Journal of Experimental Psychology: General
1998, Vol. 127, No. 2, 189–219

0096-3445/98/$3.00

What's in a Location? Comparing Object-Based and Space-Based Models of Feature Integration in Visual Search

Morris Goldsmith
University of Haifa

What is the unit of selection for feature integration in visual search: location or perceptual object? Feature integration theory (A. Treisman, 1988) asserts that it is location. Two alternative models are put forward and tested in a series of 4 experiments using a special conjunctive-search task. In this task, each stimulus item consists of 2 overlapping forms (perceptual objects). In general, the search was more efficient when the search features were linked to the same perceptual object than when they were linked to different perceptual objects at the same stimulus location. This same-object advantage, however, was shown to depend on stimulus discriminability and density, grouping strength, and hierarchical object structure. The results support a hierarchical object-based model, with important implications for feature integration, visual search, late versus early selection, and object-based versus space-based views of attention.

For many busy researchers, it is an all too common experience to open a cluttered desk drawer and hurriedly search for a particular writing instrument or other essential item. It may take a few moments, but the elusive red pen, for instance, can still be detected even though it may be partly covered by a yellow-handled pair of scissors and some colored plastic paper clips. Unless one happens to be doing work on visual search, one's ability to find the desired item under such conditions would probably not give rise to much thought.

Within the area of visual search, the dominant theoretical framework for many years has been Treisman's (1986a, 1988; Treisman & Gelade, 1980; Treisman, Sykes, & Gelade, 1977) feature integration theory (FIT). FIT is specifically concerned with the function of attention in the perception of objects. As in other analytic theories of object perception (for comprehensive reviews, see Livingstone & Hubel, 1987; Treisman, 1986b), in FIT it is assumed that visual information is initially analyzed in terms of primitive features registered automatically and in parallel across the visual field. This early analysis has clear computational advantages (Marr, 1982; Tsotsos, 1990), but it also creates a problem for the visual system: If visual features are initially analyzed and represented irrespective of the objects from which they originate, how does the visual system ultimately know which features go together? How are the features belonging to the same object rather than to different objects correctly recombined into an integrated object representation? According to FIT, this problem is solved by a focused-attentional feature binding mechanism.

This article examines a fundamental assumption made by FIT concerning the operation of this mechanism, namely, that the basic unit of attentional selection is *spatial location*.

Object-Based Versus Space-Based Attentional Selection

The question of whether the basic parameters of attention are object-based or space-based has become a widely debated topic in research on visual cognition generally (for reviews, see Kanwisher & Driver, 1992; Kramer & Jacobson, 1991; Egly, Driver, & Rafal, 1994). The issues have both representational and processing aspects. For instance, in terms of representation, "object-based theories of attention suggest that the visual world is parsed into objects or perceptual groups . . . and that attention is directed to these objects. Such theories are in contrast to purely spatial views suggesting that attention is directed to unparsed regions of space" (Baylis, 1994, p. 208). Also, in terms of processing,

> space-based models suggest that spotlights, zoom lenses, and gradients provide apt analogies for the allocation of attention. For example, in models based on the notion of a spotlight, attention is distributed in contiguous regions of the visual field. Stimuli that fall within this region or spotlight are extensively processed, while events that occur outside this area are ignored. (Kramer & Jacobson, 1991, p. 267)

In contrast, object-based accounts often emphasize situations in which "the spotlight metaphor breaks down" (Driver & Baylis, 1989).

This research is based on parts of a doctoral dissertation submitted to the Department of Psychology, University of Haifa, Haifa, Israel. It was conducted and written with the support of a National Science Foundation Graduate Fellowship, a Sam Cohen Fellowship from the University of Haifa, Haifa, Israel, and a Dissertation Writing Grant from the Israel Foundations Trustees. The work was carried out at the Institute of Information Processing and Decision Making, University of Haifa, Haifa, Israel.

I am grateful to David Navon for his valuable advice and comments throughout the project and to Shaul Hochstein, Asher Koriat, Art Kramer, Yehoshua Tsal, Steven Yantis, and Shaun Vecera for helpful comments on earlier versions of this article. Thanks also to Dahlia Mevorach for programming assistance.

Correspondence concerning this article should be addressed to Morris Goldsmith, Department of Psychology, University of Haifa, Haifa, 31905, Israel. Electronic mail may be sent to mgold@psy.haifa.ac.il.

The debate between the object-based and space-based views encompasses a wide range of theoretical frameworks, empirical phenomena and experimental tasks. A representative but far from exhaustive list of topics and studies includes work on selective attention (Baylis & Driver, 1992; Driver & Baylis, 1989; Kramer & Jacobson, 1991; Treisman, Kahneman, & Burkell, 1983), divided attention (Baylis & Driver, 1993; Duncan, 1984; Kramer & Watson, 1996; Vecera & Farah, 1994; Treisman et al., 1983), covert visual orienting (Egly et al., 1994; Vecera, 1994), negative priming (DeSchepper & Treisman, 1996; Tipper, Brehaut, & Driver, 1990), repetition priming (Kahneman, Treisman, & Gibbs, 1992), inhibition of return (Tipper, Driver, & Weaver, 1991; Tipper, Weaver, Jerreat, & Burak, 1994), feature integration (Prinzmetal, 1981), and visual search (Donelly, Humphreys, & Riddoch, 1991; Duncan & Humphreys, 1989, 1992; Humphreys, Quinlan, & Riddoch, 1989). Notably, the debate has also extended into the neuropsychological literature, in which a number of clinical disorders such as unilateral visual neglect (Driver & Halligan, 1991; Farah, Wallace, & Vecera, 1993; Humphreys, Olson, Romani, & Riddoch, 1996; Humphreys & Riddoch, 1993b; Tipper & Behrmann, 1996), simultaneous agnosia (Humphreys & Riddoch, 1993a; Luria, 1959), and unilateral spatial extinction (Humphreys et al., 1996; Ward, Goodrich & Driver, 1994) are now thought to involve both object-based and space-based attentional deficits (for reviews of this literature, see Egly et al., 1994; Rafal, 1996). The issues, therefore, are not only central to our understanding of visual perception, but may have important practical implications as well.

Unfortunately, however, the lines of demarcation between the object-based and space-based views are not always clear. Particularly problematic is the term *object-based* itself, which not only lacks a precise definition (Duncan, 1984; Logan, 1996) but is often used in different ways by different researchers. To illustrate the problem and help clarify the usage in this article, consider a seminal object-based demonstration by Duncan (1984). He presented observers with stimulus displays containing two perceptual objects (an elongated box and a superimposed straight line; cf. Figure 2, which appears later in this article) and asked them to make judgments about two properties of the display that pertained either to the same object (e.g., the texture and tilt of the line) or to different objects (e.g., the texture of the line and the size of the box). Although the two objects were centered at the same spatial location, the judgments were more accurate when they pertained to the same object than when they pertained to different objects, presumably reflecting the advantage of attending to one rather than two separate object representations.

Leaving aside for the moment certain methodological problems with this demonstration (which will be addressed later), the most basic issue concerns the nature of the object-based representation that is being accessed. Given, for example, that both the box and the line in Duncan's (1984) experiment correspond to two different sets of points in space, in what sense was the unit of selection object-based rather than space-based? Indeed, how are the object-based and space-based views to be distinguished?

The approach taken most often by object-based theorists is to acknowledge the obvious fact that objects occupy different sets of points in space, but to assert that "the chunk of information dealt with by focal attention is determined by Gestalt grouping, not by anything specifically spatial" (Duncan, 1984, p. 515). Thus, although a space-based theorist might postulate a flexible attentional "spotlight" that can conform to the precise shape of, say, a line or a box, by common usage this would in fact imply an object-based account. In addition, the object-based view also holds that "grouping may be viewed in terms of a quantitative metric of strength rather than a qualitative distinction between objects" (Kramer & Jacobson, 1991, p. 273). Thus, although spatial factors (e.g., proximity) may affect the ease of attentional processing, the object-based view holds that they do so indirectly, by means of the contribution of spatial factors to gestalt grouping strengths (Wertheimer, 1923). Note, then, that on this view, object-based and space-based effects are not mutually exclusive. Rather, space-based attention might be assimilated into a more general, object-based framework (see General Discussion).

This type of object-based view, however, differs substantially from a much stricter usage, in which *object-based* is equated with the use of spatially invariant, *object-centered* (Marr, 1982) representations, thus denying the possibility that spatial variables might moderate an object-based effect (Vecera, 1994; Vecera & Farah, 1994; see also Humphreys et al., 1996). Thus, for instance, Vecera and Farah (1994) repeated Duncan's (1984) experiment, but included a condition in which the two objects (line and box) were spatially separated. They reasoned that if the cost of dividing attention between the properties of the line and the properties of the box stemmed from the need to access two spatially invariant object representations (as opposed to a single representation in the same-object condition), then this cost should be no greater when the objects are spatially separated than when they are superimposed. Surprisingly, this in fact is what Vecera and Farah found (1994, Experiments 1 and 2), supporting the existence of a truly object-based representation in the Duncan (1984) task (but for contrary evidence, see Kramer, Weber, & Watson, 1997). In other tasks, however, the coexistence of gestalt-based and proximity-based effects was taken to imply an underlying grouped location representation (e.g., Vecera, 1994; Vecera & Farah, 1994, Experiments 3 and 4).

In keeping with common usage (see, e.g., Egly et al., 1994; Logan, 1996), this article puts forward and examines an object-based account of feature integration in which gestalt grouping factors both define the unit of selection and influence the efficiency of attentional processing. To anticipate, the issue is not whether feature integration is based on representations that are spatially invariant (clearly it is not). Rather, the question is how much object-based (gestalt) structure is built into the representations on which attention operates, and what are the consequences of this structure for the feature-integration process.

Object-Based Versus Space-Based Feature Integration

As mentioned earlier, FIT is a widely accepted theory about the role of attention in the perception of objects. It is also a *space-based* theory. In FIT, attention is assumed to operate like a spatiotemporal *spotlight* (Treisman & Gelade, 1980; Treisman & Gormican, 1988; cf. Eriksen & Hoffman, 1972; Posner, 1980) or *window* (Treisman & Sato, 1990) that serially selects particular spatial locations in the visual field (see Figure 1):

> The initial assumption is that different sensory features, such as colours, orientations, sizes, or directions of movement are coded in specialized modules ... automatically, without focused attention, and spatially in parallel.... Each module forms different feature maps for the different values on the dimensions it codes.... When features must be located and conjoined to specify objects, attention is required. Attention selects within a "master map of locations" which shows *where* all the feature boundaries are located, but not *which* features are located *where*.... When attention is focused on a particular location in the master map, it allows automatic retrieval of whatever features are currently active in that location through links to the corresponding locations in the different modular feature maps. (Treisman, 1988, p. 203)

In FIT all features present in the same attentional fixation are bound into a conjoined object representation. This temporary object representation or *object file* (Kahneman & Treisman, 1984)—the product of the feature-integration process—provides the basis for object recognition (see Figure 1) and constitutes the unit of selection for subsequent object-based processing (Treisman, 1988). Thus, somewhat paradoxically, FIT is actually cast within a more general, object-based attentional framework (see, e.g., Kahneman & Treisman, 1984; Kahneman et al., 1992; Treisman, 1988, 1992a, 1993). According to FIT, however, it is the feature-integration process that transforms a preattentive space-based representation into an integrated object representation: object-based attentional processing occurs only after attention has played its initial role in feature integration. In this sense, FIT's assumptions reflect an early-selection, late object-based view of visual attention (see General Discussion). In this article, I question that view.

The Problem of Overlapping Stimuli

Notwithstanding an impressive amount of success in explaining the basic pattern of results in several experimental paradigms (see Treisman, 1988; Treisman & Gelade, 1980), in recent years FIT has been faced with an increasing

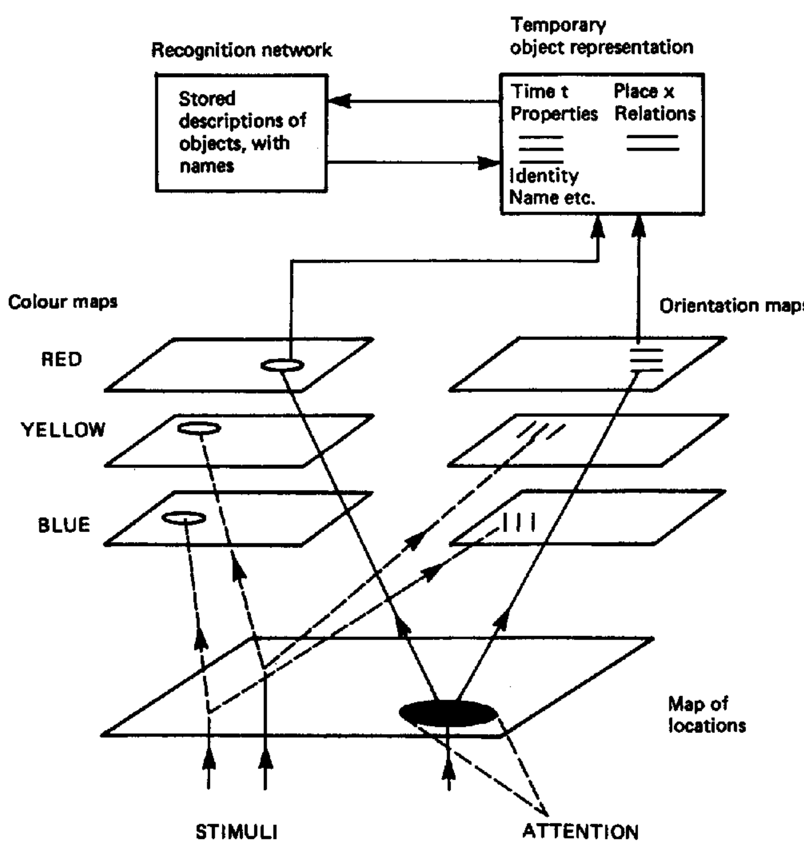


Figure 1. Model for the role of attention in feature integration theory. Reprinted from "Features and Objects: The Fourteenth Bartlett Memorial Lecture," by A. Treisman, 1988, *Quarterly Journal of Experimental Psychology, 40A,* p. 202. Copyright 1988 by the Experimental Psychology Society. Reprinted with permission.

number of theoretical and empirical challenges (e.g., Ashby, Prinzmetal, Ivry, & Maddox, 1996; Duncan & Humphreys, 1989, 1992; Navon, 1990; Navon & Ehrlich, 1995; Tsal, 1989; Tsal, Meiran, & Lavie, 1994). One problem that has gone relatively unnoticed concerns FIT's ability to explain how people process complex visual scenes. Unlike the well-differentiated stimulus displays typically used in laboratory experiments, in which the items are "neatly isolated on a homogeneous background" (Wolfe, 1994a, p. 228), in more natural environments there is often a great deal of spatial overlap between the parts and surfaces of different objects (e.g., an animal in a bush; a pen in a drawer) or between the various parts of a single object (see Wolfe, 1994b, 1996). Thus, as Treisman and Gelade (1980) acknowledge, the claim that

> object identification depends on focal attention, directed serially to different locations, to integrate the features registered within the same spatiotemporal "spotlight" into a unitary percept . . . is of course highly oversimplified; it begs many questions, such as how we deal with spatially overlapping objects. (p. 134)

The basic problem posed by overlapping stimuli is illustrated schematically in Figure 2. The illustration depicts two relatively complex stimulus items, each composed of two overlapping forms (i.e., S and square) occupying the same location. The two stimuli differ from one another in their coloring schemes: Whereas for one stimulus it is the S that is colored red and the square that is black, for the other it is the *square* that is colored. It is not clear from the basic FIT model (see Figure 1), however, how this difference can possibly be perceived: A unitary attentional spotlight or window encompassing both overlapping forms at each stimulus location should conjoin the same set of elementary features (i.e., S-shape, square, red, black)[1] in either case. Note that unlike previous demonstrations using overlapping stimuli (e.g., Duncan, 1984; Rock & Gutman, 1981), in this example the two overlapping forms are of about the same size, and are perceptually overlapped in depth as well as in the frontal plane: The coloring (red or black) of the four points of overlap is designed to induce a perception of the S as being in front of the square's right corners, but behind the square's left corners. Hence, neither three-dimensional location (Duncan, 1984) nor spatial scale (Watt, 1988) can be used to isolate the individual features. How, then, can our obvious ability to distinguish the different coloring schemes be explained within the FIT framework?

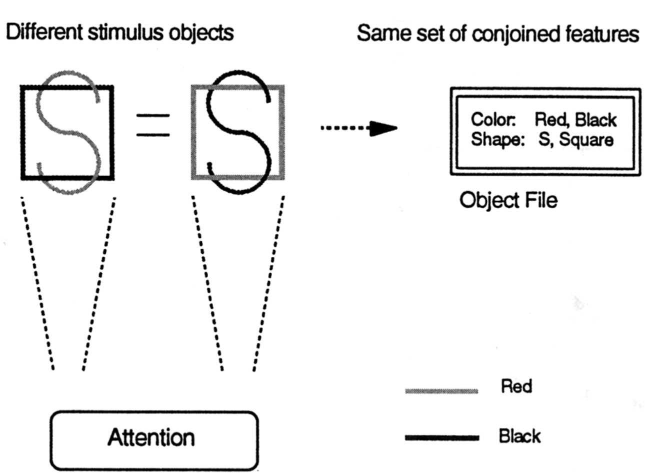


Figure 2. The problem posed by overlapping stimuli for the feature integration theory: Two stimuli are displayed, each composed of two overlapping forms, each with a different color scheme. A unitary attentional spotlight or window directed to each stimulus location would conjoin the same set of color and shape features in either case.

Two Alternative Models

As a point of departure for addressing this question, consider once again how the attentional medium is described in FIT:

> Attention selects within a "master map of locations" which shows *where* all the feature boundaries are located, but not *which* features are located *where*. Thus it distinguishes "filled" from "empty" locations, where "filled" implies the presence of any discontinuity at the feature level. (Treisman, 1988, p. 203)

Clearly, the master map of locations is not oblivious to the existence of objects, and in fact, the filled locations are assumed to be the product of preattentive segmentation and grouping processes (Marr, 1982; Neisser, 1967). However, in this space-based representation, *location* is the unit of selection, and "objectness" is an attribute of specific locations. That is, locations are represented and accessed directly, whereas objects (viz., potential objects) are represented and accessed only by virtue of occupying a given location. Because filled (or not filled) is an attribute of each location, it is possible to select only occupied locations. However, if more than one object occupies a given location (i.e., is centered at the same spatial coordinates), all must be selected.

Now consider a somewhat different possibility, an *object-based* master map in which *perceptual objects* are the unit of selection and location is an attribute of specific objects. Here it is the *gestalt* (Wertheimer, 1923) or *uniform connectedness* (Palmer & Rock, 1994) of the form that is selected, rather than its spatial location per se (see, e.g., Duncan, 1984; Kramer & Watson, 1996). Such an object-based master map might consist of a set of primitive object *tokens* or *markers* (Fox, 1978; Marr, 1982; Pylyshyn, 1989), each linked to its own individual features in the feature maps. Attention can now select *this* or *that* object marker (Pyly-

[1]Geometric shapes and English letters are probably not primitive visual features in the same sense that, for instance, color, size, and orientation are. The implicit assumption (e.g. Treisman & Gelade, 1980) is that the features actually detected and processed are more primitive shape properties (e.g., curvature, intersection, etc.; see Treisman & Gormican, 1988). Treating shape as a feature dimension, however, serves to simplify the exposition, and so this convention has been adopted in this article.

shyn, 1989), whether in a random scan or using the links from specific feature maps, even if two or more markers occupy the same location. Only the features linked to the selected marker are conjoined into an integrated object file.

These two characterizations of the attentional medium imply two different models of feature integration, a space-based and an object-based model, respectively (see Figure 3). In the space-based model (Figure 3A), it is the scope of the spatially circumscribed attentional fixation that determines which features are integrated into a single object description. Such space-based feature binding is implied by the glue metaphor commonly associated with FIT: "Any features which are present in the same central 'fixation' of attention are combined to form a single object. Thus focal attention provides the 'glue' which integrates the initially separable features into unitary objects" (Treisman & Gelade, 1980, p. 98; see also Briand & Klein, 1987). Although Treisman (1990) has since expressed some reservations regarding the applicability of the glue metaphor (see also Navon, 1990; Johnston & Pashler, 1990), this type of space-based feature binding is in fact embodied in FIT's various search models: For instance, the *group-processing model* (Treisman & Gormican, 1988; Treisman & Sato, 1990) suggests that target identification is based on a pooled response to activation from one or more target feature maps, summed across a spatially circumscribed attended region. In effect, this pooled response to activation within the focus of attention constitutes a specific instantiation of a space-based attentional binding mechanism.

In contrast, the object-based model (Figure 3B) assumes both an underlying object-based representation and a mandatory object-based mode of attentional feature binding (see also Prinzmetal, 1981). In this model, strong preattentive links between perceptual objects and their features, together with other aspects of the preattentive object structure (e.g., grouping strengths and hierarchical structure; see Experiments 3 and 4 later), constrain the attentional feature-integration process. Hence, rather than arbitrarily conjoining

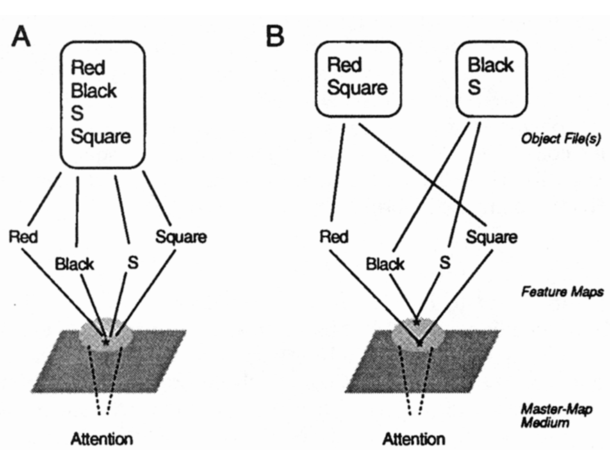


Figure 3. Graphic depiction of the space-based (A) and object-based (B) models of feature integration. The stimulus in each case is a black S and overlapping red square at the same stimulus location.

spatially proximal features into a single object description, attention is directed to particular object markers with their individual sets of preattentively linked features, and these are assigned to separate object files. Note that, as discussed earlier, the object markers in an object-based representation need not be spatially invariant or "free floating" (cf. Treisman & Gelade, 1980). If spatial location is coded as an attribute of the object marker, then perceptual objects can be selected according to spatial parameters (e.g., all the objects or object parts in a certain location or region) as well as through their links to the feature maps. The binding process, however, is still constrained by the need to access the separate object markers. According to this model, then, although attention still has an important role in creating the type of temporary object-file representations required for higher level processing, that role might better be characterized as isolating, individuating, and perhaps elaborating preattentively determined stimulus structure (Navon, 1990), rather than as integrating new structure (see General Discussion).

Overview of the Experiments

Previous research examining space-based versus object-based feature integration has focused primarily on the effects of interitem grouping. Prinzmetal (1981), for instance, using an "illusory conjunction" paradigm (Treisman & Schmidt, 1982), found that conjunction errors (wrongly combining the features of two different objects) were more likely to occur between items in the same perceptual group (defined by collinearity or color similarity) than between items in different groups, even when the spatial distance between the items was controlled (see also Baylis, Driver, & McLeod, 1992; Prinzmetal & Keysar, 1989; Prinzmetal & Mills-Wright, 1984). Similarly, several studies have shown that interitem grouping can either facilitate or impair the efficiency of conjunctive search (e.g., Banks & Prinzmetal, 1976; Bundesen & Pedersen, 1983; Egeth, Virzi, & Garbart, 1984; Humphreys et al., 1989; Treisman, 1982). Collectively, these results make a strong case for the sensitivity of the feature-integration process to potential object groupings. In general, however, such findings are subject to interpretation in terms of the effects of grouping on the spatial bounds of attention (see Logan, 1996; Treisman, 1982, 1992b), in which case either the underlying representation or the attentional mode of processing, or both, could conceivably be space based.

This study attempts to demonstrate effects of objectness on the feature-integration process that cannot be accounted for in terms of the spatial extent of attention. To that end, the logic of Duncan's (1984) study with overlapping stimuli was adapted for use in a visual-search paradigm. Using stimuli similar to those depicted in Figure 2, a special conjunctive-search task was devised, involving two commonly used feature dimensions, shape (e.g., the letters S and V) and color (e.g., red and blue). As in most conjunctive-search tasks, the target is defined as a particular combination of the two features (e.g., red and S), and the distractors are created by combining either the target shape with the nontarget color (blue and S), or the target color with the nontarget shape (red and V). In this task, however, each stimulus item is composed of both a *relevant* form (i.e., one of the letters) and an overlapping *neutral* form that is common to all stimuli (e.g., a rectangle). Furthermore, each stimulus is colored according to either of two *stimulus coloring schemes* (as depicted in Figure 2): (a) *Type U* (unified)—the relevant form (letter) is colored and the neutral form (rectangle) is achromatic; or (b) *Type D* (divided)—the neutral form (rectangle) is colored and the relevant form (letter) is achromatic. Importantly, the target conjunction is defined as the co-occurrence of the target color and the target letter-shape at any single stimulus location, regardless of whether it is the letter (relevant form) or the overlapping rectangle (neutral form) that is colored (i.e., regardless of the stimulus coloring scheme). Thus, for example, both of the stimuli depicted in Figure 2 would qualify as instances of the target red and S.

What are the predictions for this task? The space-based model predicts that in order to test each stimulus item for the target conjunction of features, regardless of that item's particular color scheme, the attentional binding mechanism will simply select the item's location. The processing times for Type U and Type D stimuli should then be equivalent, because the same object file will be created and tested in either case (see Figure 2). By contrast, according to the object-based model, the two processing times should differ: The processing time for a Type U stimulus should be faster than for a Type D stimulus, because in the former case both of the relevant features (color and shape) may be bound and tested by directing attention to a single object marker and ensuing object file (corresponding to the letter), whereas in the latter case attention must be divided between the letter's and the rectangle's object markers and files in order to test for both features (cf. Duncan, 1984; Treisman et al., 1983).

Each of the four experiments uses a variant of this basic paradigm: Participants search through matrices of stimulus items whose color schemes conform to one of three *matrix color-scheme* conditions: (a) *U-only* matrices that include only Type U stimuli, (b) *D-only* matrices that include only Type D stimuli, and (c) *mixed* matrices that include both Type U and Type D stimuli distributed randomly among the stimulus locations. Although the space-based model predicts no effect for matrix color scheme on search latencies, the object-based model predicts that performance should be fastest for U-only matrices, and moreover, this advantage should be reflected in the slopes of the functions relating search latency to the number of displayed items: Assuming that the extra processing cost for Type D stimuli is additive for each additional item that must be scanned, the U-only slopes should be shallower than both the D-only and the mixed slopes.[2]

Experiment 1 provides the basic test of the two models. Experiment 2 then inquires whether the type of processing, object-based or space-based, might depend on the type of search that is conducted: serial (attentional) or parallel (preattentive). Finally, Experiments 3 and 4 examine a more sophisticated version of the object-based model that incorporates grouping strengths and hierarchical object-based structure.

Experiment 1

Experiment 1 was designed to provide the initial test of the object-based and space-based models, using the conjunctive-search task with overlapping stimuli just described. The participants searched for two target conjunctions (blue S, red V) in separate blocks, with the three matrix color-scheme conditions (U-only, D-only, or mixed) also presented in separate blocks for each target.

Method

Participants. Twelve 1st-year psychology students at the University of Haifa, Haifa, Israel, participated in the experiment for course credit. All had normal or corrected-to-normal acuity and normal color vision.

Stimuli and apparatus. Stimulus presentation and data acquisition for this and all subsequent experiments were controlled by an Apollo DN-4000 computer workstation (Hewlett-Packard Co., Palo Alto, California) with color monitor. The stimuli were overlapping forms, each stimulus composed of both a letter form, either S or V, and a square form, centered at the same display position. At a viewing distance of 110 cm, both letter forms subtended a visual angle of 1.56° in height and 0.88° in width. The square form subtended a visual angle of 1.20° × 1.20°.

Each stimulus was colored according to one of two *stimulus color schemes*: For Type U (unified) stimuli, the letter was colored, either red (CIE [Commission Internationale de l'Eclairage] coordinates: $x = .610$, $y = .342$, lm = 8.2 cd/m^2) or blue ($x = .151$, $y = .064$, lm = 3.1 cd/m^2), and the square was black (lm = 0); for Type D (divided) stimuli, the square was colored (red or blue) and the letter was black. In order to perceptually overlap the stimuli in depth, the two right points of intersection between the letter and the square bore the color of the letter (red, blue, or black), whereas the two left points of intersection bore the color of the square. The stimuli were displayed on a light background (CIE coordinates: $x = .351$, $y = .375$, lm = 29.6 cd/m^2).

Pilot testing was conducted using a set of simple-feature search tasks to ensure that the discriminability of the task-relevant features would be closely matched for the Type U and Type D stimuli. No differences were found in the search for a particular letter shape (i.e., S or V) using letters that were colored (Type U stimuli) versus letters that were black (Type D stimuli). Likewise, no differences were found in the search for a particular color (i.e., red or blue) using colored letters (Type U stimuli) versus colored squares (Type D stimuli). These tests used both the overlapping forms and the individual component forms as stimuli.

---

[2]With regard to the comparison between the D-only and the mixed conditions, the object-based model is ambivalent: On the one hand, mixed slopes might be expected to fall half way between U-only and D-only slopes, because on the average, only half of the stimuli on each display (i.e., the Type D stimuli) would in fact require divided attention. On the other hand, it may be difficult for participants to determine which stimuli on a given display require divided attention and which do not, and hence, mixed matrices might actually be processed less efficiently than D-only, because of the lack of consistency in the coloring scheme (and ensuing processing algorithm) across items.

Array size was varied by presenting either 1, 2, 4, 8, or 16 stimuli on each display. (The single-stimulus displays, which do not allow mixed color schemes, were included for exploratory purposes and are not considered further.) Arrays of 16 stimuli were presented in a 4 × 4 matrix, covering a 8.4° × 8.4° field. Smaller arrays used subsets of these same stimulus positions, as follows: Two-stimulus arrays were presented in either the top two or the bottom two of the four central positions; four-stimulus arrays were presented in the four central positions; eight-stimulus arrays were presented in the two middle rows of the matrix.

Each array was displayed in one of three matrix color schemes: U-only matrices included only Type U stimuli; D-only matrices included only Type D stimuli; mixed matrices included equal numbers of Type U and Type D stimuli, randomly mixed among the stimulus positions.

Targets were defined as the conjunction of a particular color and a particular letter-shape at the same stimulus position, regardless of the stimulus coloring scheme. Two target conjunctions were used (in separate blocks), blue and S and red and V. On target-absent trials, half of the distractors shared the target shape and the other half shared the target color, and these were randomly assigned to the occupied array positions. On target-present trials, one of the distractors was randomly replaced by a target. For mixed matrices, the color scheme of the target stimulus was chosen at random.

Procedure. Each participant took part in a single session consisting of six blocks, one block for each combination of target conjunction (two levels) and matrix color scheme (three levels). Block order was counterbalanced across participants, with matrix color scheme as the inner blocking factor.

Each block was introduced by a special display of the current target conjunction in the appropriate stimulus color scheme or schemes. At the beginning of the session, participants were shown a series of six such displays in the appropriate (block) order and were instructed regarding the task. They were told that they would be presented with six blocks of matrices containing varying numbers of items, each item composed of one of two letters (S or V) and an overlapping square. In some blocks the letters would be colored and the squares would be black; in other blocks the squares would be colored and the letters would be black; and in yet other blocks colored-letter and colored-square items would be presented on the same display. Nevertheless, as far as they were concerned, it was unimportant whether it was the letter or the square that was colored: The target was defined as a single composite item containing both the color blue and the letter S, or both the color red and the letter V (as appropriate).

Each trial began with the display of a black fixation point in the center of the screen, approximately 500 ms before the onset of the stimulus array. The array remained on the screen until the participant responded by pressing a key to indicate that the target was present (dominant hand) or absent (nondominant hand). One second later, the fixation point for the next trial appeared. Participants were asked to respond as quickly as possible on each trial while making as few errors as possible. Error feedback was provided by a short tone.

Sixteen repetitions of array size (five levels) and target presence (two levels) were randomly intermixed, for a total of 160 experimental trials in each of the six blocks. An additional 40 practice trials (4 repetitions) were presented at the beginning of each block. Participants were allowed to rest between blocks. The sessions lasted about 1 hr.

Results and Discussion

For each participant, mean search latencies (response times [RTs]) and error rates were computed for each cell of the design, after trimming outliers (RT < 200 ms or RT > 3,500 ms; fewer than 0.1% of the total number of trials). The mean RTs are based on correct responses only. Because preliminary analyses indicated that there was essentially no difference in the pattern of results for the two target conjunctions, the results were collapsed across targets. Figure 4 plots the mean RT as a function of array size, target presence, and matrix color scheme. In addition, the slopes, *y*-intercepts, and squared correlations (r^2) of the RT × Array Size functions for each participant were computed by linear regression analysis. The means of these, together with the error rate for each condition,[3] are presented in Table 1.

Inspection of the results indicates, as predicted by the object-based model, that matrix color scheme had a substantial effect on the search rates, with the U-only matrices searched more efficiently than the D-only and mixed matrices. A two-way analysis of variance (ANOVA), Color Scheme × Target Presence, on the search slopes confirmed this conclusion: The main effects of color scheme, $F(2, 22) = 17.81, p < .0001$, and target presence, $F(1, 11) = 29.61, p < .001$, were both significant, as was the Color Scheme × Target Presence interaction, $F(2, 22) = 9.81, p < .001$. The color-scheme effect was more pronounced for the target-absent trials than for the target-present trials, but it was significant in both cases, $F(2, 22) = 16.59, p < .0001$, and $F(2, 22) = 14.96, p < .0001$, respectively.

In addition, planned comparisons to check the source of the color-scheme effects were carried out separately for the target-present and the target-absent trials: On target-present trials, the U-only matrices were searched more efficiently than the D-only matrices, $F(1, 22) = 23.80, p < .0001$, than the mixed matrices, $F(1, 22) = 20.99, p < .0001$, and than the D-only and mixed matrices combined, $F(1, 22) = 29.83, p < .0001$. There was no difference between the D-only and the mixed matrices, $F < 1$. Similarly, on target-absent trials, the U-only matrices were searched more efficiently than the D-only matrices, $F(1, 22) = 14.43, p < .001$, than the mixed matrices, $F(1, 22) = 31.92, p < .0001$, and than the D-only and mixed matrices combined, $F(1, 22) = 29.75, p < .0001$. Here the search rate for the D-only matrices was marginally faster than for the mixed matrices, $F(1, 22) = 3.43, p < .08$.

A further set of analyses focused on the target-present trials for the mixed matrices only, examining whether the color scheme of the target itself (target color scheme) had any effect on search performance. The average RT × Array Size search functions for Type U and Type D targets, and the corresponding error rates, are plotted in Figure 5. A two-way ANOVA, Target Color Scheme × Array Size, indicates that the search latencies were indeed faster for Type U targets than for Type D targets, $F(1, 11) = 60.71, p < .0001$. There was also a significant interaction with array size, $F(3, 33) = 7.11, p < .001$. (The average slopes of the best fitting linear

[3]The error rates for this and all other experiments reported here were quite low, and the within-subject correlations between error rates and search latencies across the experimental conditions were also generally low and positive (r averaging .03, .12, .25, and .16 in Experiments 1 to 4, respectively). In no case was there any indication of a speed–accuracy tradeoff.

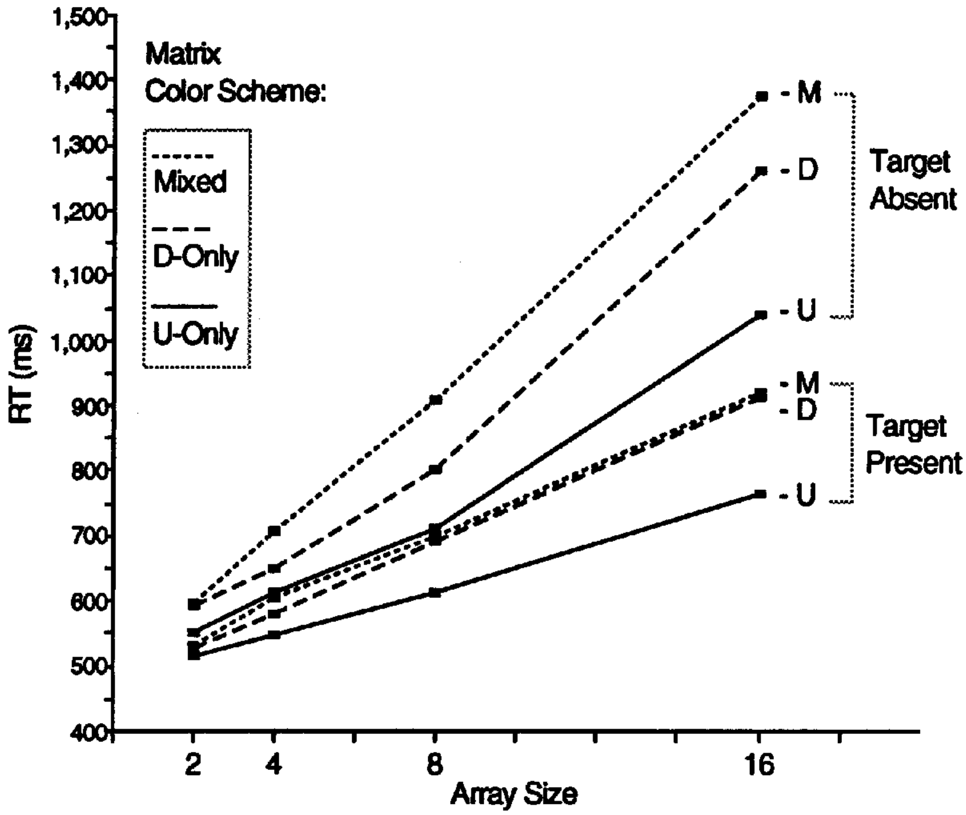


Figure 4. Results from Experiment 1: Mean response time (RT) as a function of array size for each Matrix Color Scheme × Target Presence condition. M = mixed; D = D-only; U = U-only.

search functions for Type U and Type D targets were 23.7 ms and 32.3 ms, respectively.) In addition, a similar ANOVA on the error rates indicates that significantly more errors (misses) were made when the target was Type D than when it was Type U, $F(1, 11) = 6.46$, $p < .05$. There was no interaction with array size, $F(3, 33) = 1.19$, *ns*.

In sum, the results regarding both matrix color scheme and target color scheme clearly favor the object-based model: First, they imply an object-based representation underlying feature integration. Unless specific feature–object linkages (rather than just feature–location linkages) are represented in the master-map medium, the differences in such linkages between the Type U and the Type D stimuli could not exert an effect. Second, the color-scheme effects were observed even though the specific feature–object linkages for each item were irrelevant to the task. This suggests that not only is the object-based structure of the stimuli represented in the attentional medium, but that also the attentional feature-binding process is in fact constrained by this preattentively derived structure (cf. Duncan, 1984; Kahneman & Henik, 1981; Prinzmetal, 1981).

We should, however, consider some potential objections. One involves the possibility that the Type D matrices were more difficult to process spatially than were the Type U matrices. For instance, if the participants focused on a small part of the contour of the letter component of each item, they could perhaps extract both shape and color information from the Type U stimuli, but they would need either to shift or to expand their attentional focus in order to extract the color information from the square portion of the Type D stimuli. Such an account of the data cannot strictly be ruled out. However, it is questionable whether focusing on a small portion of the letter would in fact be beneficial for the extraction of shape information, particularly if it also takes time to narrow one's focus (Eriksen & St. James, 1986). It is also unclear how such a search strategy itself can be accounted for within the FIT framework (see Figure 2), which assumes that attention selects filled locations rather than subsets of "pixels." Nevertheless, this objection is

Table 1
Mean Slopes (in Milliseconds), Intercepts, r^2, and Error Rates (Percent Errors) of the Search Functions for Each Condition in Experiment 1

Matrix color scheme	Target present				Target absent			
	Slope (ms)	Int.	r^2	% error	Slope (ms)	Int.	r^2	% error
U-only	18.1	473	.97	5.0	35.0	467	.97	1.4
D-only	27.8	469	.97	5.3	48.8	460	.98	0.6
Mixed	27.2	484	.97	5.3	55.5	481	.99	1.3

Note. Int. = intercept.

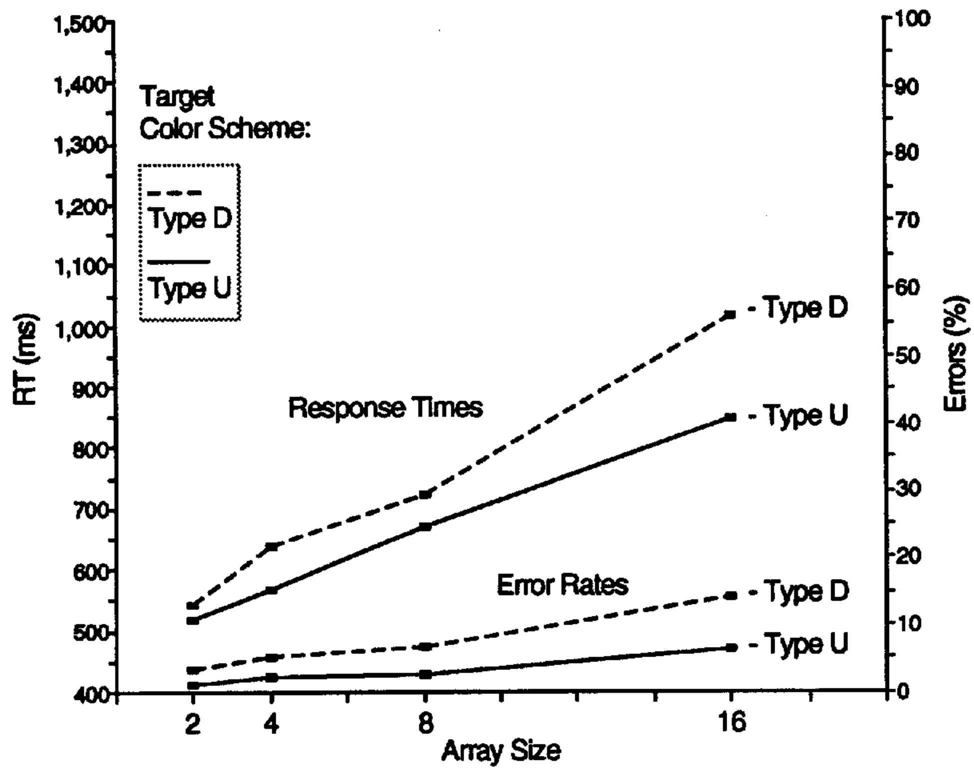


Figure 5. Results from Experiment 1 for mixed matrices, target-present trials only: Mean response time (RT) and error rate as a function of array size and target color scheme.

taken seriously, and further converging evidence was brought to bear against it in Experiments 3 and 4.

A second objection concerns whether the task used here constitutes a fair test of the space-based view. After all, doesn't the task itself require that the stimulus items be parsed into their component parts (i.e., letter and square), so that the target letter-shape can be distinguished from the distractor letter shape? Here a point of clarification is in order. Like the object-based model, the FIT framework allows both for the operation of preattentive segmentation and grouping processes and for the registration of stimulus features at various spatial scales (see, e.g., Heathcote & Mewhort, 1993; Treisman, 1992a, 1992b; and see Experiments 3 and 4 here). Thus, the space-based model too would hold that the relevant letter-shape features are available in the preattentive representation (see Rensink & Enns, 1995). Unlike the object-based model, however, which assumes that the registered features are linked to specific perceptual objects or gestalts, the space-based model assumes that they are linked to a common spatial location (hence the importance of equating the spatial scale of the letter and square components and of centering them on the same spatial coordinates). This is the key issue examined here, with the results so far supporting the object-based view.

Finally, however, a third objection concerns possible differences in the internal target definition or "template" (Duncan & Humphreys, 1989, 1992; Treisman, 1991) that is used to perform the search. Because of the blocked presentation, participants may have adopted a more complex target template in the D-only and mixed conditions (e.g., one that includes the square as well as the letter) than in the U-only condition, thereby increasing the difficulty of the feature-matching operations in the former conditions. Several points may be raised against this explanation of the results, however. First, the very fact that participants would need to adopt a different template for the various matrix color schemes would in itself suggest that they cannot simply search for conjunctions of color and form at particular locations in a purely space-based manner. Second, differences between target templates could not explain the effects of target color scheme in the mixed matrices, because these effects were observed within a particular matrix color-scheme condition.

Third, randomly mixing the matrix color-scheme conditions within blocks (Goldsmith, 1995, Experiment 1), thereby reducing the possibility that participants will employ a different target template for each condition (cf. Treisman, 1991, Experiment 1), yields essentially the same pattern of effects observed here: The mean search slopes (in milliseconds) for the U-only, D-only, and mixed conditions were 17.9, 20.5, and 22.5, respectively, on target-present trials, and 39.2, 52.1, and 52.9 respectively on target-absent trials (a significant effect in both cases). Indeed, the similar pattern for both blocked and unblocked presentations adds further support for the existence of object-based constraints on the feature-binding process, because the color-scheme effects were observed both when participants were unprepared for a particular matrix color scheme (Goldsmith, 1995, Experiment 1) and under conditions that encouraged them to process each color scheme in the most efficient way possible (Experiment 1 here). The generality of these findings is examined further, however, in the next experiment.

Experiment 2

So far, the results have supported the object-based model of feature integration in visual search. However, it is important to consider possible differences between two components of the conjunctive-search process postulated in the FIT framework: In addition to the serial attentional-scan component, FIT and related models have more recently added a preattentive (or widely spread attentional)[4] parallel-guidance component as well (e.g., Cave & Wolfe, 1990; Cohen & Ivry, 1991; Treisman & Sato, 1990; Wolfe, 1994a). This modification was motivated by findings indicating that under certain conditions, characterized generally by high feature discriminability, conjunctive-search latencies do not increase much or at all as the number of distractors is increased (e.g., Cohen & Ivry, 1991; Duncan & Humphreys, 1989; Nakayama & Silverman, 1986; Treisman & Sato, 1990; Wolfe, Cave, & Franzel, 1989). Thus, for instance, in a modification of FIT called *guided search* (Cave & Wolfe 1990; Wolfe, 1994a; Wolfe et al., 1989), preattentive top-down activation of the target features, summed in parallel across the feature maps, guides the attentional spotlight to the most probable target locations. To the extent that attention is led directly to the target location (under conditions of high feature discriminability and low noise), search latencies are more or less independent of the number of distractors. A similar revision was proposed by Treisman and Sato (1990). However, rather than activating stimulus locations containing the target features, their feature inhibition mechanism inhibits stimulus locations containing the distractor features.

Experiment 2 is designed to examine whether the object-based or space-based nature of feature binding might differ for the postulated serial-scan and parallel-guidance components of visual search. In order to address this question, the same conjunctive-search task as in the previous experiment was used, but the overall difficulty of the search was manipulated by using two levels of target-distractor discriminability (high vs. low) and two levels of display density (spread vs. packed). Both of these factors have been implicated in findings of parallel or near-parallel conjunctive search (e.g., Cohen & Ivry, 1991; Duncan & Humphreys, 1989, 1992). Here, these manipulations are intended to have a substantial combined effect on the search slopes for the U-only matrices: the packed low-discriminability condition yielding the steepest slopes, reflecting a predominantly serial, attentional search, and the spread high-discriminability condition yielding relatively flat search functions, reflecting a predominantly preattentive, guided search. In accordance with the results of Experiment 1, in which the search functions were fairly steep, matrix color scheme is expected to have a substantial impact on the efficiency of the attentional-scan component in the packed low-discriminability condition (indicating object-based feature binding). The crucial question, however, is whether the preattentive, guided-search component will be similarly affected. That is, assuming that U-only performance indicates a predominantly parallel search process in the spread high-discriminability condition, will the effects of matrix color scheme still be observed?

Method

Participants. Twelve 1st-year psychology students at the University of Haifa, Haifa, Israel, participated for course credit. All had normal or corrected-to-normal acuity and normal color vision.

Stimuli. Each stimulus was composed of both a letter form, either X or O, and a rectangle form, centered at the same display position. At a viewing distance of 90 cm, both letter forms subtended a visual angle of 0.76° in height and 0.41° in width. The rectangle form subtended a visual angle of 0.41° in height and 0.76° in width.

As in Experiment 1, each stimulus was colored according to one of two stimulus color schemes, Type U or Type D, and once again the stimuli were perceptually overlapped in depth by appropriate coloring of the four points of intersection. In this experiment, however, the stimuli were displayed on a black background (lm = 0), and a different set of colors was used in each of two stimulus-discriminability conditions: In the *high-discriminability* condition, the task-relevant colors were green (CIE coordinates: $x = .230$, $y = .650$, lm = 21.5 cd/m^2) and pink–magenta ($x = .381$, $y = .203$, lm = 11.3 cd/m^2), and the neutral color was white ($x = .330$, $y = .352$, lm = 24.0 cd/m^2). In the *low-discriminability* condition, the task-relevant colors were orange (CIE coordinates: $x = .489$, $y = .441$, lm = 10.4 cd/m^2) and red ($x = .455$, $y = .248$, lm = 8.4 cd/m^2), and the neutral color was gray ($x = .330$, $y = .352$, lm = 13.9 cd/m^2).

Pilot testing was again conducted using simple-feature search tasks to ensure that for both stimulus-discriminability conditions, the discriminability of the task-relevant features would be approximately equal for the Type U and Type D stimuli. No differences were found in the search for a particular letter shape using colored versus achromatic letters. Likewise, no differences were found in the search for a particular color using colored letters versus colored rectangles. Separate tests were conducted for each stimulus-discriminability condition and each display density.

Arrays of either 4, 9, or 16 stimuli were presented in one of two display densities, spread or packed. Spread arrays of 16 stimuli were presented in a slightly irregular 4 × 4 matrix (randomly jiggled by up to 0.05° of visual angle), covering a 6.2° × 6.2° field. The average distance between the centers of adjacent stimuli was 1.8°. Arrays of 9 and 4 stimuli were displayed on randomly chosen 3 × 3 and 2 × 2 subsets of these same stimulus positions. For the packed arrays, an additional position was added between every two spread positions, yielding a 7 × 7 matrix of potential stimulus locations. In this case, the average distance between positions was

[4]The term *preattentive* is commonly used in the visual-search literature to refer to search processes that operate in parallel across all stimuli in the visual field, before attention is (serially) focused on particular stimulus items. As such, it is somewhat of a misnomer, and in fact is better conceived of as widely spread attentional processing (see, e.g., Cavanagh, Arguin, & Treisman, 1990; Treisman & Gormican, 1988; Wolfe, 1994a; and see General Discussion).

0.9° (the stimuli were almost touching). Arrays of 4, 9, and 16 stimuli were displayed by randomly choosing 2 × 2, 3 × 3, and 4 × 4 subsets of these more densely packed positions.

A single target conjunction was used for each stimulus-discriminability condition: In the high-discriminability condition, the target was green and X among green O and pink X distractors. In the low-discriminability condition, the target was orange and X among orange O and red X distractors. The three matrix color schemes were U-only, D-only, and mixed.

Procedure. Each participant took part in two separate sessions conducted on consecutive days. There were two session types, in which either stimulus discriminability or display density was varied between sessions. When stimulus discriminability was varied between sessions, display density and matrix color scheme were varied within each session in separate blocks, with display density as the outer blocking factor. When display density was varied between sessions, stimulus discriminability replaced display density as the outer blocking factor within each session. Session type, session order, and block order were counterbalanced across participants.

The instructions, the chronology of each trial, and the manner of responding were essentially the same as in Experiment 1. However, a different practice regimen was employed: At the beginning of each session, as well as when either the discriminability or the density condition was changed at the midpoint of each session, participants were given a warm-up period consisting of three blocks (60 trials each), one block for each matrix color-scheme condition (in the proper order). Each warm-up period was then followed by nine more blocks, three consecutive blocks for each of the three color-scheme conditions. Of these, the first block of each new color-scheme condition was discarded as practice, leaving two blocks of experimental trials per condition. Within each block, 10 repetitions of array size (three levels) and target presence (two levels) were randomly intermixed. Participants were allowed to rest between blocks and were required to take a 10-min break at the midpoint of each session. The sessions lasted about 90 min.

Results and Discussion

For each participant, mean RTs (errorless trials only) and error rates were computed for each cell of the design, after trimming outliers (RT < 200 ms or RT > 5,000 ms; fewer than 0.05% of the total number of trials). Figure 6 presents the average RT × Array Size functions for each combination of stimulus discriminability, display density, target presence, and matrix color scheme. As before, the best fitting linear-regression functions were also computed, and the average slope, *y*-intercept, and r^2 for each of these, together with the error rates, are presented in Table 2.

Inspection of the results indicates a substantial effect of stimulus discriminability on the Type U search slopes, with a lesser effect of display density appearing primarily for the high-discriminability displays.[5] For the purposes of this article, however, it is enough to note that the joint manipulation of these factors did in fact yield performance ranging from slow, fairly steep search functions for the packed (and the spread) low-discriminability condition, representative of serial search, to fast, relatively flat functions for the spread high-discriminability condition, suggesting a substantial contribution of the parallel, guided-search–feature-inhibition component.[6]

The critical question is whether the effects of matrix color scheme were also modulated by differences in the search conditions. A four-way ANOVA, Discriminability × Density × Target Presence × Matrix Color Scheme on the search slopes yielded the following effects and interactions involving matrix color scheme: First, as in Experiment 1, there was a significant main effect for matrix color scheme, $F(1, 11) = 12.30$, $p < .0005$, and a (marginally) significant Color Scheme × Target Presence interaction, $F(2, 22) = 3.24$, $p < .06$. In addition, however, there was also a significant Color Scheme × Density interaction, $F(2, 22) = 9.23$, $p < .005$, a significant Color Scheme × Density × Target Presence interaction, $F(2, 22) = 7.69$, $p < .0001$, and a marginal Color Scheme × Discriminability interaction, $F(2, 22) = 2.73$, $p < .09$.

Separate analyses for the target-present and target-absent trials helped clarify these interactions: For the target-present trials, a three-way ANOVA yielded a significant Color Scheme × Discriminability interaction, $F(2, 22) = 7.67$, $p < .005$, indicating greater color-scheme effects for low-discriminability than for high-discriminability displays, but the Density × Color Scheme interaction was not significant, $F(2, 22) = 2.00$, $p > .15$. Conversely, for the target-absent trials, the Density × Color Scheme interaction was significant, $F(2, 22) = 12.85$, $p < .0005$, indicating greater color-scheme effects for packed than for spread arrays, but here the Discriminability × Color Scheme interaction was not significant, $F < 1$. Thus, both discriminability and density were found to modulate the color-scheme effects, but the pattern is different for the target-present and the target-absent trials.

[5]For the interested reader, a three-way ANOVA, Discriminability × Density × Target Presence, on the Type U search slopes yielded significant main effects for target presence, $F(1, 11) = 72.12$, $p < .0001$, and discriminability, $F(1, 11) = 29.95$, $p < .0005$, but not for density, $F < 1$. In addition, all of the two-way interactions were either significant or marginal, as was the three-way interaction: Discriminability × Density, $F(1, 11) = 18.37$, $p < .005$; Discriminability × Target Presence, $F(1, 11) = 23.90$, $p < .0005$; Density × Target Presence, $F(1, 11) = 4.13$, $p < .07$; three-way, $F(1, 11) = 8.42$, $p < .05$. On target-present trials, discriminability had a significant effect (high more efficient than low) regardless of display density, $F(1, 11) = 24.38$, $p < .0005$, whereas density had an effect (spread more efficient than packed) only in the high-discriminability condition, $F(1, 11) = 5.14$, $p < .05$. On target-absent trials, the main effect of discriminability was again significant, $F(1, 11) = 29.59$, $p < .0005$, but here there was a crossover interaction between density and discriminability, $F(1, 11) = 15.12$, $p < .005$: Spread arrays were searched more efficiently than packed arrays when discriminability was high, $F(1, 11) = 3.21$, $p < .10$ (only marginal), but packed arrays were actually searched more efficiently than spread arrays when discriminability was low, $F(1, 11) = 13.55$, $p < .005$.

[6]Inferences regarding parallel versus serial processing in visual search are typically based on comparisons of target-present performance. Target-absent performance often reflects additional verification or double-check strategies, making it less diagnostic (see Cave & Wolfe, 1990; Wolfe, 1994a).

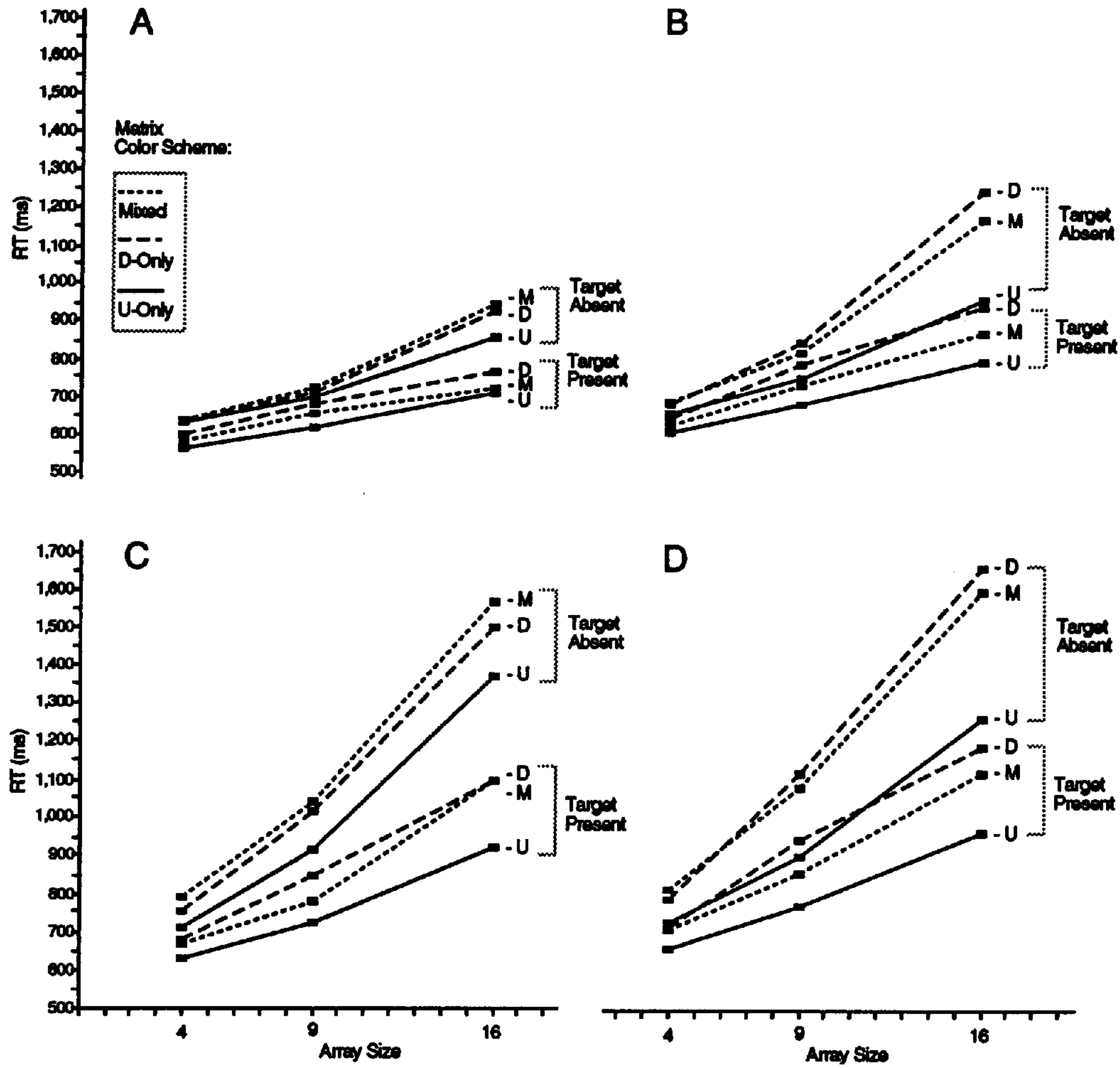


Figure 6. Results from Experiment 2. Mean response time (RT) as a function of array size, matrix color scheme, and target presence, plotted separately for each Stimulus Discriminability × Display Density condition: high-discriminability spread display (A), high-discriminability packed display (B), low-discriminability spread display (C), and low-discriminability packed display (D). M = mixed; D = D-only; U = U-only.

Of particular importance for this study is the comparison of the pattern of color-scheme effects in the two extreme conditions, spread high discriminability and packed low discriminability. In the latter condition, which presumably reflects a predominantly serial, focused-attentional search process, there were strong color-scheme effects on the search slopes for both the target-present trials, $F(2, 22) = 7.60, p < .005$, and the target-absent trials, $F(2, 22) = 12.74, p < .0001$. By contrast, in the spread high-discriminability condition, which presumably reflects a predominantly parallel, preattentive search process, there was no color-scheme effect for the target-present trials, $F < 1$ (also, $F < 1$ for all planned comparisons), and only a marginal color-scheme effect for the target-absent trials, $F(2, 22) = 2.64, p < .10$. (For the target-absent trials, however, the planned comparison between the U-only condition and the mean of the D-only and mixed conditions was significant, $F(1, 22) = 5.13, p < .05$.)

This pattern is reinforced by comparing the effects of the *target* color scheme in the mixed matrices under the different search conditions (see Figure 7). On the one hand, in the spread high-discriminability condition, there was no differ-

Table 2
Mean Slopes, Intercepts, r^2, and Error Rates of the Search Functions for Each Condition in Experiment 2

	Target present				Target absent			
Matrix color scheme	Slope (ms)	Int.	r^2	% error	Slope (ms)	Int.	r^2	% error
High-discriminability spread display								
U-only	12.2	511	.93	7.7	19.5	540	.97	4.4
D-only	13.5	549	.87	8.1	25.0	518	.95	3.1
Mixed	11.5	541	.83	8.8	26.2	515	.95	2.4
High-discriminability packed display								
U-only	15.8	539	.93	6.7	25.4	541	.98	2.2
D-only	24.1	554	.97	9.4	47.0	465	.97	2.1
Mixed	20.1	545	.95	6.2	40.4	498	.94	1.9
Low-discriminability spread display								
U-only	24.4	521	.97	7.1	56.0	457	.98	2.4
D-only	34.7	538	.99	7.8	62.6	483	.97	0.8
Mixed	36.6	494	.95	9.0	65.9	498	.97	2.0
Low-discriminability packed display								
U-only	25.4	553	.99	7.1	44.8	532	.98	1.9
D-only	39.4	569	.97	8.5	73.0	484	.98	2.8
Mixed	34.4	564	.95	7.8	65.9	531	.97	2.6

Note. Int. = intercept.

ence in RT for Type U and Type D targets, $F < 1$, nor was the Color Scheme × Array Size interaction significant, $F(2, 22) = 1.90$, *ns*. In fact, looking at the error rates, there was actually a tendency for more errors (misses) when the target was Type U than Type D, $F(1, 11) = 4.45, p < .06$. On the other hand, for the packed low-discriminability condition, Type D targets were substantially slower to detect than Type U targets, $F(1, 11) = 42.21$, $p < .0001$, and this difference increased with increasing display size, $F(1, 11) = 9.82$, $p < .0001$. Furthermore, there were also significantly more errors for Type D than Type U targets, $F(1, 11) = 13.15, p < .005$.

In several respects, then, the processing of spread high-discriminability displays was found to be insensitive to stimulus color scheme. However, performance in the *packed* high-discriminability condition was also relatively insensitive to the target color scheme in the mixed matrices: Neither the main effect of color scheme on RT nor the Color-Scheme × Array Size interaction was significant ($F < 1$ for both), though there was a significant Color-Scheme × Array Size interaction on the error rates, $F(2, 22) = 4.11$, $p < .05$. Nevertheless, returning to consider the effects of *matrix* color scheme on the search slopes (see Figure 6), unlike the spread high-discriminability condition, the packed high-discriminability condition yielded significant color-scheme effects for both target-present trials, $F(2, 22) = 5.11$, $p < .05$, and target-absent trials, $F(2, 22) = 7.44, p < .005$. Thus, only in the spread high-discriminability target-present condition was performance found to be entirely indifferent to the color scheme of the stimuli.

Of course, one might argue that the absence of color-scheme effects in that condition was due simply to the very fast response times. By this account, however, differences in the magnitude of the observed color-scheme effects across the various conditions would be expected to mirror differences in the baseline levels of Type U performance, but they did not: For example, on target-absent trials, stimulus discriminability had a substantial effect on U-only performance (see Table 2), yet discriminability did not modulate the color-scheme effects for these trials—density was the critical factor modulating these effects. (Additional results discounting this interpretation are presented in Experiments 3 and 4.)

Taken at face value, then, the results suggest that the sensitivity of the feature-binding process to object structure may in fact depend on the conditions of the search: Whereas the serial, attentional-scan component of visual search appears to be object based, the parallel, preattentive-guidance component may in fact be space based. This idea is consistent with some earlier results as well: On the one hand, under fairly difficult search conditions, Grabowecky and Khurana (1990) found that separating the target features between two adjacent (coterminating) line segments substantially impaired the search efficiency compared to a condition in which both features pertained to the same line segment. On the other hand, using highly discriminable stimuli under conditions that allowed near-parallel conjunctive search, Wolfe et al. (1990, Experiment 6) found that separating the feature-carrying parts of each stimulus item (e.g., an achromatic letter C and an attached, colored bar) did not impair the efficiency of the search. (Note that in contrast to the stimuli used so far in this study, in both of these earlier studies there was a clear spatial separation between the stimulus parts.)

Curiously, however, the characterization of the attentional-scan component as object based and the preattentive-guidance component as space based is in direct contrast to a suggestion made in connection with FIT's feature-inhibition mechanism (Treisman, 1993; Treisman & Sato, 1990). In discussing the operation of the mechanism, Treisman and

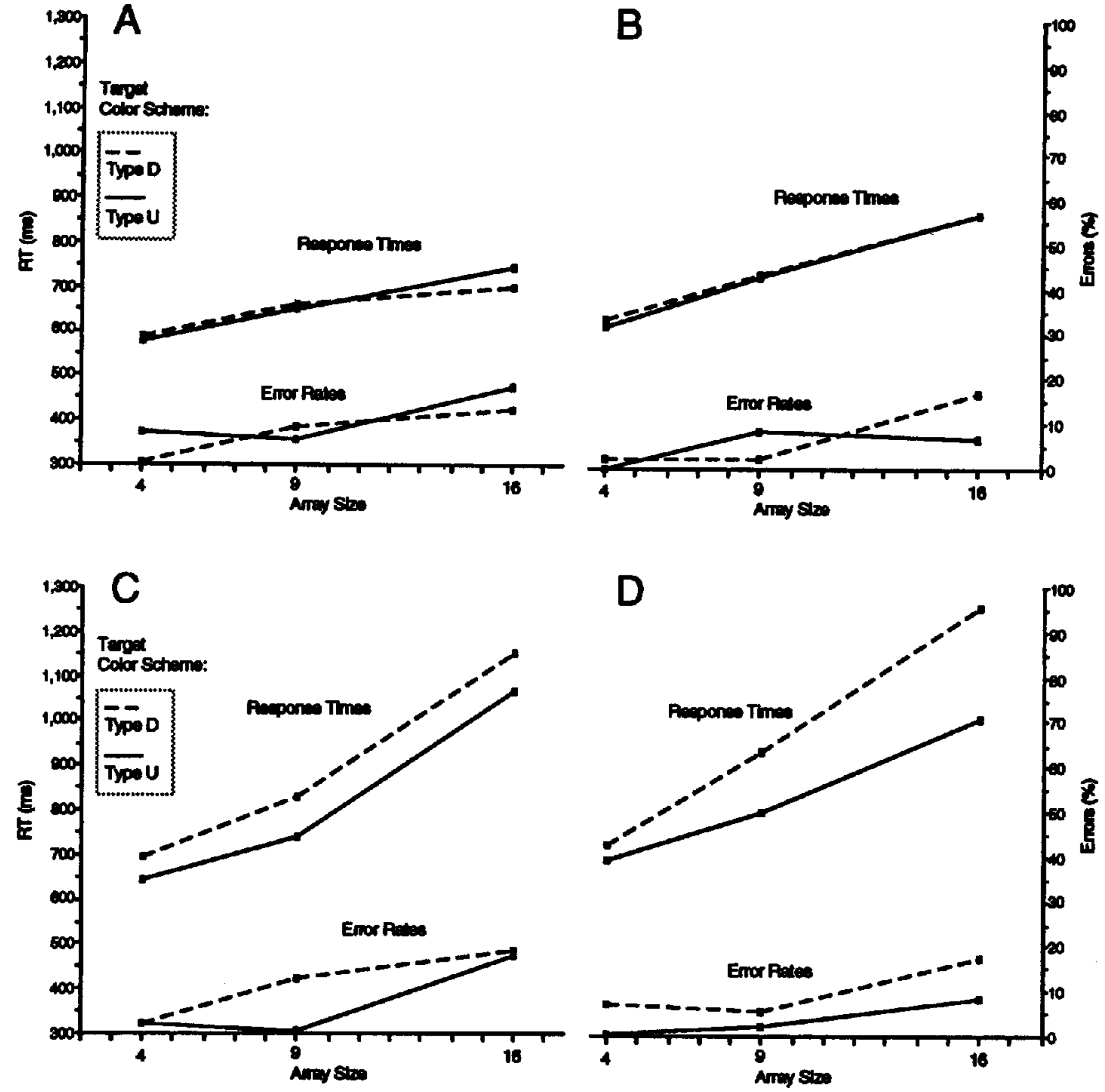


Figure 7. Results from Experiment 2 for mixed matrices, target-present trials only. Mean response time (RT) and error rate as a function of array size and target color scheme, plotted separately for each Stimulus Discriminability × Display Density condition: high-discriminability spread display (A), high-discriminability packed display (B), low-discriminability spread display (C), and low-discriminability packed display (D).

Sato (1990) state that "the locations that get inhibited are not the global areas in which patterned elements are located, but rather the specific points occupied by the inhibited features" (p. 476). This, it is suggested, could

> explain conditions in which attention appears to select objects rather than locations. . . . The selection would this time be on a finer scale, corresponding to one set of overlapped lines, and without the constraints imposed by the shape and unity of the externally controlled attention window. (p. 462; see also Treisman, 1993, for an elaboration of this idea)

From the perspective of the present study, the relegation of object-based selection entirely to the feature-inhibition mechanism would seem to be problematic: First, FIT assumes that both the attentional-scan and the feature-inhibition components operate in concert on the same underlying master-map of stimulus locations. Yet, activation or inhibition from the feature maps could serve the selection of individual overlapping forms only if the basic unit of representation in the master map is *perceptual object* rather than location. Hence, an object-based master-map representation would need to underlie both components. Moreover, assuming that a common object-based representation is used by the serial-scan and guided-search–feature-inhibition components, the results of this study so far would imply that it is the parallel-guidance processing, rather than the attentional-scan processing, that is space based, with activation or inhibition from the feature maps spreading to the global stimulus locations.

There is, however, an alternative account of these results that has yet to be considered: Perhaps under conditions that appear to implicate space-based processing, it is not the *global location* of each stimulus item that constitutes the unit of processing but rather a more *global perceptual object*. This idea, which implies a more refined version of the object-based model, is examined in the following two experiments.

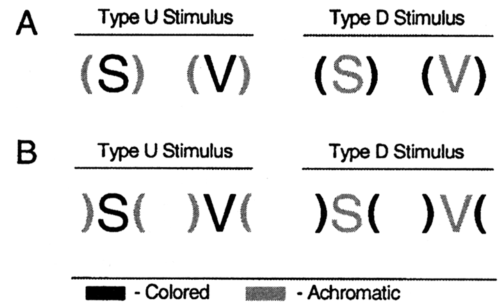
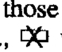
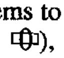


Figure 8. Stimuli used in Experiment 3: high gestalt (A) and low gestalt (B).

Experiment 3

In its most simple form, the object-based representation assumed by the object-based model includes only primitive object markers. However, many previous discussions have pointed out the need to incorporate some type of hierarchical object structure, whereby objects and their properties are represented hierarchically at different levels of globality (e.g., Garner, 1978; Marr, 1982; Navon, 1983; Palmer, 1977). Thus, in a more refined, hierarchical version of the object-based model, the display as a whole might constitute the highest level object, with lower levels coding groups of items, then individual stimuli, then their parts, and so forth (see also Baylis & Driver, 1993; Duncan & Humphreys, 1989, 1992; Treisman, 1992a). In addition, explicit grouping strengths could be included in order to capture differences in the structural cohesiveness of individual objects or groups of objects (Duncan & Humphreys, 1989, 1992; Kramer & Jacobson, 1991; Pomerantz & Pristach, 1989; Wertheimer, 1923).

Assuming a hierarchically structured object-based medium underlying feature integration, the absence of color-scheme effects need not indicate that features are being conjoined spatially. Rather, the search process could be using more unitary object representations at the global item level. In Experiment 2, for instance, the highly discriminable features (shape and color) and sparse spacing of the items may have allowed those items to be processed efficiently as global objects (i.e.,  vs. ), so that the coloring of the local parts would have little or no effect. Although speculative, this account has the advantage of invoking only a single, object-based mode of processing to explain both attentive and preattentive search performance.

Experiments 3 and 4 attempt to provide converging evidence that both the presence and the absence of color-scheme effects reflect sensitivity to objectness rather than to spatial separation per se. Using the same conjunctive-search task as before, in these experiments the critical comparisons are between the color-scheme effects obtained using different versions of Type U and Type D stimuli, in which the objectness or gestalt grouping of the relevant parts is manipulated, but the spatial separation between them is held constant.

Figure 8 depicts the stimuli used in Experiment 3. The relevant letter-shape features are S and V, and the added neutral element is a pair of flanking parentheses that either face toward the letter (*high-gestalt* stimuli) or away from the letter (*low-gestalt* stimuli). Although the spatial proximity of the component parts is equivalent for the high- and low-gestalt stimuli, the high-gestalt stimuli seem to form a much more unitary and cohesive global object (cf. Pomerantz, 1981). Previous work has shown that attention may be divided relatively easily between two parts that are perceived as forming a single, configural whole (e.g., Pomerantz, 1981; Pomerantz & Pristach, 1989; Treisman et al., 1983). Thus, the object-based model predicts that the color-scheme effects should be reduced for the high-gestalt stimuli compared to the low-gestalt stimuli. In contrast, the space-based model predicts no effect whatsoever for the gestalt manipulation. Only the spatial separation between the feature-carrying elements should affect the efficiency of Type D processing, and this separation is equivalent in both gestalt conditions.

In addition, both stimulus discriminability and display density were also manipulated. Neither factor, however, is expected to affect the basic pattern of results. With regard to stimulus discriminability, even if the high-discriminability Type D stimuli are easier to process as unitary objects than the low-discriminability stimuli (cf. the results of Experiment 2), the gestalt manipulation should affect the efficiency of such processing in either case. With regard to display density, unlike in Experiment 2, two levels of spread matrices were used. By spacing the stimuli widely enough, any gestalt effect should primarily reflect the internal cohesion or objectness of the stimulus items (intraitem grouping), rather than the interference between neighboring items (interitem grouping), or the difficulty of preattentively parsing the display (cf. Wolfe, 1994b). Nevertheless, a comparison of the magnitude of the gestalt effects between the two display densities allows a check on the extent to which interitem interference or parsing difficulty might also be contributing to these effects. Finally, in manipulating matrix color scheme, only Type U and Type D matrices were used (mixed matrices were omitted).

Method

Participants. Eight 1st-year psychology students at the University of Haifa, Haifa, Israel, participated for course credit. All had normal or corrected-to-normal acuity and normal color vision.

Stimuli. Each stimulus was composed of both a letter form, either S or V, and a pair of flanking parentheses. At a viewing distance of 90 cm., both letter forms subtended a visual angle of 0.76° in height and 0.41° in width. Each parenthesis subtended 0.76° in height and 0.36° in width (concavity). In the high-gestalt condition the parentheses faced inward (toward the letter), whereas in the low-gestalt condition they faced outward (see Figure 8). In either case, the most proximal point on the contour of each parenthesis (midpoint or endpoints) was offset 0.13° of visual angle horizontally from the most proximal part of the letter contour. Thus, the distance between the letter and the flanking parentheses was the same in both gestalt conditions.

Each stimulus was colored according to one of two stimulus color schemes, Type U (colored letter, achromatic parentheses) or Type D (achromatic letter, colored parentheses). Two matrix color schemes were used, U-only and D-only. Stimulus discriminability (high vs. low) was manipulated as in Experiment 2, using the same two sets of stimulus colors: In the high-discriminability condition, the target was green and S among green V and pink S distractors. In the low-discriminability condition, the target was orange and S among orange V and red S distractors.

Array size and display density were manipulated by presenting arrays of either 4, 9, or 16 stimuli in one of two display densities, *wide spread* or *medium spread*. Wide-spread arrays of 16 stimuli were presented in a slightly irregular 4 × 4 matrix (randomly jiggled by up to 0.05° of visual angle), covering a 8.0° × 8.0° field. The average distance between the centers of adjacent stimuli was 2.2°. Arrays of 9 and 4 stimuli were displayed on randomly chosen 3 × 3 and 2 × 2 subsets of these same stimulus positions. For the medium-spread arrays, the average distance between (jiggled) centers of adjacent stimuli was 1.8°, and these arrays (all three array sizes) were randomly positioned within the larger, wide-spread field.

Procedure. Each participant took part in two separate sessions conducted on consecutive days. There were two session types, in which either stimulus discriminability or stimulus gestalt was varied between sessions. When stimulus discriminability was varied between sessions, stimulus gestalt and matrix color scheme were varied within each session in separate blocks, with stimulus gestalt as the outer blocking factor. When stimulus gestalt was varied between sessions, stimulus discriminability replaced stimulus gestalt as the outer blocking factor within each session. Display density was manipulated within blocks, together with array size and target presence. Session type, session order, and block order were counterbalanced across participants.

The instructions, the chronology of each trial, and the manner of responding were essentially the same as in the previous experiments. Also, the same type of practice and test regimen was employed as in Experiment 2: At the beginning of each session, as well as when either the discriminability or the gestalt condition was changed at the midpoint of each session, participants were given a warm-up period consisting of two short blocks (60 trials each), one block for each matrix color-scheme condition (in the proper order). The warm-up blocks were then followed by six longer blocks (96 trials each), three consecutive blocks for each of the two color-scheme conditions. Of these, the first block of each new color-scheme condition was discarded as practice, leaving two blocks of experimental trials per condition. Within each of these blocks, eight repetitions of density (two levels), array size (three levels), and target presence (two levels) were randomly intermixed.

Results and Discussion

For each participant, mean RTs (errorless trials only) and error rates were computed for each cell of the design, after trimming outliers (RT < 200 ms or RT > 4,000 ms; fewer than 0.1% of the total number of trials). Figures 9 and 10 present the average RT × Array Size functions for each combination of stimulus gestalt, stimulus discriminability, display density, target presence, and matrix color scheme. The average slope, *y*-intercept, and r^2 for the best fitting linear functions, together with the error rates, are presented in Table 3.

A five-way ANOVA, Gestalt × Discriminability × Density × Target Presence × Color Scheme was performed on the slopes of the search functions. The reported results are confined to the effects and interactions of current interest (those involving matrix color scheme and gestalt). Focusing first on the effects of matrix color scheme (across the gestalt conditions), the main color-scheme effect was significant, $F(1, 7) = 89.36$, $p < .0001$, as was the Color Scheme × Target Presence interaction, $F(1, 7) = 34.05$, $p < .001$, the Color Scheme × Discriminability interaction, $F(1, 7) = 9.56$, $p < .05$, and the Color Scheme × Density interaction, $F(1, 7) = 9.96$, $p < .05$. The pattern resembles the one observed in Experiment 2: The color-scheme effect was less pronounced when the target was present, when the stimulus discriminability was high, and when the stimuli were more widely spread. In this experiment, however, the color-scheme effect remained significant under all of the search conditions (see later discussion).

Turning now to the gestalt manipulation, the main gestalt effect was marginally significant, $F(1, 7) = 5.37$, $p < .06$, but more importantly, the Gestalt × Color Scheme interaction, $F(1, 7) = 12.65$, $p < .01$, and the Gestalt × Color Scheme × Target Presence interaction, $F(1, 7) = 7.78$, $p < .05$, were both significant. Analyses of simple effects clarified these interactions: First, as predicted, the color-scheme effect was more pronounced for the low-gestalt stimuli (averaging 41.8 ms/item) than for the high-gestalt stimuli (averaging 20.2 ms/item), but it was significant in both cases, $F(1, 7) = 47.44$, $p < .0005$, and $F(1, 7) = 133.05$, $p < .0001$, respectively. Second, although the predicted interaction (between gestalt and color scheme) was more pronounced for the target-absent trials than for the target-present trials, it too was significant in both cases, $F(1, 7) = 12.07$, $p < .05$, and $F(1, 7) = 10.09$, $p < .05$, respectively. Third, when the slopes for the U-only and D-only conditions were separately analyzed, the gestalt manipulation had no effect on Type U performance, $F(1, 7) = 1.21$, *ns*, but it had a significant effect on Type D performance, $F(1, 7) = 8.64$, $p < .05$, and this effect was more pronounced for the target-absent than for the target-present trials, $F(1, 7) = 6.76$, $p < .05$, but was significant in both cases.

Clearly, these results accord well with the hierarchical object-based model: Binding the features of Type D stimuli was facilitated when the feature-carrying parts were more strongly grouped to form a unitary perceptual object despite the fact that the spatial proximity of the parts was held constant. Moreover, it seems unlikely that these effects derived primarily from differences in the amount of interference from neighboring items (i.e., interitem grouping) or

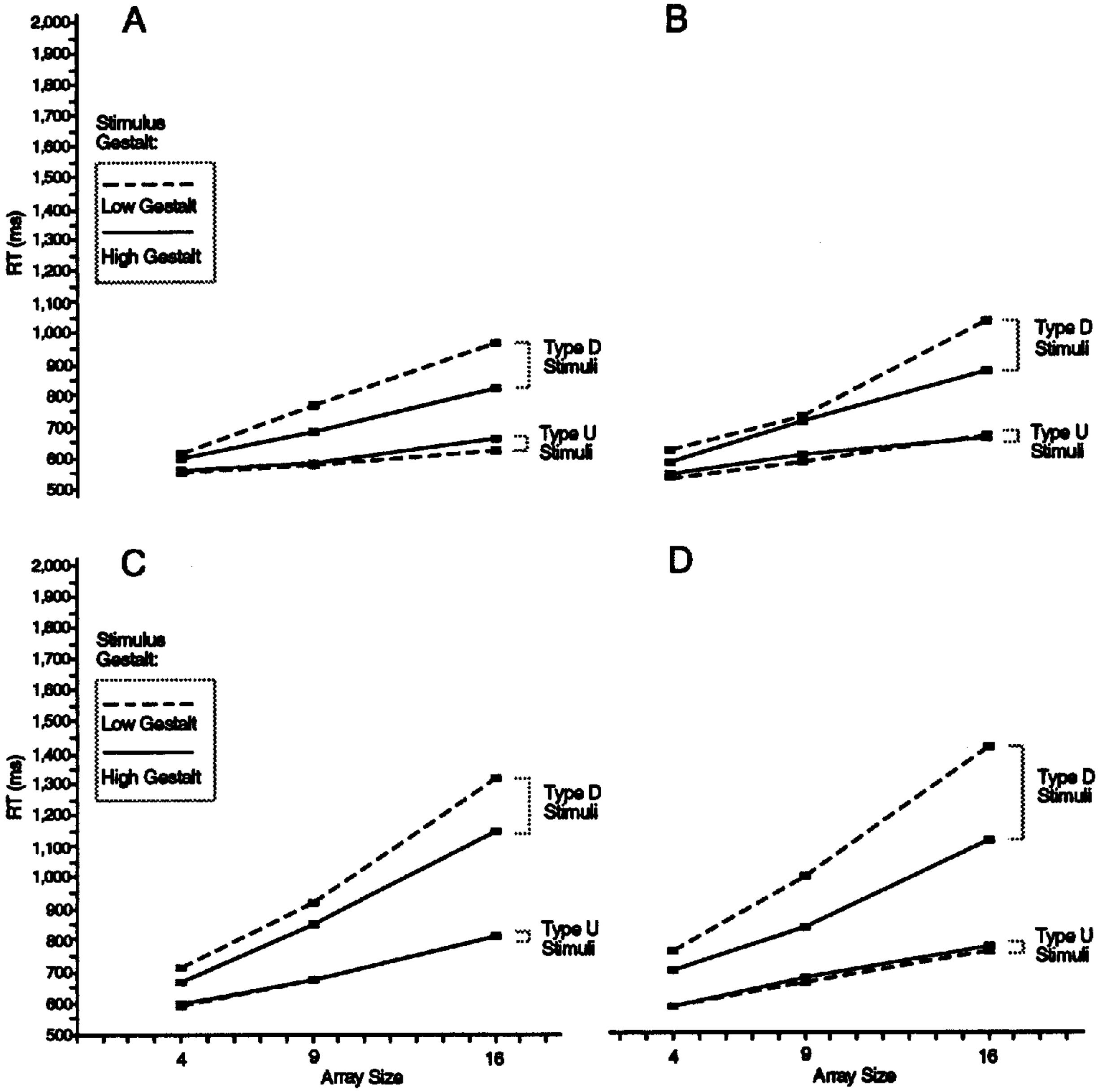


Figure 9. Target-present results from Experiment 3. Mean response time (RT) for target-present trials as a function of array size, matrix color scheme, and stimulus gestalt, plotted separately for each Stimulus Discriminability × Display Density condition: high-discriminability wide-spread display (A), high-discriminability medium-spread display (B), low-discriminability wide-spread display (C), and low-discriminability medium-spread display (D).

from the relative difficulty of preattentively parsing the display (Wolfe, 1994b). Not only was there no Gestalt × Color Scheme × Density interaction ($F < 1$), but the Gestalt × Color Scheme interaction was significant both for the medium-spread displays, $F(1, 7) = 15.15$, $p < .01$, and for the wide-spread displays, $F(1, 7) = 7.26$, $p < .05$., analyzed separately. Instead, the results suggest that the feature-binding process is sensitive to the internal cohesiveness of the global perceptual objects (i.e., intraitem grouping; cf. Rensink & Enns, 1995).

More generally, the results reinforce the conclusion that the color-scheme effects examined in this article do indeed reflect object-based constraints on the feature-binding process. Although there may have been a potentially viable space-based account of the color-scheme effects in the previous experiments (see discussion in Experiment 1), object-based grouping of the component parts can be seen to exert an effect above and beyond any putative effect of spatial separation per se.

Finally, it is worth noting the finding of substantial

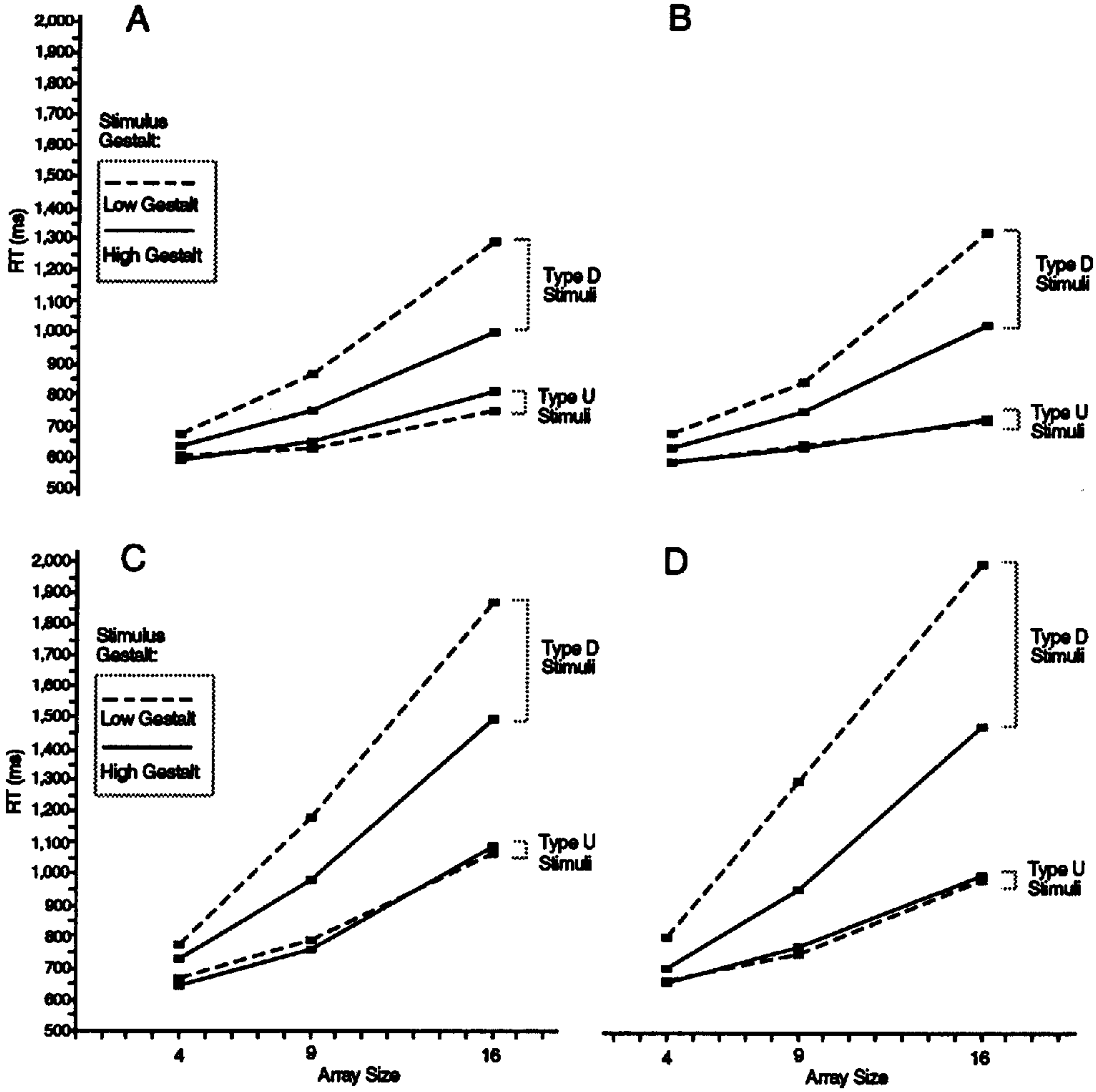


Figure 10. Target-absent results from Experiment 3. Mean response time (RT) for target-absent trials as a function of array size, matrix color scheme, and stimulus gestalt, plotted separately for each Stimulus Discriminability × Display Density condition: high-discriminability wide-spread display (A), high-discriminability medium-spread display (B), low-discriminability wide-spread display (C), and low-discriminability medium-spread display (D).

color-scheme effects for all combinations of the discriminability, density, and gestalt conditions, in particular, even in the high-discriminability condition, in which the U-only slopes were quite flat. This result has two important implications: First, it shows that the magnitude of obtained color-scheme effects does not depend simply on the absolute level of U-only performance (an issue raised in Experiment 2). Second, it shows that good grouping between individual component parts does not necessarily eliminate color-scheme effects, even under conditions allowing parallel search for Type U stimuli. Instead, this may depend on the availability of an effective global object representation. Note that despite the relative unitization of the items in the high-gestalt condition, the closed parentheses may actually tend to conceal the letter-shape of each item. That is, the dominant shape of the global item seems to be round rather than S- or V-shaped. Of course, the failure to obtain parallel search with the Type D stimuli in this experiment could simply stem from the fact that the stimulus components were not spatially overlapping (but see Wolfe et al., 1990,

Table 3
Mean Slopes, Intercepts, r^2, and Error Rates of the Search Functions for Each Condition in Experiment 3

Condition	High gestalt Slope (ms)	Int.	r^2	% error	Low gestalt Slope (ms)	Int.	r^2	% error
High-discriminability wide-spread display								
Target present								
U-only	8.7	515	.79	3.6	6.0	526	.71	4.2
D-only	18.7	520	.91	6.0	29.3	498	.96	6.5
Target absent								
U-only	18.8	501	.91	1.6	12.6	539	.85	1.6
D-only	31.2	495	.95	1.3	52.5	438	.87	1.3
High-discriminability medium-spread display								
Target present								
U-only	9.8	516	.88	6.0	11.7	485	.90	3.1
D-only	23.9	499	.95	10.7	34.7	466	.93	9.9
Target absent								
U-only	12.2	527	.88	2.9	10.7	539	.81	2.3
D-only	32.8	480	.96	2.6	55.0	412	.96	3.9
Low-discriminability wide-spread display								
Target present								
U-only	17.8	522	.94	6.3	18.5	511	.90	5.7
D-only	40.7	494	.97	12.0	50.6	495	.96	9.9
Target absent								
U-only	37.5	470	.96	0.8	33.2	519	.95	1.3
D-only	64.3	445	.97	3.1	92.2	384	.97	1.6
Low-discriminability medium-spread display								
Target present								
U-only	16.3	518	.92	4.9	15.0	523	.88	6.0
D-only	35.0	540	.97	10.7	55.0	522	.89	12.0
Target absent								
U-only	28.9	530	.89	1.6	27.1	538	.84	1.6
D-only	65.2	418	.98	4.7	99.7	406	.99	5.8

Note. Int. = intercept.

Experiment 6). These issues are addressed further in the following experiment.

Experiment 4

Experiment 4 is designed to address some additional implications of the hierarchical object-based model. In particular, it examines the idea that search efficiency for complex stimuli depends not only on the grouping strength between the feature-carrying elements of the stimuli, but also on whether the distinguishing features can be processed effectively at some unitary level of object structure. Like Experiment 3, this experiment employs a gestalt manipulation, but here it is the *type* of gestalt rather than the strength of the gestalt that is varied. In fact, in this experiment there is a very strong grouping between the task-relevant and task-neutral stimulus components in both gestalt conditions. However, in the *global gestalt* condition, the gestalt yields a distinguishing holistic feature at the global-item level (see, e.g., Garner, 1978; Kimchi & Goldsmith, 1992; Navon, 1983; Pomerantz & Pristach, 1989; Rensink & Enns, 1995), whereas in the *local gestalt* condition the effective search feature is available only at the local component level.

Figure 11 depicts the stimuli used in this experiment. Regardless of the gestalt condition, the target conjunction is

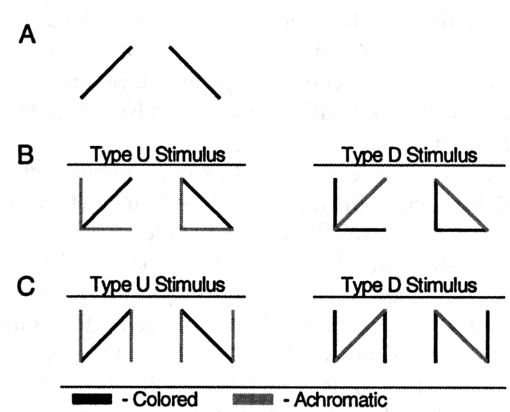


Figure 11. Stimuli used in Experiment 4, demonstrating the no-context (A), global gestalt (B), and local gestalt (C) conditions.

defined in terms of line orientation and color (e.g., 45° and green). However, although the target orientation is nominally 45° versus 135° (right oblique vs. left oblique; see Figure 11A), the task-neutral elements added to these stimuli are expected to configure with the task-relevant components, changing the effective search features between the global gestalt and local gestalt conditions: For the global gestalt stimuli (Figure 11B), a neutral L component has been added to each item, so that the target and distractor orientations can now be distinguished in terms of global arrow and triangle shapes, respectively. These global shapes and their configural features (e.g., junction vs. closure) should be more effective in distinguishing the targets and distractors than the component (oblique) line orientation alone. Indeed, Pomerantz, Sager, and Stover (1977) have previously demonstrated a "configural superiority" effect for such stimuli: Response times in a discrimination task using arrow and triangle stimuli were more than twice as fast as when oblique lines were presented alone.

For the local gestalt stimuli (Figure 11C), the added component is a pair of flanking vertical lines, which creates a reflected (mirror image) N for stimuli with the target line orientation, and a normal N for stimuli with the distractor orientation. Here, the added context should tend to yield a configural *inferiority* effect (Pomerantz, 1981; Pomerantz et al., 1977): Distinguishing the two types of N shapes should be more difficult than distinguishing the individual line orientations alone, because in this case there is no apparent holistic feature that could compensate for the structural masking of the diagonal lines (cf. Rensink & Enns, 1995, who found that searching for line segments of differing lengths became substantially more difficult when they were embedded in more global Müller-Lyer configurations of equal overall length).

Using these stimuli, stimulus color scheme was manipulated as in the previous experiments: For Type U stimuli, the diagonal-line component was colored and the additional component (L or ||) was achromatic, whereas for Type D stimuli, the additional component was colored and the diagonal-line component was achromatic. Stimulus discriminability (high vs. low) was manipulated as in Experiments 2 and 3, but only a single display density (spread) was used. In addition, in order to obtain data regarding configural superiority and inhibition effects, a no-context baseline task using the diagonal-line stimuli alone was also included (see *Method* section).

What are the predictions for this experiment? First, it is expected that across the matrix color-scheme conditions, the search should be more efficient for the global gestalt than for the local gestalt stimuli: The holistic features of the arrow and triangle shapes should allow for relatively easy search (cf. Treisman & Paterson, 1984), whereas discriminating reflected from normal Ns should be more difficult. Second, and more important, it is predicted that there will be little or no color-scheme effect for the global gestalt stimuli, but that a substantial color-scheme effect will be obtained for the local gestalt stimuli. For the global gestalt stimuli, the processing should be based on the global item representation, in which case only a single object marker–file needs to be accessed for both color and shape features regardless of the stimulus coloring scheme. In contrast, because the global representation for the local gestalt stimuli presumably does not offer a discriminating holistic feature, processing needs to be based on the local component level, in which case Type U stimuli may be tested by assessing a single object marker (i.e., the diagonal line), whereas the processing of Type D stimuli requires two or more object markers (i.e., the individual lines) to be separately accessed. These gestalt effects are expected to hold regardless of the stimulus (color) discriminability.

Method

Participants. Eight 1st-year psychology students at the University of Haifa, Haifa, Israel, participated for course credit. All had normal or corrected-to-normal acuity and normal color vision.

Stimuli. Each stimulus was composed of a diagonal line, oriented relative to upright at either 45° (/ [right oblique]) or 135° (\ [left oblique]), and a pair of task-neutral context lines (see Figure 11). In the global gestalt condition, the pair of context lines formed an L shape. The L was positioned below and to the left of the diagonal line in such a way that a left-oblique diagonal line would create a triangle shape, whereas a right-oblique diagonal line would create an arrow shape. In the local gestalt condition, the context lines were vertical and parallel, flanking the diagonal line. In this case, a right-oblique diagonal line would create an N shape, whereas a left-oblique diagonal line would create a reflected N shape. At a viewing distance of 90 cm., the length of the diagonal line subtended a visual angle of 0.97°, and the length of each of the context lines subtended a visual angle of 0.68°.

Each stimulus was colored according to one of two stimulus color schemes, Type U (colored diagonal line, achromatic context lines) or Type D (achromatic diagonal line, colored context lines). Two matrix color schemes were used, U-only and D-only. Stimulus discriminability (high vs. low) was manipulated as in Experiments 2 and 3, using the same two sets of stimulus colors: In the high-discriminability condition, the target was green and / among green \ and pink / distractors, whereas in the low-discriminability condition the target was orange and / among orange (and red / distractors.

Arrays of either 4, 9, or 16 stimuli were presented at a fixed display density. Arrays of 16 stimuli were presented in a slightly irregular 4 × 4 matrix (randomly jiggled by up to 0.05° of visual angle), covering a 7.0° × 7.0° field. The average distance between the centers of adjacent stimuli was 2.2°. Arrays of 9 and 4 stimuli were displayed on randomly chosen 3 × 3 and 2 × 2 subsets of these same stimulus positions.

Procedure. Each participant took part in two separate sessions conducted on consecutive days. Stimulus discriminability was varied between sessions. Gestalt type and matrix color scheme were varied within each session in separate blocks, with gestalt type as the outer blocking factor. Session order and block order were counterbalanced across participants. The instructions, the chronology of each trial, and the manner of responding were essentially the same as in the previous experiments.

In addition, an initial baseline task for each discriminability condition was also included at the beginning of each session, in which participants performed a standard conjunctive-search task using the oblique-line stimuli alone (i.e., without any contextual lines). In this task, the target was defined simply as a green right-oblique line or an orange right-oblique line, as appropriate.

In each session, participants performed three consecutive blocks in each matrix color-scheme condition (two levels) for each gestalt type (two levels), plus an additional three blocks in the initial baseline task. Within each block, 14 repetitions of target presence (two levels) and array size (three levels) were randomly intermixed, for a total of 84 trials per block. The first block of each new color-scheme condition (and of the baseline task) was discarded as practice. Participants were allowed to rest between blocks and were required to take a 10 min break at the midpoint of each session (between gestalt conditions). The sessions lasted about 90 min.

Results and Discussion

For each participant, mean RTs (errorless trials only) and error rates were computed for each cell of the design, after trimming outliers (RT < 200 ms or RT > 4,000 ms; fewer than 0.1% of the total number of trials). Figure 12 presents the average RT × Array Size functions for each combination

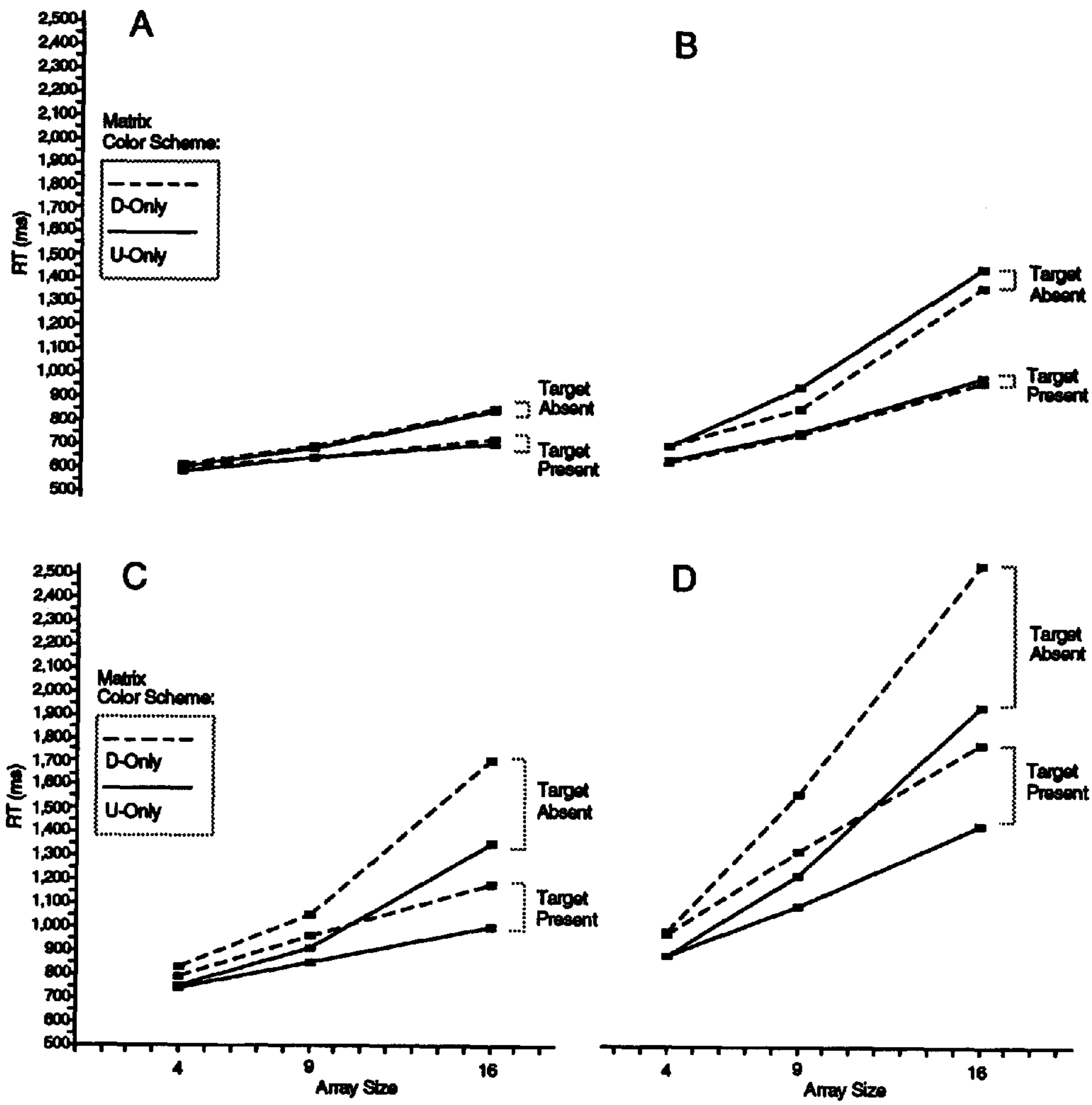


Figure 12. Results from Experiment 4. Mean response time (RT) as a function of array size, matrix color scheme, and target presence, plotted separately for each Gestalt Type × Stimulus Discriminability condition: global gestalt high discriminability (A), global gestal low discriminability (B), local gestalt high discriminability (C), and local gestalt low discriminability (D).

of gestalt type, stimulus discriminability, target presence, and matrix color scheme. The average slope, *y*-intercept, and r^2 of the best fitting linear search functions, together with the error rates, are presented in Table 4.

A visual inspection of the results discloses two very different patterns of effects for the two gestalt conditions. A four-way ANOVA, Gestalt × Discriminability × Target Presence × Color Scheme, was conducted on the search slopes to confirm this impression: First, the main effect for gestalt was significant, $F(1, 7) = 53.53$, $p < .0005$, as was the Gestalt × Target Presence interaction, $F(1, 7) = 27.84$, $p < .005$. Substantially better search efficiency was found for the global gestalt stimuli than for the local gestalt stimuli, and this advantage was more pronounced for the target-absent trials than for the target-present trials, but was significant in both cases. Second, and more important, the Gestalt × Color Scheme interaction was significant, $F(1, 7) = 18.31$, $p < .005$, as was the Gestalt × Color Scheme × Discriminability interaction, $F(1, 7) = 6.04$, $p < .05$, and the Gestalt × Color Scheme × Target Presence interaction, $F(1, 7) = 17.23$, $p < .005$.

Analyses of simple effects clarified these interactions: On the one hand, with the global gestalt stimuli, no color-scheme effect was obtained ($F < 1$), and this was the case regardless of stimulus discriminability and target presence. On the other hand, the local gestalt stimuli yielded a substantial color-scheme effect, $F(1, 7) = 27.44$, $p < .005$, modulated by both stimulus discriminability and target presence: It was more pronounced in the low-discriminability condition than in the high-discriminability condition, $F(1, 7) = 6.81$, $p < .05$, and for the target-absent trials than for the target-present trials, $F(1, 7) = 24.01$, $p < .005$, but was significant in all cases.

Figure 13 presents supplemental results that can shed some light on the effects of the added contextual lines for the two types of gestalt stimuli. Here, the U-only search functions for each gestalt (context) type are compared to the functions that were obtained for the no-context oblique-line stimuli in the initial baseline task. These baseline comparisons can yield only a rough indication of the context effects, because the no-context conditions always preceded the context conditions and, hence, differential practice and transfer might also be contributing. Nevertheless, the general pattern is clear: The global gestalt stimuli exhibit a context-facilitation effect (except for the low-discriminability target-absent trials), whereas the low-gestalt stimuli yield a context-inhibition effect (except for the high-discriminability target-present trials).

These results implicate a crucial role for hierarchical object structure in visual search: It appears that the feature-binding process is sensitive not only to the grouping strength of the component parts of complex stimulus items (Experiment 3), but also to the quality of the task-relevant feature representations at different levels of the object hierarchy. The component parts of the stimulus items were strongly grouped into global configurations in both the global gestalt and the local gestalt conditions. Yet stimulus color scheme had no effect on search efficiency in the former case and a substantial effect in the latter case. Of course, one cannot be certain that grouping strength did not also vary along with the intended manipulation of the configural features (see Pomerantz & Pristach, 1989, regarding the difficulty of distinguishing these two factors empirically). However, considering the qualitative difference between the results of Experiments 3 and 4, it seems reasonable to conclude that the critical factor in this experiment was not grouping strength per se, but rather, the extent to which the task-relevant features could be accessed effectively at a unitary level of object structure. These results, then, join other recent findings (e.g., Bilsky & Wolfe, 1995; Donelly et al., 1991; Enns & Rensink, 1990a, 1990b, 1991; Gilchrist, Humphreys, Riddoch, & Neumann, 1997; He & Nakayama, 1992; Kimchi, in press; Rensink & Enns, 1995; Wolfe, Friedman-Hill, & Bilsky, 1994) in suggesting a rather complex role for object structure in visual search (see General Discussion).

The results also shed light on two other issues that were raised earlier: First, they reinforce the idea that color-scheme effects do not derive from the mere spatial separation of the feature-carrying elements: The stimulus components of the global gestalt stimuli were not spatially overlapping, yet no color-scheme effects were observed (see also Wolfe et al., 1990, Experiment 6). Second, the results cast further doubt on the idea that the preattentive guided-search mechanism is special in the sense of being space based: In this experiment, no color-scheme effects were observed for the global gestalt stimuli in both the high- and the low-discriminability conditions, even though the slopes in the latter condition were quite steep. Thus, immunity to color-scheme effects appears to depend on whether or not the stimuli can be processed most effectively at a unitary level of object structure, regardless of whether that processing is slow and serial or fast and parallel.

Table 4
Mean Slopes, Intercepts, r^2, and Error Rates of the Search Functions for Each Condition in Experiment 4

Condition	Global gestalt: Slope (ms)	Int.	r^2	% error	Local gestalt: Slope (ms)	Int.	r^2	% error
High discriminability								
Target present								
U-only	10.3	536	.96	3.7	21.7	651	.87	7.6
D-only	9.8	554	.91	5.2	32.6	659	.99	9.5
Target absent								
U-only	20.2	511	.96	1.6	51.2	505	.94	4.2
D-only	19.8	526	.96	2.7	74.0	479	.96	3.9
Low discriminability								
Target present								
U-only	30.4	584	.96	7.7	44.6	674	.98	9.4
D-only	29.0	580	.94	6.4	63.5	699	.97	12.2
Target absent								
U-only	63.7	495	.97	1.8	84.7	489	.98	4.2
D-only	57.0	506	.97	1.8	123.8	442	.98	2.7

Note. Int. = intercept.

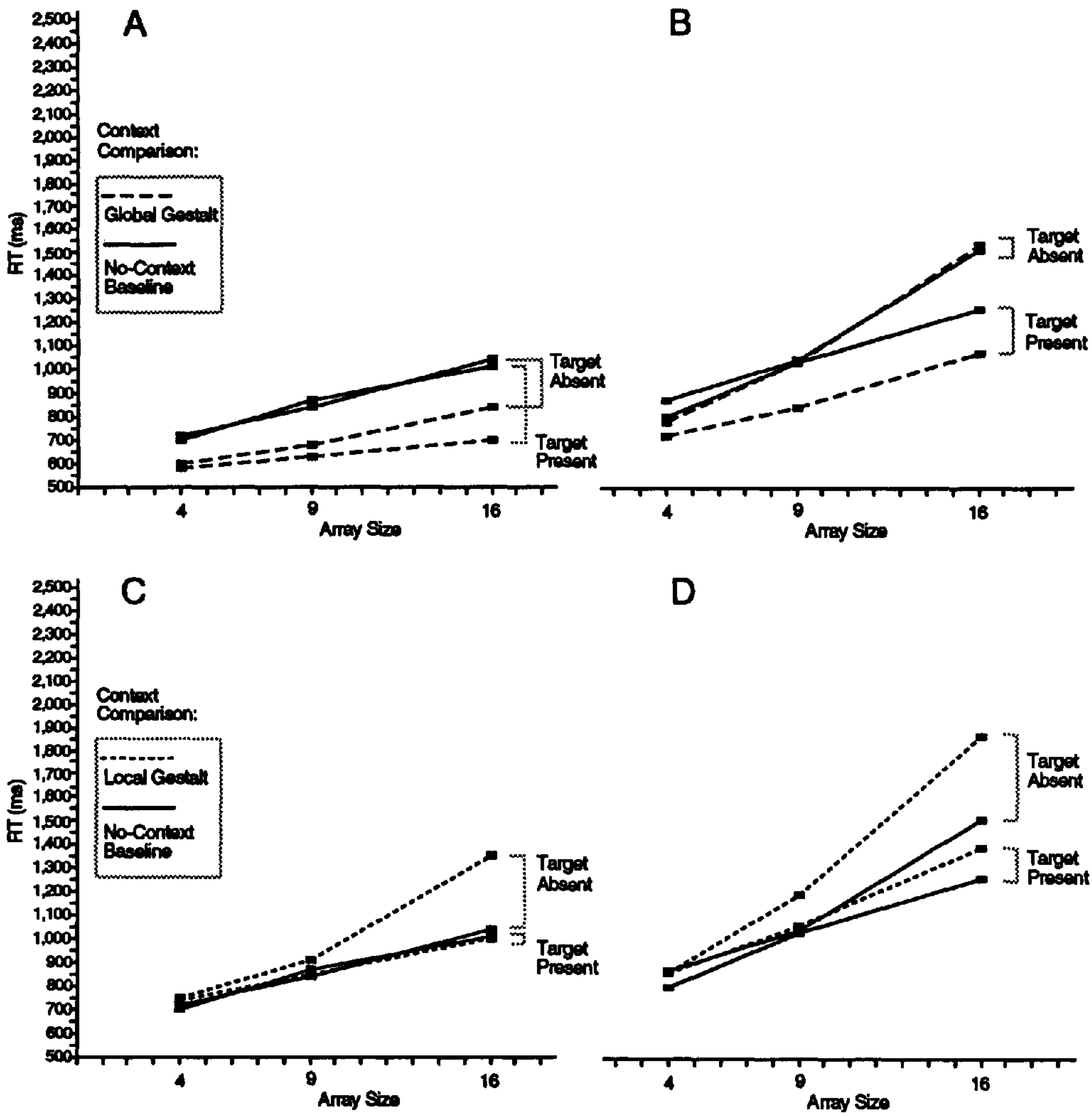


Figure 13. Context-baseline effects in Experiment 4. Mean response time (RT) as a function of array size and target presence for the U-only conditions versus the no-context baseline conditions, plotted separately for each Gestalt Type × Stimulus Discriminability condition: global gestalt high discriminability (A), global gestalt low discriminability (B), local gestalt high discriminability (C), and local gestalt low discriminability (D). The no-context baseline functions are plotted twice.

General Discussion

This study attempted to clarify some basic assumptions made by feature-integration theory concerning the representation and processing of stimulus features in visual search. The overall strategy has been to remain as faithful as possible to the general FIT framework, both in the specification of theoretical alternatives and in the interpretation of experimental results. This approach was adopted because of FIT's dominance of research on visual search over the past 15 years, making its status quo view a good point of departure (but see Duncan & Humphreys, 1989, 1992; Grossberg, Mingolla, & Ross, 1994; Humphreys & Müller, 1993; Logan, 1996; Navon, 1990). By accepting the basic tenets of FIT, my work can point to those assumptions which, at a minimum, seem to require modification or elaboration in light of the findings. Of course, it is important to explore some more far-reaching theoretical alternatives as well.

In this discussion, I first review the results and consider their implications for the basic FIT framework. I then examine the broader implications of the findings with regard to some more general issues, including the putative role of attention in feature integration, the issue of early versus late attentional selection, and the question of how the object-based and space-based views of attention might be reconciled.

Object-Based Feature Integration: Some Proposed Modifications

Summary of the Results

Experiment 1 showed that searching for a conjunction of features was more difficult when the two features pertained to two different forms at the same stimulus location (Type D stimuli) than when they pertained to the same form (Type U stimuli). This effect was found when the matrix color-scheme conditions were presented in separate blocks (Experiment 1 here), when they were manipulated within blocks (Goldsmith, 1995, Experiment 1), and when performance for the two *target* color schemes was compared in the mixed color-scheme condition (both in Experiment 1 here and in Goldsmith, 1995, Experiment 1). These results support the object-based model: First, to account for those effects, it must be presumed that specific feature–object linkages, rather than just feature–location linkages, are embodied in the master-map representation. Second, the results imply that the feature-binding process is constrained by the object-based stimulus structure (cf. Duncan, 1984; Kahneman & Henik, 1981; Prinzmetal, 1981), because performance was affected even though the specific feature–object linkages were irrelevant to the task.

The remaining experiments attempted both to strengthen and to refine these results. In Experiment 2, the magnitude of the color-scheme effect was found to depend on the conditions of the search: Whereas substantial color-scheme effects were evidenced under most of the search conditions, no such effects were observed under conditions that allowed near-parallel search (target-present trials for spread, highly discriminable stimuli). Although such a result could be taken to imply that the preattentive, parallel-guidance component of visual search is space based, an alternative account was put forward in terms of hierarchical object-based structure: The color-scheme effects may have been attenuated because the highly discriminable and widely spaced stimuli could be processed efficiently at a global level of object structure.

Experiments 3 and 4 provided converging evidence for the idea that color-scheme effects depend on object-based rather than spatial parameters. In Experiment 3, the cost of processing Type D stimuli was shown to depend on the grouping strength between the feature-carrying parts, even though the spatial separation between them was held constant: Performance improved when the parts were more strongly grouped to form a more cohesive object. Nevertheless, unlike in Experiment 2, the color-scheme effects were observed in Experiment 3 under conditions allowing near-parallel search for Type U stimuli, apparently because the stimuli were not highly discriminable at the global object level. Experiment 4 manipulated the level of object structure at which the search could be most effectively conducted. When the stimuli could be discriminated most effectively at a local level of object structure, substantial color-scheme effects were observed. However, when the stimuli could be discriminated most effectively at the global level of object structure, no color-scheme effects were observed—both when the search was fast and parallel (high-discriminability condition) and when it was slow and serial (low-discriminability condition).

Although, as discussed earlier, some of these findings might perhaps be explained by postulating some combination of the object-based and space-based models, they are generally embarrassing to the space-based account. Indeed, the most parsimonious account would seem to be provided by the hierarchical object-based model. Of course, the results do not point to any specific instantiation of this model, but rather, provide constraints for a general class of models that must account for the role of object structure in visual search. What modifications are needed for FIT to accommodate the current results? Several aspects need to be incorporated, including (a) specific feature–object linkages and object-based feature binding and (b) intraobject grouping strengths and hierarchical object structure.

Master Map of Perceptual Objects

Incorporating specific feature–object linkages, rather than feature–location linkages, into the underlying master-map representation is a subtle but important change. Instead of a master map of locations, in which objectness (filled) is an attribute of the location, what is required is a master map of perceptual objects, in which location is an attribute of the object (see also Wolfe, 1996). As discussed earlier, FIT seems to have approached such a view with regard to the top-down feature-inhibition component (Treisman & Sato, 1990) but not for the attentional-scan component, which is still assumed to select locations "independently of the objects they contain" (Treisman, 1993, p. 171). Admittedly, this modification may be of little consequence for the many visual-search experiments that employ simple, discretely spaced stimulus items. Wolfe (1994a), for instance, notes that

> there has been considerable discussion about whether attention is directed to locations or to objects. . . . [B]ut for the current version of Guided Search, the issue is not critical. In an array of items, activation will be higher at the items than in intervening blank space. Objects and locations will be, in effect, the same thing. (p. 238)

Nonetheless, an object-based representation would seem to be imperative for any model that purports to explain how visual attention operates in more ecologically realistic scenes containing structurally complex and spatially overlapping objects (e.g., Wickens & Long, 1995; Wolfe, 1994b, 1996). Moreover, adopting the notion of object-based feature binding may also necessitate a corresponding shift in stance with regard to the role of attention in the perception of objects (see later discussion).

Intraobject Grouping and Hierarchical Object Structure

FIT assumes that grouping and hierarchical structure can be captured generally in terms of interitem grouping between spatially contiguous items and in terms of a master-map representation at multiple spatial scales (see, e.g., Heathcote & Mewhort, 1993; Logan, 1996; Nakayama, 1990; Treisman, 1982, 1990, 1992a, 1992b). Treisman (1982), for instance, found conjunctive search to be serial across groups, rather than across items, whenever the display contained homogeneous groups of distractors. She "predicted and explained 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" (Treisman, 1982, p. 211; see also Treisman & Sato, 1990). Presumably, such a finding can be accounted for without explicitly representing perceptual objects or groups in the master-map medium (Treisman, 1990, 1992b). A comparable assumption is made in the guided search model, in which "there is an implicit grouping. . . . Similar items will have similar top-down activations. Spatially contiguous groups of items will produce lower bottom-up activations" (Wolfe, 1994a, p. 213).

The combined results of the experiments here, however, are difficult to accommodate in purely spatial terms. First, the color-scheme effects with overlapping stimuli (Experiments 1 and 2) indicate that as far as the feature-binding process is concerned, different line segments at the same stimulus location can group into separate perceptual objects or object parts. Second, both the strength and the type of grouping between the parts of individual items can have substantial effects on search performance. Experiment 3 demonstrated the effects of *intra*item grouping—grouping that derived from the configural goodness of the gestalt formed by the component elements (based on good continuation and containment), rather than from spatial proximity per se (see also Rensink & Enns, 1995). Like the color-scheme effects themselves, such effects cannot be mediated by the mere influence of proximity or feature homogeneity on the spatial bounds of an attentional spotlight (or window). Rather, to account for such effects, FIT would need to postulate a master-map representation having more or less cohesive units of selection—a notion that clearly belongs to an object-based representational scheme. Similarly, Experiment 4 demonstrates that when object parts group, they group to form a more global object that has its own set of global features. Although, this idea has been addressed by FIT's assumption that certain emergent features are directly registered and linked to global stimulus locations (Treisman, 1992a, 1992b; Treisman & Paterson, 1984), the units of selection at the various levels of globality would seem to be captured better in terms of global objects and their parts than in terms of more or less diffuse locations.

These results, then, suggest that FIT's assumptions are inadequate because they attempt to account for objectness and object grouping implicitly—by means of the spatial bounds of the attentional spotlight—rather than by explicitly representing these aspects in the preattentive master map (for some related arguments, see Duncan, 1995; Wolfe, 1996). Of course, explicit incorporation poses its own problems, such as the need to specify exactly how perceptual objects, with their hierarchical structure and grouping strengths, are in fact derived and represented. Despite a vast amount of work on perceptual organization and stimulus-object structure (for reviews, see, e.g., Banks & Krajicek, 1991; Kimchi, 1992; Kimchi & Goldsmith, 1992; Robertson & Lamb, 1991; Treisman, 1986b), there is still no consensual definition of *perceptual object* (Duncan, 1984; Logan, 1996) nor any standard procedure for extracting the object-based structure of a visual display.

How, then, might grouping strengths and hierarchical object structure be explicitly incorporated into the master-map representation? One rather far-reaching proposal comes from Duncan and Humphreys (1989, 1992) in a framework for visual search called *attentional engagement theory* (AET). AET assumes a hierarchically structured preattentive representation in which objects and their parts are represented as *structural units* at different levels of scale, with each structural unit linked to every other unit by a grouping strength. During the search process, there is neither an attentional spotlight nor feature integration but rather the structural units compete for access to visual short-term memory (cf. Bundesen, 1990; Grossberg et al., 1994; Logan, 1996), with priority of entry determined by a weight assigned to each unit. This weight increases with increasing similarity between the unit's features and the target definition or template. In addition, the weight assigned to one unit also affects the weights of other units in proportion to the strength of perceptual grouping between them (*weight linkage*).

Although many of the specific details of AET as opposed to FIT are unnecessary to account for the results of this study, AET embodies several aspects of the type of hierarchical object-based architecture implied by my findings (and see Humphreys & Müller, 1993): (a) The underlying preattentive representation is a structured hierarchy of perceptual objects; (b) selection is indefinitely flexible, using object descriptions at any level or levels in the hierarchy in accordance with the demands of the task; (c) despite this flexibility, the selection is nevertheless constrained by the initial, preattentively determined object structure; and (d) objectness (grouping) is a graded attribute, explicitly represented in the weight linkages. Interestingly, some of these same hierarchical object-based aspects have also emerged in work conducted within the guided-search framework (Bilsky & Wolfe, 1995; Wolfe et al., 1994), suggesting the need for a "preattentive item map" in that model (Wolfe, 1996).

An Early-Object-Based Late-Selection View of Feature Binding

The supposition of a relatively sophisticated and structured preattentive object-based representation raises an important issue regarding the nature of the attentional processing: What role does attention play in an object-based account of feature integration? Can object-based feature binding be incorporated into FIT without compromising its central assumption concerning the "integrative" function of attention?

To answer this question, let us reexamine the presumed role of attention in the FIT framework from the perspective of another central issue in the attention literature—the issue of early versus late attentional selection. FIT constitutes an *early selection* (Broadbent, 1982; Treisman, 1969) theory of feature integration, in which the preattentive processing stage is limited to analyzing simple features of the two-dimensional image (e.g., color, orientation, length, curvature) and registering these features into a set of separate spatiotopic maps (see Figure 1). These simple features are generally assumed to be derived at the earliest stages of processing in the visual cortex (Cavanagh, Arguin, & Treisman, 1990; Treisman, Cavanagh, Fischer, Ramachandran, & von der Heydt, 1990). The computation of all higher level object properties (e.g., feature conjunctions and spatial relations) must then be carried out at the focused attentional stage of processing. Hence, the original claim was that attention was required to bind together the separable features of objects, which would otherwise be "free-floating" (Treisman & Gelade, 1980). This is apparently one source of the term *feature integration* (Navon, 1990).

In subsequent formulations (e.g., Treisman, 1986a, 1988; Treisman & Gormican, 1988; Treisman & Souther, 1985), however, FIT added preattentive links that tie each stimulus feature to a particular location in the master-map representation (see Figure 1). In such a scheme, attention might be likened to a peephole that secures a circumscribed workspace that can be used to test the features linked to each stimulus location (Navon, 1990) and, if necessary, to analyze their structural relations (Treisman, 1990). Although this may seem to be a more modest role, attention still fulfills an essential function by bringing the preattentively linked features together into a temporary workspace (object file) that allows higher level perceptual and cognitive processing.

Such a role is broadly consistent with the object-based model proposed here. This model, however, implies a *late-selection* view (Deutsch & Deutsch, 1963; Duncan, 1980), in which "the perceptual units that engage visual attention during search are emergent units that are the outcome of considerable preattentive visual processing" (Grossberg, et al., 1994, p. 470; see also Duncan & Humphreys, 1989, 1992; Heathcote & Mewhort, 1993). In this model, rather than synthesizing features into conjoined object descriptions, attention might be thought of as *navigating* through a complex and highly structured preattentive representation, specifying selected portions for higher-level processing (see also Duncan & Humphreys, 1989, 1992; Kramer & Watson, 1996).

This point, however, requires some clarification. The term *preattentive*, as it is used in this article and throughout the visual-search literature, refers to processing that occurs before attention is focused on a particular stimulus, that is, when attention is (still) widely spread across the entire display or visual field (Treisman & Gormican, 1988; Wolfe, 1994a). Such preattentive processing should be distinguished from that which occurs *without* attention, for instance, when attention is focused on some other stimulus (Rock, Linnett, Grant, & Mack, 1992). Indeed, although in the latter case there appear to be some severe limitations on the type of information that is consciously available about the unattended stimuli (e.g., Ben-Av, Sagi, & Braun, 1992; Mack, Tang, Tuma, Kahn, & Rock, 1992; Rock & Gutman, 1981; Rock et al., 1992; but see DeSchepper & Treisman, 1996; Moore & Egeth, 1997), recent findings from the visual-search literature indicate that the products of preattentive (or widely spread attentive) processing may be rather refined. For example, preattentive parallel search has been observed for targets defined by relatively high-level properties of the stimuli, such as depth from shading, surface representation, line relations, and three-dimensional structure (e.g., Duncan & Humphreys, 1989; Enns, 1990; Enns & Rensink, 1990a, 1990b, 1991; He & Nakayama, 1992; Heathcote & Mewhort, 1993; Kleffner & Ramachandran, 1992). Such findings are in addition to those indicating that preattentive search processes are sensitive to perceptual grouping and hierarchical object structure (e.g., Bilsky & Wolfe, 1995; Bravo & Blake, 1990; Donelly et al., 1991; Gilchrist et al., 1997; Humphreys et al., 1989; Rensink & Enns, 1995; Wolfe et al., 1994). Thus, "although there are severe limitations on parallel processing in the visual system, there is clear evidence for some quite sophisticated processing beyond the mere extraction of basic visual features" (Wolfe et al., 1994. p. 549).

Of course, the very issue of early versus late selection in visual search is predicated on the assumption that the focused-attentional processing stage does indeed use a preattentively derived stimulus representation (Neisser, 1967). Although this assumption is often treated as axiomatic, one could in fact dispute the use of a preattentive representation during the attentional processing stage altogether (e.g., Navon, 1990; Navon & Kastan, 1997). A less radical alternative, however, would be to assume mutual interactions between the "preattentive" representation and the attentional processing, such that it is continually refreshed and reorganized during the course of that processing (Grossberg et al., 1994; Logan, 1996). Thus, assuming that attention can influence both feature extraction (e.g., Posner, Snyder, & Davidson, 1980; Prinzmetal, Presti, & Posner, 1986) and perceptual organization (e.g., Mack et al., 1992; Peterson & Gibson, 1991; Prinzmetal & Keysar, 1989; Tsal & Kolbet, 1985; Yantis, 1992), a further role of attention in an object-based model might be to recursively regulate the quality and organization of the representational medium in which it operates. Note that on this view, "attention may influence perceptual organization, but it is perceptual organization that determines feature integration" (Prinzmetal et al., 1986, p. 368).

In sum, like the early-selection, space-based (or late object-based) view of feature integration advanced in FIT, the late-selection, early object-based model proposed in this article also implies that "a process of attentional integration is good for something" (Treisman, 1990, p. 461), whether that function is to secure access for higher level processing rather than to integrate, and whether attention is viewed as interacting with, rather than entirely dependent on, a preattentively derived object-based representation. By assuming a more highly processed and structured preattentive representation, however, the early object-based model does leave less (if anything) for attention to do in the way of elaborating the ensuing object-file representation (cf. Treisman, 1990).

Reconciling Object-Based and Space-Based Views of Attention: An Assimilative Approach

In this article, object-based and space-based accounts of feature integration have been cast in opposition. However, many recent discussions imply the need for a rapprochement between the object-based and space-based views generally (see, e.g., Egly et al., 1994; Humphreys et al., 1996; Logan, 1996). The reason for this is not merely to avoid controversy. Rather, the fact is that considerable evidence exists to support both sides of the argument. Thus, for example, after reviewing the available evidence, Egly et al. (1994) conclude: "The evidence we have considered indicates a visual attention system that allows us to select locations, objects, or both, as relevant sources of information. Any comprehensive theory of visual attention will therefore require both space-based and object-based components" (p. 164).

Perhaps the most straightforward approach to reconciliation assumes that attentional selection may sometimes be object based and sometimes be space based, depending on task demands, participant strategies, types of impairment, or specific parameters of the stimulus display. There are two distinct variants of this idea, however. Vecera and Farah (1994), for instance, offer evidence suggesting that "qualitatively different attentional mechanisms exist and can be elicited with the same stimulus set" (p. 158). This evidence is predicated on the relatively restrictive use of the term *object based* mentioned earlier, in which attentional selection is based on the spatially invariant representation of an object's shape or structure, the *what* as opposed to the *where* (Farah et al., 1993; Ungerlieder & Mishkin, 1982). According to this view, space-based and object-based attentional effects reflect the operation of relatively independent, neurologically distinct mechanisms, which select among stimuli on the basis of spatial and non-spatial properties, respectively (Humphreys et al., 1996).

A different variant of this idea is often implied by object-based theorists who focus on the *which* as a fundamental unit of attentional selection, as distinguished from both the *what* and the *where* (see Kanwisher & Driver, 1992). In this view, although the things or chunks of information selected by attention are best thought of as perceptual objects (defined by uniform connectedness and other gestalt grouping factors, including common motion), there is also a need to add spatial information to the underlying object-based representation. One may do so explicitly, by including location information as part of the object representation (e.g., Kahneman & Treisman, 1984; Kahneman et al., 1992), or one might posit functionally distinct but interacting representational systems, for example, a space-based system that represents spatial locations and an object-based segmentation and grouping system that represents perceptual objects that link those locations (e.g., Egly et al., 1994; Farah et al., 1993; Grossberg et al., 1994; Humphreys & Riddoch, 1993a, 1993b). Either way, because both object information and spatial information are included in the preattentive representation or representations, attentional processing can conceivably be based on both types of variables. Thus, for instance, in the initial version of the object-based model presented in this article, although the stimulus features are linked to specific object tokens, these tokens can be selected on the basis of their location (cf. Tsal & Lavie, 1988, 1993), and the time needed to move attention from object to object may depend, at least in part, on the spatial distance between them (cf. Shulman, Remington, & McLean, 1979; Tsal, 1983).

A somewhat different approach, however, follows from the notion of objectness as a graded attribute. In this *assimilative* approach, spatial variables are incorporated into (rather than added to) the overall object structure through their critical role in mediating (a) the hierarchical segmentation of the visual field into perceptual objects (e.g., Beck, Prazdny, & Rosenfeld, 1983; Logan, 1996; Marr, 1982), and (b) the gestalt grouping strengths between these objects (e.g., Kramer & Jacobson, 1991; Pomerantz & Schwaitzberg, 1975; Wertheimer, 1923). Thus, for example, the hierarchical object-based model proposed in this article predicts that both overall search efficiency and color-scheme effects should be affected by the spatial proximity of the stimuli (e.g., Experiment 2)—not because of ensuing differences in the spatial extent of the attentional "spotlight," but rather, because of differences in the strength of grouping between the stimulus objects (cf. Experiment 3). Similarly, although by the additive view, the finding that attention can be reoriented to a different perceptual object more efficiently when the two objects are placed closer together (Vecera, 1994) could be taken to imply the joint operation of separate object-based and space-based attentional components (Egly et al., 1994), by the assimilative view this finding might simply reflect the stronger grouping between the two more proximal objects.

In general, then, the assimilative view treats spatial proximity as one of many grouping factors—albeit, a very powerful one—that affects the efficiency of object-based selection (Baylis & Driver, 1992; Logan, 1996; Prinzmetal, 1981). Of course, this view may not be useful in accounting for all types of space-based attentional effects, for instance, those that occur when there are no objects at all in the visual field (e.g., Posner, 1980), or all types of object-based effects, for instance, those tapping spatially invariant shape representations (e.g., Vecera & Farah, 1994). Indeed, visual attention is probably too multifarious and complex a set of phenomena to be captured entirely by any single approach or conception (cf. Allport, 1989; Koriat & Goldsmith, 1996, 1997). Nevertheless, the assimilative object-based view should prove to be of value in channeling research efforts away from the mere demonstration of space-based or object-based effects, to the more ambitious goal of understanding how the various grouping factors, both spatial and nonspatial, interact to determine the effective stimulus structure (e.g., Duncan & Humphreys, 1992; Gilchrist et al., 1997; Kimchi, in press; Kramer & Watson, 1996; Palmer & Rock, 1994; Rensink & Enns, 1995). This would seem to be a major challenge for the assimilative approach, in order to avoid becoming an ad-hoc explanatory framework.

In conclusion, the object-based view advocated in this article calls for a more thorough and comprehensive examination of the object-based stimulus structure underlying feature integration and visual search. Future work should focus on the details of how that structure is determined, represented, and processed.

References

Allport, A. (1989). Visual attention. In M. I. Posner (Ed.), *Foundations of cognitive science* (pp. 631–682). Cambridge, MA: MIT Press.

Ashby, G. F., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review, 103,* 165–192.

Banks, W. P., & Krajicek, D. (1991). Perception. In M. R. Rozenzweig & L. Porter (Eds.), *Annual review of psychology, 1991* (Vol. 42, pp. 305–331). Palo Alto, CA: Annual Reviews.

Banks, W. P., & Prinzmetal, W. (1976). Configurational effects in visual information processing. *Perception and Psychophysics, 19,* 361–367.

Baylis, G. C. (1994). Visual attention and objects: Two-object cost with equal convexity. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 208–212.

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effects of grouping. *Perception and Psychophysics, 51,* 145–162.

Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 451–470.

Baylis, G. C., Driver, J., & McLeod, P. (1992). Movement and proximity constrain miscombinations of color and form. *Perception, 21,* 201–218.

Beck, J., Prazdny, K., & Rosenfeld, A. (1983). A theory of textural segmentation. In J. Beck, B. Hope, & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 1–38). New York: Academic Press.

Ben-Av, M. B., Sagi, D., & Braun, J. (1992). Visual attention and perceptual grouping. *Perception and Psychophysics, 52,* 277–294.

Bilsky, A. B., & Wolfe, J. M. (1995). Part/whole information is useful in Size × Size but not Orientation × Orientation conjunction searches. *Perception and Psychophysics, 57,* 749–761.

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

Briand, K. A., & Klein, R. M. (1987). Is Posner's "beam" the same as Treisman's "glue"? On the relation between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 228–241.

Broadbent, D. E. (1982). Task combination and selective intake of information. *Acta Psychologica, 50,* 253–290.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547.

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

Cavanagh, P., Arguin, M., & Treisman, A. (1990). Effect of surface medium on visual search for orientation and size features. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 479–491.

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. B. (1991). Density effects in conjunction search: Evidence for a coarse location mechanism of feature integration. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 891–901.

DeSchepper, B., & Treisman, A. (1996). Visual memory for novel shapes: Implicit coding without attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 27–47.

Deutsch, J. A., & Deutsch, D. (1963). Attention: Some theoretical considerations. *Psychological Review, 70,* 80–90.

Donelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 561–570.

Driver, J., & Baylis, G. C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 448–456.

Driver, J., & Halligan, P. W. (1991). Can visual neglect operate in object-centered coordinates? An affirmative single-case study. *Cognitive Neuropsychology, 8,* 475–496.

Duncan, J. (1980). The locus of interference in the perception of simultaneous stimuli. *Psychological Review, 87,* 272–300.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General, 113,* 501–517.

Duncan, J. (1995). Target and nontarget grouping in visual search. *Perception and Psychophysics, 57,* 117–120.

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

Duncan, J., & Humphreys, G. W. (1992). Beyond the search surface: Visual search and attentional engagement. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 578–588.

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.

Egly, R., Driver, J., & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General, 123,* 161–177.

Enns, J. T. (1990). Three-dimensional features pop out in visual search. In D. Brogan (Ed.), *Visual search* (pp. 37–45). London: Taylor and Francis.

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

Enns, J. T., & Rensink, R. A. (1990b). Sensitivity to three-dimensional orientation in visual search. *Psychological Science, 1,* 323–326.

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98,* 335–351.

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

Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the focus of attention: A zoom lens model. *Perception and Psychophysics, 40,* 225–240.

Farah, M. J., Wallace, M. A., & Vecera, S. P. (1993). "What" and "where" in visual attention: Evidence from the neglect syndrome. In I. H. Robertson & J. C. Marshall (Eds.), *Unilateral* *neglect: Clinical and experimental studies* (pp. 123–137). Hillsdale, NJ: Erlbaum.

Fox, J. (1978). Continuity, concealment, and visual attention. In G. Underwood (Ed.), *Strategies of information processing* (pp. 23–66). New York: Academic Press.

Garner, W. R. (1978). Aspects of a stimulus: Features, dimensions, and configurations. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, NJ: Erlbaum.

Gilchrist, I. D., Humphreys, G. W., Riddoch, J., & Neumann, H. (1997). Luminance and edge information in grouping: A study using visual search. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 464–480.

Goldsmith, M. (1995). *What's in a location? Comparing object-based and space-based models of feature integration in visual search.* Unpublished doctoral dissertation, University of Haifa, Haifa, Israel.

Grabowecky, M., & Khurana, B. (1990). Features were meant to be integrated [Abstract]. *Investigative Ophthalmology and Visual Science, 31,* 105.

Grossberg, S., Mingolla, E., & Ross, W. D. (1994). A neural theory of attentive visual search: Interactions of boundary, surface, spatial, and object representations. *Psychological Review, 101,* 470–489.

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

Heathcote, A., & Mewhort, D. J. K. (1993). Representation and selection of relative position. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 488–516.

Humphreys, G. W., & Müller, H. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25,* 43–110.

Humphreys, G. W., Olson, A., Romani, C., & Riddoch, M. J. (1996). Competitive mechanisms of selection by space and object: A neuropsychological approach. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 365–393). Washington, DC: American Psychological Association.

Humphreys, G. W., Quinlan, P. T., & Riddoch, M. J. (1989). Grouping processes in visual search: Effects with single- and combined-feature targets. *Journal of Experimental Psychology: General, 118,* 258–279.

Humphreys, G. W., & Riddoch, M. J. (1993a). Interaction between space-based and object-based systems revealed through neuropsychology. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 143–162). Cambridge, MA: MIT Press.

Humphreys, G. W., & Riddoch, M. J. (1993b). Interactive attentional systems and unilateral neglect. In I. H. Robertson & J. C. Marshall (Eds.), *Unilateral neglect: Clinical and experimental studies* (pp. 139–168). Hillsdale, NJ: Erlbaum.

Johnston, J. C., & Pashler, H. (1990). Close binding of identity and location in visual feature perception. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 843–856.

Kahneman, D., & Henik, A. (1981). Perceptual organization and attention. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 181–211). Hillsdale, NJ: Erlbaum.

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

Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24,* 175–219.

Kanwisher, N., & Driver, J. (1992). Objects, attributes, and visual attention: Which, what and where. *Psychological Science, 1,* 26–31.

Kimchi, R. (1992). The primacy of wholistic processing and the global/local paradigm: A critical review. *Psychological Bulletin, 112,* 24–38.

Kimchi, R. (in press). Uniform connectedness and grouping in the perceptual organization of hierarchical patterns. *Journal of Experimental Psychology: Human Perception and Performance.*

Kimchi, R., & Goldsmith, M. (1992). Structure and process in perceptual organization. In B. Burns (Ed.), *Percepts, concepts and categories: The representation and processing of information* (pp. 77–105). Amsterdam: Elsevier.

Kleffner, D., & Ramachandran, V. (1992). On the perception of shape from shading. *Perception and Psychophysics, 52,* 18–36.

Koriat, A., & Goldsmith, M. (1996). Memory metaphors and the real-life/laboratory controversy: Correspondence versus storehouse conceptions of memory. *Behavioral and Brain Sciences, 19,* 167–188.

Koriat, A., & Goldsmith, M. (1997). The myriad functions and metaphors of memory. *Behavioral and Brain Sciences, 20,* 27–28.

Kramer, A. F., & Jacobson, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception and Psychophysics, 50,* 267–284.

Kramer, A. F., & Watson, S. E. (1996). Object-based visual selection and the principle of uniform connectedness. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 395–414). Washington, DC: American Psychological Association.

Kramer, A. F., Weber, T. A., & Watson, S. E. (1997). Object-based attentional selection—Grouped-arrays or spatially invariant representations?: Comment on Vecera and Farah (1994). *Journal of Experimental Psychology: General, 126,* 3–13.

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

Logan, G. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103,* 603–649.

Luria, A. L. (1959). Disorders of "simultaneous form" perception in a case of bilateral occipitoparietal brain injury. *Brain, 83,* 437–449.

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

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Moore, C. M., & Egeth, H. (1997). Perception without attention: Evidence of grouping under conditions of inattention. *Journal of Experimental Psychology: Human Perception and Performance, 23,* 339–352.

Nakayama, K. (1990). The iconic bottleneck and the tenuous link between early visual processing and perception. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 411–422). Cambridge, England: Cambridge University Press.

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

Navon, D. (1983). How many trees does it take to make a forest? *Perception, 12,* 239–254.

Navon, D. (1990). Does attention serve to integrate features? *Psychological Review, 97,* 453–459.

Navon, D., & Ehrlich, B. (1995). Illusory conjunctions: Does inattention really matter? *Cognitive Psychology, 29,* 59–83.

Navon, D., & Kastan, R. (1997). *Does attentional processing necessarily depend on the output of preattentive processing?* Unpublished manuscript.

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

Palmer, S. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9,* 441–474.

Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Psychonomic Bulletin and Review, 1,* 29–55.

Peterson, M. A., & Gibson, B. S. (1991). Directing spatial attention within an object: Altering the functional equivalence of shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 170–182.

Pomerantz, J. R. (1981). Perceptual organization in information processing. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 141–180). Hillsdale, NJ: Erlbaum.

Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 635–649.

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

Pomerantz, J. R., & Schwaitzberg, S. D. (1975). Grouping by proximity: Selective attention measures. *Perception and Psychophysics, 18,* 355–361.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32,* 3–25.

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

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

Prinzmetal, W., & Keysar, B. (1989). Functional theory of illusory conjunctions and neon colors. *Journal of Experimental Psychology: General, 118,* 165–190.

Prinzmetal, W., & Mills-Wright, M. (1984). Cognitive and linguistic factors affect visual feature integration. *Cognitive Psychology, 16,* 305–340.

Prinzmetal, W., Presti, D. E., & Posner, M. I. (1986). Does attention affect visual feature integration? *Journal of Experimental Psychology: Human Perception and Performance, 12,* 361–369.

Pylyshyn, Z. W. (1989). The role of location indexes in spatial perception: A sketch of the FINST spatial index model. *Cognitive Psychology, 32,* 65–97.

Rafal, R. (1996). Visual attention: Converging operations from neurology and psychology. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 139–192). Washington, DC: American Psychological Association.

Rensink, R. A., & Enns, J. T. (1995). Preemption effects in visual search: Evidence for low-level grouping. *Psychological Review, 102,* 101–130.

Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology, 23,* 299–330.

Rock, I., & Gutman, D. (1981). Effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 275–285.

Rock, I., Linnett, C. M., Grant, P., & Mack, A. (1992). Perception without attention: Results of a new method. *Cognitive Psychology, 24,* 502–534.

Shulman, G. L., Remington, R. W., & McLean, J. P. (1979). Moving attention through visual space. *Journal of Experimental Psychology: Human Perception and Performance, 5,* 522–526.

Tipper, S. P., & Behrmann, M. (1996). Object-centered not scene-based visual neglect. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 1261–1278.

Tipper, S. P., Brehaut, J., & Driver, J. (1990). Selection of moving and static objects for the control of spatially directed action. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 492–504.

Tipper, S. P., Driver, J., & Weaver, B. (1991). Object-centered inhibition of return of visual attention. *Quarterly Journal of Experimental Psychology, 43A,* 289–298.

Tipper, S. P., Weaver, B., Jerreat, L., & Burak, A. (1994). Object-based and environment-based inhibition of return of visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 478–499.

Treisman, A. (1969). Strategies and models of selective attention. *Psychological Review, 76,* 282–299.

Treisman, A. (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. (1986a). Features and objects in visual processing. *Scientific American, 254,* 114–124.

Treisman, A. (1986b). Properties, parts, and objects. In K. Boff, L. Kaufman, and J. Thomas (Eds.), *Handbook of perception and human performance: Vol. 2. Cognitive processes and performance* (pp. 35-1–35-70). New York: Wiley.

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

Treisman, A. (1990). Variations on the theme of feature integration: Reply to Navon (1990). *Psychological Review, 97,* 460–463.

Treisman, A. (1991). Search, similarity, and integration of features between and within dimensions. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 652–676.

Treisman, A. (1992a). Perceiving and re-perceiving objects. *American Psychologist, 47,* 862–875.

Treisman, A. (1992b). Spreading suppression or feature integration? A reply to Duncan and Humphreys (1992). *Journal of Experimental Psychology: Human Perception and Performance, 18,* 589–593.

Treisman, A. (1993). Representing visual objects. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: Synergies in experimental psychology, artificial intelligence, and cognitive neuroscience* (pp. 165–175), Cambridge, MA: MIT Press.

Treisman, A., Cavanagh, P., Fischer, B., Ramachandran, V. S., & von der Heydt, R. (1990). Form perception and attention: Striate cortex and beyond. In L. Spillman & J. Werner (Eds.), *Visual perception: The neurophysiological foundations* (pp. 273–316). Orlando, FL: Academic Press.

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

Treisman, A., & Gormican, S. (1988). Feature analysis in early vision: Evidence from search asymmetries. *Psychological Review, 95,* 15–148.

Treisman, A., Kahneman, D., & Burkell, J. (1983). Perceptual objects and the cost of filtering. *Perception and Psychophysics, 33,* 527–532.

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

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

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

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

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

Tsal, Y. (1983). Movements of attention across the visual field. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 523–530.

Tsal, Y. (1989). Do illusory conjunctions support the feature integration theory? A critical review of theory and findings. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 394–400.

Tsal, Y., & Kolbet, L. (1985). Disambiguating ambiguous figures by selective attention. *Quarterly Journal of Experimental Psychology, 37A,* 25–37.

Tsal, Y., & Lavie, N. (1988). Attending to color and shape: The special role of location in selective visual processing. *Perception and Psychophysics, 44,* 15–21.

Tsal, Y., & Lavie, N. (1993). Location dominance in attending to color and shape. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 131–139.

Tsal, Y., Meiran, N., & Lavie, N. (1994). The role of attention in illusory conjunctions. *Perception and Psychophysics, 55,* 350–358.

Tsotsos, J. K. (1990). Analyzing vision at the complexity level. *Brain and Behavioral Sciences, 13,* 423–469.

Ungerlieder, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Vecera, S. (1994). Grouped locations and object-based attention: Comment on Egly, Driver, and Rafal (1994). *Journal of Experimental Psychology: General, 123,* 316–320.

Vecera, S., & Farah, M. J. (1994). Does visual attention select objects or locations? *Journal of Experimental Psychology: General, 123,* 146–160.

Ward, R., Goodrich, S. J., & Driver, J. (1994). Grouping reduces visual extinction: Neuropsychological evidence for weight linkage in visual selection. *Visual Cognition, 1,* 101–130.

Watt, R. J. (1988). *Visual processing: Computational, psychophysical, and cognitive research.* Hillsdale, NJ: Erlbaum.

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt: II [Laws of organization in perceptual forms]. *Psychologische Forschung, 4,* 301–350.

Wickens, C. D., & Long, J. (1995). Object-based versus space-based models of visual attention: Implications for the design of head-up displays. *Journal of Experimental Psychology: Applied, 1,* 179–193.

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

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

Wolfe, J. M. (1996). Extending guided search: Why guided search needs a preattentive "item map." In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual selective attention* (pp. 247–270). Washington, DC: American Psychological Association.

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

Wolfe, J. M., Friedman-Hill, S. R., & Bilsky, A. B. (1994). Parallel processing of part–whole information in visual search tasks. *Perception and Psychophysics, 55,* 537–550.

Wolfe, J. M., Yu, K. P., Stewart, M. I., Shorter, A. D., Friedman-Hill, S. R., & Cave, K. R. (1990). Limitations on the parallel guidance of visual search: Color × Color and Orientation × Orientation conjunctions. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 879–892.

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

Received October 16, 1996
Revision received June 4, 1997
Accepted July 17, 1997 ■