the luminance target was at an on or off location. The data are collapsed across array sizes. By following the findings of Klein (1988), we expected the reaction time to be slower to a target in the clump condition when the target appeared in a location previously occupied by one of the visual search stimuli. Furthermore, we expected this inhibition to be strongest for the target-absent trials; however, inhibition was not obtained for
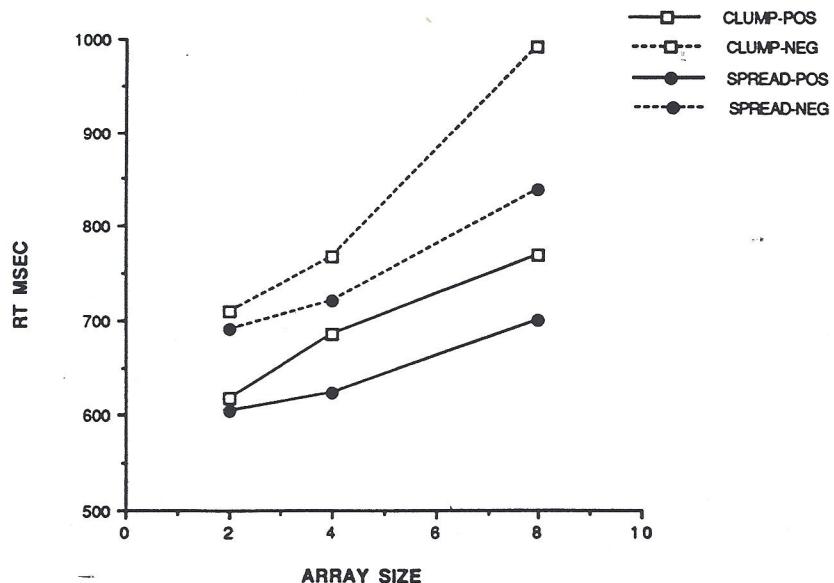

Figure 2. Mean reaction time for the visual search task of Experiment 1.

Table 1

The Reaction Time for the Luminance Task of Experiment 1

Condition	Clump		Spread	
	Target present	Target absent	Target present	Target absent
On	293.05	301.07	301.38	301.27
Off	283.18	291.31	284.03	301.63

Note. On represents the conditions under which the target appeared in a location occupied on the same trial by an item in the visual search task. Off represents the conditions under which the target appeared in a location not occupied on the same trial by an item in the visual search task.

any of the conditions. No significant effects were obtained in an analysis of variance with spatial position (spread vs. clump), luminance target position (on vs. off), and search target (present vs. absent) as variables.

There are a number of differences between our task and the task used by Klein (1988) that could account for the discrepancy between the results. For example, the difference between the target and the distractors was confined to elements of shape, whereas the target in our experiment was composed of a conjunction of color and shape. In addition, the time between the response to the visual search and the appearance of the target in the luminance-detection task was different in the two studies. These differences could account for the discrepant results. Furthermore, recent attempts by Wolfe and Pokorny (1990) to replicate Klein's results under conditions similar to those used by Klein were unsuccessful. This issue is not pursued further in this article.

Our main finding in Experiment 1 is that there appears to be a strong difference in visual search between the clump and the spread conditions. The search rate in the clump condition was slow, and the ratio of target-absent to target-present trials was approximately 2:1, as in the early studies of visual search (e.g., Egeth et al., 1984; Treisman & Gelade, 1980). On the other hand, search rate in the spread condition was faster, and the ratio of target-absent to target-present trials was less than 2:1. These results can be explained by our hypothesis that the coarse location mechanism can be used in the spread condition but not in the clump condition. An alternative explanation of these results, however, is that the clump condition was slower because of lateral masking between neighboring objects. We examine this possibility in Experiment 2.

Experiment 2

Our explanation and the lateral masking explanation make different predictions concerning the perception of individual features (i.e., the various colors and shape primitives). We claim that the difference between the clump and spread conditions is due to the operation of the coarse location mechanism for feature integration in the latter condition only. Our hypothesis assumes that registration of the individual features is the same in both conditions. A lateral-masking account predicts that the difference is due to impaired perception of the individual features in the clump condition because of the interference between the adjacent objects. Interference

may be between letters, colors, or both. If the lateral-masking explanation is valid, there ought to be similar differences between the clump and the spread condition even when the visual search task is composed of detection of a single feature. We examine this possibility in this experiment.

Method

The visual search task was similar to the one used in Experiment 1, with one major difference. The task involved determining the presence or absence of a simple feature rather than detecting a conjunctively defined target. One group of 9 subjects had to detect the presence of yellow among blue distractors. The shape of the stimuli in this condition for both target and distractors could be either X or O and was determined randomly for each item in the display. A second group of 9 subjects had to detect the presence of an X among O distractors. All of the letters in this condition could be either blue or yellow; color was determined randomly for each item in the display. The only other difference between the present experiment and Experiment 1 is that the two experimental blocks in this experiment consisted of 180 trials each.

Results and Discussion

The reaction time to the color task is presented in the upper part of Figure 3. The reaction time to the shape task is
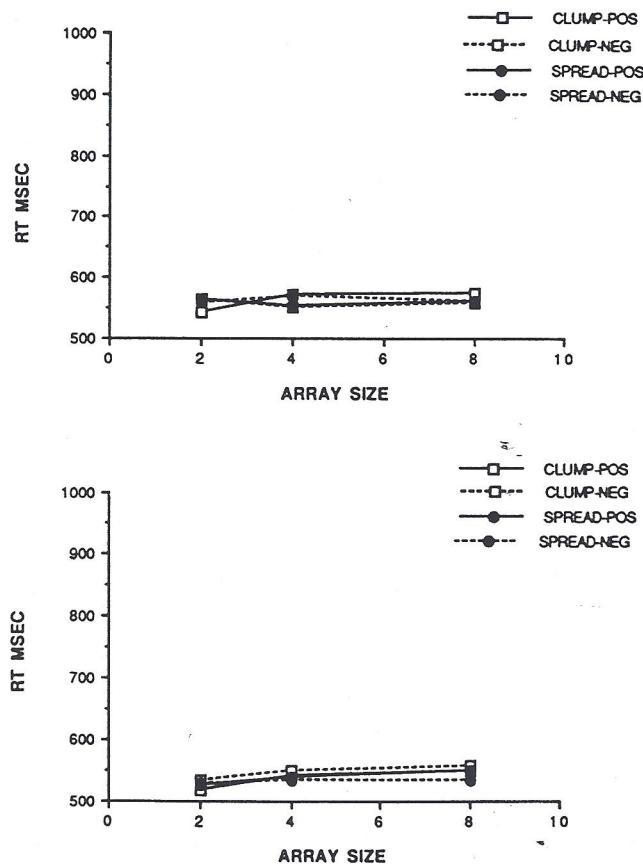

Figure 3. Mean reaction time for the visual search task of Experiment 2. (The upper panel depicts the results of the color task, and the lower panel depicts the results of the shape task.)

presented in the bottom panel. Only correct responses are included. The results of this experiment are very different from those of Experiment 1. The important finding is that unlike the conjunction task, there was no difference between the clump and the spread conditions in either feature task. Similar to other studies with a feature visual search (e.g., Treisman & Gelade, 1980), the slope of the reaction time function with regard to array sizes was essentially flat, with only a slight increase of reaction time with array size. An analysis of variance performed on these data confirmed these observations. The only significant result in the shape task was the main variable of array size, $F(2, 16) = 9.36$, $p < .05$. Subjects were faster in both the spread and the clump conditions when the array consisted of 2 items. The Array Size × Target (present vs. absent) interaction approached significance, $F(1, 8) = 5.01$, $p < .053$. The reason for the difference between small array size and the larger array sizes is not clear, although it is often observed with single-feature tasks (Ivry & Cohen, 1990). The only significant result in the color task was a triple interaction of all three variables, $F(2, 16) = 9.7$, $p < .05$. The reason for this interaction is that the response to the target-present trials with Array Size 2 in the clump condition was faster, but there was no such difference in the target-absent trials. The error rates did not differ between the different conditions.

The main point is clear: The interaction observed between the spatial position (clump vs. spread) and array sizes was not present in the single-feature tasks. Therefore, the lateral-masking explanation cannot account for the results obtained in Experiment 1.

The lack of lateral masking between simple features has been observed previously. For example, Bjork and Murray (1977), through the use of a different paradigm, reported that lateral masking between features did not occur when features were not similar to each other. The two colors and two shapes we used are also dissimilar to each other. Our experiment showed that lack of lateral masking between dissimilar features is also observed in visual search conditions like the one used in Experiment 1.

Experiment 3

Experiment 1 indicated that there is a basic difference between the searches in the clump and spread conditions. Note that this difference holds regardless of the details of the search functions obtained in these two conditions. Nevertheless, although the slope of the array-size function and the ratio of target-absent to target-present trials in the clump condition of Experiment 1 were similar to those obtained in studies with dense arrays (e.g., Egeth et al., 1984; Treisman & Gelade, 1980), the functions in the spread condition only approximated those obtained by Pashler (1987) for spread arrays. The 1.5:1 ratio of absent to present trials in the spread condition of Experiment 1 was higher than the 1:1 ratio Pashler obtained. One possible reason for the difference between the two studies is that we mixed the spread and clump conditions within a block. Subjects may have adopted a conservative search because fast responses would produce errors in the clump condition. In Experiment 3 we test the spread condition by itself.

Method

The method for Experiment 3 was identical to the one used in Experiment 1, with one important difference: Only the spread condition was included. Thus, according to our hypothesis, subjects ought to be able to use the coarse location mechanism consistently in this experiment. Ten subjects took part in this experiment.

Results and Discussion

The reaction time to the correct responses in the various conditions of the visual search are presented in Figure 4. The ratio of target-absent to target-present trials in this experiment,
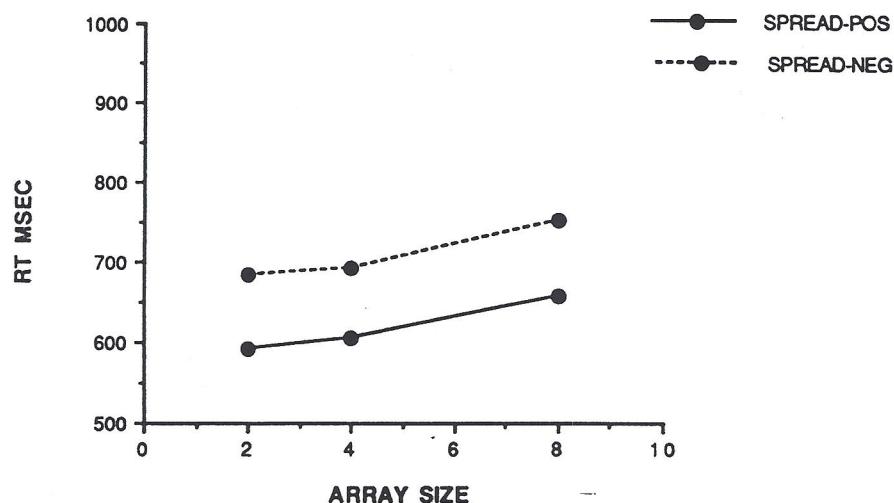

Figure 4.   Mean reaction time for the visual search task of Experiment 3.

1.04, was similar to that obtained by Pashler (1987; 11.3 ms per item in the target-absent condition and 10.9 ms per item in the target-present condition). The range of the ratio of target-absent to target-present trials was fairly small and varied between 0.51 to 1.49.

The combined results of Experiments 1–3 confirm our main hypothesis: People can use the coarse location mechanism for feature integration when the objects in the visual field are sufficiently spread apart. The search rate under these conditions differed from the search rate under clump conditions in two of its characteristics: The search rate was relatively quick and exhaustive (as indicated by the 1:1 ratio between the target-absent and target-present trials). Pashler (1987) provided evidence which suggested that search under these conditions is also done in parallel. When objects are crowded together, the coarse location mechanism cannot be used. Instead, a slower and limited-capacity focal attention mechanism is used. Search rate with this mechanism is slow and self-terminating, as indicated by the slow search per item and the 2:1 slope ratio.

Note that although the search in the spread condition is different from the search in the clump condition, it is also different from search of simple features. A comparison of Figures 3 and 4 reveals that the slope of the search in the simple feature task is flatter than the one in the spread condition of the conjunctively defined target. This difference is expected because detection of the conjunctively defined target with the coarse location mechanism involves detection of overlap between location information of the different features. This computation is not required in the detection of simple features. The additional step required for the operation of the coarse location mechanism will add noise, which will in turn increase the slope of the search (e.g., Cave & Wolfe, 1990; Eriksen & Spencer, 1969).

Experiment 4

The first three experiments were confined to relatively small array sizes. Our hypothesis is that the coarse location mechanism can be used even when the number of the objects in the field is large as long as objects are spread apart. Indirect support for this conjecture can be found by comparing the studies of Pashler (1987) and Wolfe et al. (1989). Pashler, with displays similar to the clump condition, obtained search functions that were slow and self-terminating for a conjunctively defined target with large array sizes. In contrast, Wolfe et al., with displays similar to the spread condition, obtained a shallow search function in a conjunction search on target-present trials with large array sizes. Target-absent trials were inconsistent between subjects and fairly steep. This situation often happens with large array sizes even when the search is for a simple feature.

Experiment 4 tests the generality of the coarse location mechanism by manipulating density with large array sizes. Treisman (1982, Experiment 3) also examined the effect of density on visual search with large array sizes. Contrary to our prediction, Treisman (1982) did not find any effect of density on the search rate. The search rate was slow and self-terminating in both the dense and sparse conditions used in her experiment. In her experiment, Treisman (1982) used desaturated colors. Wolfe et al. (1989) showed that under these conditions the search rate is relatively slow and self-terminating even with the spread condition. In addition, although there was a clear difference in density between the sparse and dense conditions in Treisman's (1982) experiment, the distance between neighboring items in the experiment was not reported. Thus, it is not clear whether the dense and sparse conditions in Treisman's experiment correspond to the clump and spread conditions of our experiments.

Method

Subjects. Twenty-four undergraduate psychology students from Indiana University participated in partial fulfillment of a course requirement. Twelve subjects were assigned to the clump condition, and 12 were assigned to the spread condition.

Apparatus and stimuli. The stimuli were presented on an NEC MultiSync SD color monitor that was controlled by a CompuAdd microcomputer. Subjects were required to put their chin on a chin rest and viewed the display from a distance of 100 cm. As in Experiment 1, the target was a yellow X, and the distractors were yellow Os and blue Xs. The size of each character subtended approximately 0.46° in height and 0.29° in width.

Design. Subjects performed a visual search task similar to Experiment 1. Display contained 4, 8, 16, or 24 items. The items were positioned on two imaginary circles. The radii of the inner and outer circles extended approximately 2.8° and 4.17° from the center. As in Experiment 1, the items in each position were selected with the constraint that runs of identical colors or letters could not exceed three items. For displays containing 4 and 8 items, all of the items were located on either the inner or outer circle (randomly selected). For displays of 16 and 24 items, half of the items were located on each circle. Thus, for the clump condition, the items were arranged in one clump for the two smallest display sizes and two clumps for the two largest display sizes. The distance between the center of one item and the center of an adjacent item was approximately 0.69° within a clump. In the spread condition, the position of the letters was selected in such a way that the distance between two adjacent letters was at least 1.37°. When present, the target replaced one of the distractor items, chosen at random.

Subjects received a short practice block of 24 trials. This was followed by five experimental blocks of 80 trials each, or 10 trials per condition, for a total of 400 trials.

Procedure. There were two differences between the present procedure and that of Experiment 1. First, the luminance task was eliminated. Second, the message "ERROR" was presented on the screen for 500 ms following incorrect responses.

Results and Discussion

Mean reaction time for the correct responses in the various conditions are shown in Figure 5. There was a clear difference between the clump and spread conditions. The data were entered into a 2 (spatial position: clump vs. spread) × 2 (target: present vs. absent) × 4 (array size: 4, 8, 16, or 24) analysis of variance. All of the main effects were significant. Subjects were slower in the clump condition than in the spread condition, $F(1, 22) = 16.24$, $p < .05$; subjects were slower in the target-absent condition than in the target-present condition, $F(1, 22) = 103.24$, $p < .05$; and subjects were slower when the array size increased, $F(3, 66) = 114.14$, $p < .05$. The
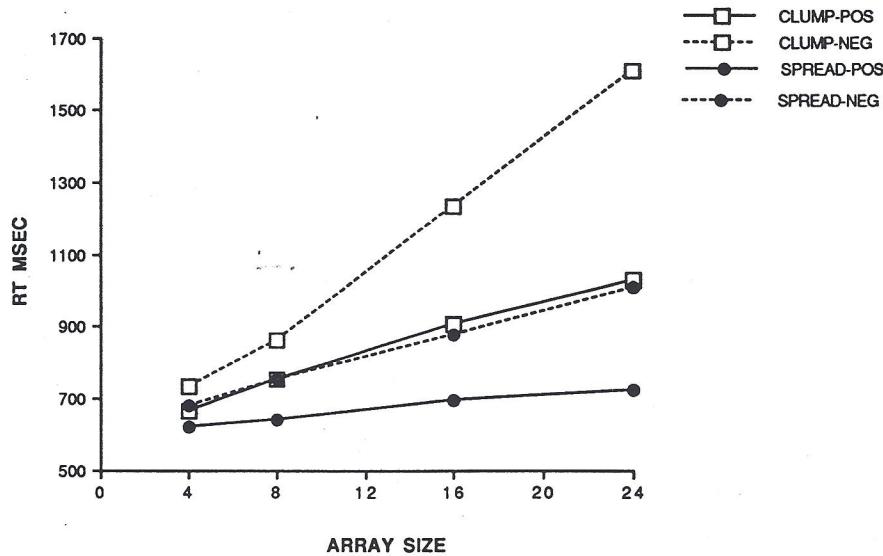
Figure 5. Mean reaction time for the visual search task of Experiment 4.

interactions among the main effects were also significant, as was the three-way interaction: Spatial Position × Target, $F(1, 22) = 6.89$, $p < .05$; Spatial Position × Array Size, $F(3, 66) = 26.92$, $p < .05$; Spatial Position × Target, $F(3, 66) = 60.77$, $p < .05$; Spatial Position × Target × Array Size, $F(3, 66) = 11.33$, $p < .05$. The error rates did not differ between the clump and the spread conditions.

The results in the clump condition are similar to those obtained in Experiment 1: The average search was 18.2 ms per item in the target-present condition and 43.6 ms per item in the target-absent trials. The ratio of absent to present trials was 2.4. These results are comparable to the ones obtained by Pashler (1987) under similar conditions. The slopes for the target-present and target-absent trials in the spread condition were 5.2 ms and 16.3, respectively. These results are comparable to those reported by Wolfe et al. (1989).

The ratio of the absent to present trials in the spread condition was different from the 1:1 ratio obtained in Experiment 3. Search in the target-present trials was minimally affected by array size. Search in the target-absent trials, however, was more appreciably affected by array size. Similar results were obtained by Wolfe et al. (1989) and are sometimes obtained with comparable array sizes even when the search is for a simple feature (e.g., Treisman & Gelade, 1980). Comparison of the slope ratio of individual subjects reveals a clearly different distribution between the clump and the spread conditions. The slope ratio of all but 1 of the subjects in the clump condition ranged from 1.79 to 2.81 (the remaining subject's ratio was 3.21). In contrast, the slope ratio in the spread condition ranged from 0.96 to 6.48. Furthermore, only 1 of the subjects in this condition had a slope ratio within the range of the clump condition (2.46). Two of the subjects had a ratio close to 1 (0.96 and 1.27), and the ratio of the remaining 9 subjects varied from 2.94 to 6.48. One possible explanation for these results is that when large array sizes are used, some subjects in the spread condition adopt a conservative criterion in the target-absent trials to avoid misses.

A second difference between the results of Experiment 3 and 4 is that the search rate on the target-present trials was faster with large array sizes of Experiment 4 (5.2 ms) than a comparable search with the small array sizes of Experiment 3 (10.9 ms). This decrease in search rate with conditions that include large display sizes is similar to the findings of Wolfe et al. (1989). It is possible that eye movements contribute to this effect.[3] If one assumes that the number of eye movements in the wrong direction (i.e., away from the target) is similar with different array sizes and that wrong eye movements slow reaction time by a fixed amount of time, then wrong eye movements will cost more time per item with small array sizes. This possibility has not been tested yet, however. Future studies are needed to settle this question.

General Discussion

The experiments reported in this study demonstrate a large density effect in visual search for a target defined by the conjunction of color and shape. When the items in the display were spread apart, the search rate was relatively fast. Moreover, there were no significant differences between the slopes of the target-present and target-absent search functions over the small array sizes. In contrast, when items in the display were clumped together, the search rates were slower, and the slope ratio of target-absent to target-present trials was roughly 2:1, suggesting a self-terminating search. The density effects were not obtained when the search was for a target defined by a single feature.

_____

[3] We do not suggest that eye movements change the pattern of results in visual search (e.g., whether the search is exhaustive or self-terminating). There is in fact recent evidence (Klein & Farrel, 1990) which suggests that eye movements do not affect the pattern of results obtained in visual search. Our suggestion is that it may cause differences between search with large array sizes and search with small array sizes.

These results provide a resolution to some discrepancies in the visual search literature for conjunctively defined targets (Pashler, 1987; Wolfe et al., 1989). Pashler found that for arrays with relatively few items, the search rate is fast and exhaustive. For large-size arrays, the search rate is slow and self-terminating. Wolfe et al. obtained search functions that became flatter for large arrays, a result almost opposite to Pashler's. Our findings indicate that these discrepant results are due to differences in display density between the experiments. The density was low for both Wolfe et al. and for displays with few objects in the work of Pashler, resembling the spread condition of our study. The search rate under these conditions is relatively fast. The density used by Pashler with large array sizes resembles the clump condition of our study and resulted in a slow search rate.

Our density findings are in accord with the hypothesis that two mechanisms can be used for feature integration. A coarse location mechanism can use coarse location information to bind features into objects. This mechanism is fast but can only operate when objects are not crowded together in the visual scene. If objects are located near each other, a second, slower mechanism is used. We developed (Cohen & Ivry, 1989) an early version of this hypothesis to account for results from illusory-conjunction studies. The current experiments provide evidence with the visual search paradigm.

The assumption that coarse location information becomes available with feature perception and that this information can be used by a coarse location mechanism can also account for other findings in visual search. Egeth et al. (1984) found that subjects can selectively search some items when the target is defined by a conjunction of color and shape. For example, if the target is a yellow X and the distractors are yellow Os and blue Xs, subjects may search among the yellow items only. The assumption that location information is available with the registration of the features may explain this phenomenon. Subjects may use the coarse location information that is registered with the yellow feature to guide their search.

Two findings reported by Wolfe et al. (1989) also fit with our findings. First, Wolfe et al. found that when a target is composed of a conjunction of three features (e.g., large yellow X), and the distractors differ from the target in two of these features (e.g., large blue Os, small blue Xs, and small yellow Os), the detection of the target is only minimally affected by the number of distractors (see also Quinlan & Humphreys, 1987). The explanation for this finding is identical to our account of the current results. The coarse location mechanism can be used to detect the target in the triple conjunction situation because the stimuli were spread apart. The coarse location mechanism can be used to detect overlap of location information between any two of the three features that define the target. In the example given previously, overlapping location information of either yellow and X, large and yellow, or large and X is a positive signal for the presence of the target.

Wolfe et al. (1989) also found that a particular spatial arrangement of line orientations (e.g., finding L among Ts when the length of the lines in the two letters is equated) requires a slow and self-terminating search rate. According to our framework, these findings imply that the spatial arrangement of line orientations that discriminate Ls and Ts requires greater resolution than that contained in the coarse location

information available for the coarse location mechanism. Thus, the slow search mechanism is required in this situation.

Relation to Other Models of Feature Integration

Our hypothesis is that two separate mechanisms for feature integration exist. Two other theories, the original feature integration theory proposed by Treisman (e.g., Treisman & Gelade, 1980; Treisman & Schmidt, 1982) and a recent theory proposed by Wolfe and Cave called the *guided search model* (Cave & Wolfe, 1990; Wolfe et al., 1989; see also Treisman & Sato, 1990, for a similar model), posit only a single mechanism for feature binding. Can these theories explain the data from the visual search and from the illusory-conjunction paradigms?

Feature Integration Theory

The original theory proposed by Treisman and Gelade (1980; see also Kahneman & Treisman, 1984) assumed that features are registered in spatiotopic feature maps. Feature detection can be performed by checking activity in the relevant feature maps. Because feature registration occurs in parallel, the reaction time to determine the presence of features is minimally affected by the number of distractors. Feature integration requires that the locations on the feature maps be connected to a master spatiotopic map. When attention is directed to an activated location on the master map, the corresponding location in each feature map is activated, and the features at this location are conjoined. The information available in the master map is not sufficient to determine the presence of a particular conjunction of features and to attract the binding mechanism: The representations are in terms of locations that contain objects without specifying the features at the activated locations. Thus the theory postulated that feature integration requires a focal attentional mechanism that operates by serially scanning the representations on the master map. The theory did not provide any other mechanism for feature integration.

How can this theory account for the density effects reported in this article? One possible way is to assume that the focal attention mechanism can scan the visual field much faster when objects are spread apart than when objects are clumped. In other words, the presence of nearby objects interfere with the scanning of a given object, which results in a much slower search rate when objects are clumped. There is a problem, however, with this explanation. The search rates obtained in our experiments and by Wolfe et al. (1989) were very fast (5 ms per item). Indeed, the recent reports of fast search rate with conjunctively defined targets (Nakayama & Silverman, 1986; Wolfe et al., 1989) have led Treisman to propose that subjects may strategically modulate the output from locations on the master map of selected features when the features are highly discriminable (Treisman, 1988; Treisman & Sato, 1990). This modification of feature integration theory yields a model similar to the guided search model (discussed next).

Guided Search Model

Recently, Cave and Wolfe (1990; see also Wolfe et al., 1989) proposed an alternative theory, the guided search

model. Similar to the feature integration theory, the model assumes that there is a bottom-up analysis of the features in feature maps. The guided search model, however, postulates that the activation of the features present at a particular location is summed in the master map (which they called the *activation map*). The bottom-up processing of the features is not sufficient to distinguish between conjunctively defined targets; however, knowledge of the identity of the target can guide the feature integration mechanism to the location of candidate conjunctions. For example, when searching for a yellow X target among yellow O and blue X distractors, there is more activation in the feature map of color when yellow is detected (because of the top-down knowledge that yellow is searched for) than when blue is detected. Similarly, there is more activation in the shape map when X is detected than when O is detected. Consequently, when the yellow X target is present, the summed activation sent from the color and shape maps at the particular location of the target to the master map will be higher than summed activation at any other locations. The visual system can quickly find the location with higher activation, and thus the target can be detected quickly (see Cave & Wolfe, 1990, for additional details).

In its current form, the guided search model cannot account for the density effects in visual search. With the additional assumption that the locations in the activation map are coarsely coded, however, the model can explain the results. In fact, with this modification the guided search model may be viewed as one possible implementation of the coarse location mechanism. Note that the guided search model assumes that there is only a single binding mechanism. When the pattern of activation of a conjunctively defined target is distinct (i.e., when items are spread apart), the binding mechanism is attracted to the location of that target. When the pattern of activation of the target is not sufficiently distinct, the binding mechanism needs to scan the activated locations.

Although the density effects in visual search do not distinguish between the one- and two-mechanism accounts, findings in illusory-conjunction experiments give some support to the two-mechanism account. In Experiment 2 of Cohen and Ivry (1989), subjects were asked to perform two conjunctive tasks on each trial. One task was similar to that used in the current experiments: Two colored letters were presented, and the subject was required to report the color of a cued letter. For a second task, two achromatic digits were presented, one small and one large, and the subject reported the smaller digit. The digits were presented at the center of the display, and the instructions emphasized accuracy; thus this task required subjects to focus their attention at the display center. The colored letters were presented in the periphery. Subjects committed illusory conjunctions between the colored letters only when the two colored letters were spatially adjacent. This pattern of results was similar to that obtained when the digit task was simpler and required identification of a single digit (Cohen & Ivry, 1989, Experiment 1). These results can be explained by the two-mechanism account by assuming that the attentional binding was operating in the middle of the screen and the coarse location mechanism was operating in the periphery. Thus, variations in the digit task that affect the attentional binding mechanism do not affect the performance in the letter task. The one-mechanism account would have to

assume that subjects were first able to identify the digit and then shift their binding mechanism to the periphery. Most critical, the one-mechanism model would have to propose that the added difficulty of the digit task did not affect the speed of shifting attention. Though this explanation seems tenuous, it cannot be rejected, because the quantitative parameters of shifting attention are not known.

In summary, we believe that the current evidence supports the two-mechanism account. One-mechanism explanations can also account for the data, however, and thus additional studies are needed to resolve the issue. We outline two types of studies that may bear on this debate. First, if two distinct mechanisms are used, they may have different characteristics. For example, feature binding may differ in terms of whether it occurs within or outside the focus of attention. In Experiment 1 we pursued this question by investigating inhibitory effects following conjunction search, predicting that inhibition would only be obtained for those displays that required the focal binding mechanism. Unfortunately, we were unable to obtain inhibitory effects, failing to replicate Klein (1988). The point we emphasize here is that experiments involving similar logic may be useful for comparing one-mechanism and two-mechanism hypotheses. Second, studies with neurologically impaired patients may also be indicative. In particular, some parietal lobe patients suffer from a deficit in shifting their attention (e.g., Posner, Walker, Friedrich, & Rafal, 1984). It is possible that such patients are impaired with their attentional binding mechanism but can use the coarse location mechanism. We are currently pursuing this line of investigation (see Cohen & Rafal, 1991).

Open Issues

Several other issues concerning the binding mechanism are still open. We mention two of the questions that we are currently exploring.

Effect of irrelevant features. The distractors in our experiments always shared one feature with the target (the blue X distractor shared the shape, and the yellow O distractor shared color). Under these conditions, search in the clump condition was slow. We are now investigating displays with four distractors: two as in the current experiments and two that do not share any features of the target. If the items are clumped, will search be slow when neighboring distractors do not share any of the features with the target? For example, if the target is a yellow X and neighboring distractors are green Ss and red Ts, will the search be similar to the clump condition or the spread condition? This question is important because in a typical visual scene, objects are often clumped together. In many of these situations, however, we may search for an object that has different features from adjacent objects. Can the coarse location mechanism be used in these situations?

Number of possible targets. In the current experiments, subjects searched for a single target. Treisman and Gelade (1980) found that search for two different targets defined by simple features is not affected by the number of items in the display. We have shown that search for conjunctively defined targets is minimally affected by the number of distractors for low-density displays. What happens in the spread-condition search when the subjects must determine the presence of

either one of two conjunctively defined targets? The answer depends on specific processing assumptions. For example, Cave and Wolfe (1990) suggested that activity of features is linearly summed in the master map; this assumption leads to clear predictions. It is possible to arrange situations under which two targets can be detected by this linear mechanism in parallel and other situations under which two targets cannot be searched in parallel but rather have to be searched sequentially. For example, such a mechanism predicts that when the two possible targets are yellow X and red T, and the possible distractors are blue X, yellow O, red T, and green S, it is possible to search for two targets in parallel. When the possible targets are yellow X and blue O, and the distractors are yellow O and blue X, it is impossible to search both targets in parallel because both the color and shape of each distractor are features of the targets. Therefore, feature summation will not be sufficient to differentiate the target from the distractors. Wolfe and Cave (1990) reported that in the latter situation, subjects perform two sequential searches, as is predicted by a linear mechanism. It remains to be determined whether search for the two targets is performed in parallel in the former situation, in which the hypothetical linear mechanism can operate on both targets simultaneously.

References

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

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

Cohen, A., & Ivry, R. (1989). Illusory conjunction inside and outside the focus of attention. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 650–663.

Cohen, A., & Rafal, R. D. (1991). Attention and feature integration: Illusory conjunctions in a patient with a parietal lobe lesion. *Psychological Science, 2*, 106–109.

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

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

Eriksen, C. W., & Spencer, T. (1969). Rate of information processing in visual perception: Some results and methodological considerations. *Journal of Experimental Psychology Monographs, 79*(2, Pt. 2).

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. In D. E. Rumelhart & J. L. McClelland (Eds.), *Parallel distributed processing: Vol. 1. Foundations* (pp. 77–109). Cambridge, MA: MIT Press.

Houck, M. R., & Hoffman, J. E. (1986). Conjunction of color and form without attention: Evidence from an orientation-contingent color aftereffect. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 186–199.

Ivry, R. B., & Cohen, A. (1990). Dissociation of short- and long-range apparent motion in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 16*, 317–331.

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

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

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*, 430–431.

Klein, R., & Farrel, M. (1990). Search performance without eye movements. *Perception & Psychophysics, 46*, 476–482.

Livingstone, M. S., & Hubel, D. H. (1987). Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *Journal of Neuroscience, 7*, 3416–3468.

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

Nakayama, K., & Silverman, G. H. (1986). Conjunctive visual search for stereo-motion and stereo-color is parallel. *Nature, 320*, 264–265.

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

Posner, M. I., Walker, J. A., Friedrich, F. J., & Rafal, R. D. (1984). Effects of parietal lobe injury on covert orienting of visual attention. *Journal of Neuroscience, 4*, 1863–1874.

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.

Snyder, C. R. R. (1972). Selection, inspection and naming in visual search. *Journal of Experimental Psychology, 92*, 428–431.

Townsend, J. T. (1976). Serial and within-stage independent parallel model equivalence on the minimum completion time. *Journal of Mathematical Psychology, 14*, 219–239.

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

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. M. (1986). Properties, parts and objects. In K. Boff, L. Kaufman, & J. Thomas (Eds.), *Handbook of perception and human performance* (pp. 1–70). New York: Wiley.

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

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

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

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

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

Wolfe, J. M., & Cave, K. R. (1990). Deploying visual attention: The guided search model. In T. Troscianko & A. Blake (Eds.), *AI and the eye* (pp. 79–103). Chichester, England: Wiley.

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

Wolfe, J. M., & Pokorny, C. W. (1990). Inhibitory tagging in visual search: A failure to replicate. *Perception & Psychophysics, 48*, 357–362.