Spatial Attention and the Apprehension of Spatial Relations

Gordon D. Logan

Seven experiments examined the role of spatial attention in apprehending spatial relations *above*, *below*, *left*, and *right*. In Experiment 1, visual search was difficult when targets differed from distractors only in the spatial relation between their elements. Reaction time increased linearly with display size with a slope greater than 60 ms/item. In Experiment 2, search was easy (the slope was flat) when targets differed from distractors in the identity of their elements. In Experiments 3 and 4, target position was cued with a discrepant color, and performance was better when attention was pulled toward spatial-relation targets than away from them. Experiments 5–7 generalized the results over different displays and extended practice. The results suggest that apprehending spatial relations requires spatial attention.

This article concerns the role of attention in the interface between perception and cognition. Conceptual representations can be independent of perceptual representations. We are able to think about things we do not see and "see" things we do not think about. However, conceptual representations are often addressed to perceptual representations; we think about what we see. The emerging area of visual cognition addresses these phenomena (e.g., Pinker, 1984). This article concerns the role of attention in visual thinking. It concerns the role of attention in establishing relations between conceptual and perceptual representations.

This article focuses on the role of spatial attention in apprehending spatial relations such as *above, below, left,* and *right*. Spatial relations are important in orienting to the environment (Pick & Acredolo, 1983), recognizing objects (Biederman, 1987; Hummel & Biederman, 1992; Marr & Nishihara, 1978), comprehending perceived (Biederman, 1981) and imagined scenes (Johnson-Laird, 1983), and directing attention around perceived (Logan, in press) and imagined spaces (Bryant, Tversky & Franklin, 1992; Franklin & Tversky, 1990). Spatial relations play an important role in language (Clark, 1973; Herskovits, 1986; Jackendoff, 1983; Levelt, 1984; Talmy, 1983; Vandaloise, 1991). They are an important topic in cognitive neuropsychology (Farah, Brunn, Wong, Wallace, & Carpenter, 1990; Kosslyn, 1987). The present investigation contributes to

these literatures by addressing the real-time processes by which spatial relations are apprehended.

The experiments tested a theory of the computations involved in apprehending spatial relations proposed by Logan (in press). According to the theory, apprehension involves coordinating perceptual and conceptual representations of space, and attention plays important roles in the process. The perceptual representations are analog arrays of objects and surfaces; the conceptual representations are propositions, like *above (dash, plus)*, that consist of a relation (*above*) and one or more arguments (*dash, plus*). The process of apprehension involves (a) spatially indexing the arguments of the relation, (b) imposing a reference frame on the argument that serves as the reference object, and (c) determining whether the object corresponding to the other argument falls within a region of acceptability for the relation in question (Logan, in press; Logan & Sadler, in press). These are attentional operations, analogous to operations discussed in the literature on visual spatial attention. The experiments used methods from that literature to test the theory's assumption that these operations are used in apprehending spatial relations.

Spatial indexing involves selecting an object in the perceptual representation and establishing correspondence between it and a symbol that stands for it in the conceptual representation. Spatial indexing is a common selection mechanism in theories of visual spatial attention, such as Treisman's feature integration theory (Treisman & Gelade, 1980; Treisman & Schmidt, 1982) and Pylyshyn's FINST theory (Pylyshyn, 1984, 1989; Pylyshyn & Storm, 1988). Spatial indexing is logically necessary in apprehending spatial relations because the arguments of spatial relations are highly schematized (i.e., they are represented as points, lines, regions, or volumes; Herskovits, 1986; Jackendoff & Landau, 1992; Talmy, 1983). Any object can be above any other. Consequently, the number of objects that can serve as arguments for a given spatial relation is indefinitely large. It is unlikely that the relation could be computed for all possible pairs of objects, so the arguments it is computed for must be selected.

Correspondence concerning this article should be addressed to Gordon D. Logan, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, Illinois 61820. Electronic mail may be sent to glogan@s.psych.uiuc.edu.

Spatial indexing is also necessary because of a semantic asymmetry between the arguments of spatial relations. One argument serves as a reference object and the other serves as a located object. The relation specifies the position of the located object with respect to the reference object (Clark, 1973; Herskovits, 1986; Jackendoff & Landau, 1992; Olson & Bialystock, 1983; Talmy, 1983). So, for example, "the dash is above the plus" specifies the position of the dash with respect to the plus. Spatial indexing is important to establish correspondence between arguments of the relations and these semantic roles; "the dash is above the plus" is different from "the plus is above the dash," and spatial indexing is necessary to maintain the distinction.

Spatial relations between objects are defined in terms of reference frames imposed on reference objects (Clark, 1973; Garnham, 1989; Levelt, 1984; Talmy, 1983). A reference frame is a three-dimensional coordinate system that has an origin, an orientation, a direction, and a scale (Hinton & Parsons, 1981; Palmer, 1989). Imposing the reference frame involves setting these parameters. The reference frame that is used to compute a spatial relation may be intrinsic to the reference object or it may be imposed on the reference object by the viewer. In either case, the origin of the reference frame is set to coincide with the location of the reference object and its orientation is set to coincide with the intrinsic orientation of the reference object or determined by the viewer. Reference frame computation is not a common operation in current theories of visual spatial attention, but it is important nevertheless because Logan (in press) showed that it is a necessary step in directing attention from cues to targets. Reference frame computation requires attention because an infinite number of reference frames could be imposed on any given display, and many of them would be inconsistent with each other. The one that is actually used must be chosen somehow, and Logan (in press) proposed that that choice is a function of attention.

Spatial relations between objects are defined in terms of regions of acceptability that are defined in terms of the reference frame imposed on the reference object (Logan & Sadler, in press; see also Herskovits, 1986; Talmy, 1983). For an object to be *above* a reference object, it must fall within a region defined with respect to the reference object and the reference frame. The best examples of *above* fall within a region aligned with the upward projection of the reference frame out of the reference object (i.e., they lie directly above the reference object). Acceptable but poorer examples of *above* fall in a region above the origin of the reference frame defined by projecting the origin horizontally, orthogonal to the up–down axis (Logan & Sadler, in press). Regions of acceptability are analogous to spotlights (Eriksen & St. James, 1986; Posner, 1980; Treisman & Gormican, 1988) and gradients (LaBerge & Brown, 1989) in the attention literature in that they can be moved around space at will and they emphasize some parts of space more than others.

Logical necessity dictates that apprehension of spatial relations requires spatial attention. However, logical necessity does not guarantee experimental results, so it is important to determine empirically whether spatial attention is involved in the apprehension of spatial relations. The apprehension of spatial relations can be investigated in several different experimental procedures. The conceptual representation can be viewed as a program that specifies a computation to be carried out on the perceptual representation. Several different programs are possible.

In cuing tasks, subjects are given an argument and a relation and asked to report an object that stands in the given relation to the given object (e.g., "What is behind the chair?"; see Bryant et al., 1992; Franklin & Tversky, 1990; Logan, in press). In this case the reference object is located (the chair), the appropriate reference frame is applied to it, and the region of acceptability associated with the relation (behind) is searched to find the object in question.

In relation judgment tasks, subjects are given two arguments and asked to report the relation between them (e.g., "Is the dash above or below the plus?"; Logan, 1980). In this case, the two arguments are located, the reference frame is imposed on the reference object, and the region in which the located object falls is compared with the regions of acceptability for the possible relations.

In verification tasks, subjects are given two arguments and a relation and asked if they apply to a given display (e.g., "is there a dash above a plus in this display?"; Clark, Carpenter, & Just, 1973). Verification can be done like a cuing task, by locating the reference object, imposing a reference frame on it, and seeing whether the located object falls within the relevant region of acceptability. Or it can be done like a relation judgment task, by locating the two objects, imposing a reference frame on the reference object, determining which region of acceptability the located argument falls within, and comparing that region with the one associated with the given relation.

In any case, regardless of the task, the person must locate the arguments in the perceptual representation, bind them to the symbols that refer to them in the conceptual representation, impose a reference frame on the reference object, and determine whether the located object falls in the region of acceptability specific to the relation in question. These processes of spatial attention are the main focus of this article. I chose to investigate them in verification tasks.

The Experiments

The experiments were like the sentence–picture verification experiments of the 1970s (e.g., Carpenter & Just, 1975; Clark & Chase, 1972; Clark et al., 1973). Subjects were presented with a sentence (e.g., "Is there a dash above a plus?") and then with a picture (e.g., a display of dashes and pluses), and their task was to indicate whether the sentence described the display, just as it was in sentence–picture verification experiments. But unlike sentence–picture verification experiments, my experiments involved visual search. My pictures contained several potential targets (e.g., several pairs of dashes and pluses), and subjects had to indicate whether any one of the potential targets matched the sentence.

My theoretical focus is different from the focus of sentence–picture verification experiments. Theories of

sentence–picture verification assumed that the picture and the sentence were both represented conceptually. Reaction time (RT) effects were interpreted in terms of a symbolic comparison process that compared the constituents of the conceptual representations (Carpenter & Just, 1975; Clark & Chase, 1972). By contrast, I assume that the picture is represented perceptually as well as conceptually, and I interpret some of the RT effects in terms of attentional processes that operate on the perceptual representation to produce a conceptual representation of the picture. My experiments, using multielement pictures, provide more opportunities for attention to influence performance than the traditional single-element pictures in the sentence–picture verification literature.

My experiments are like the conjunction search experiments of the 1980s and 1990s (e.g., Cave & Wolfe, 1990; Treisman & Gelade, 1980). Subjects searched through multielement displays to find a target that consisted of a "conjunction" of two components. My idea of conjunction is different, however. The conjunction search literature usually addresses single objects (e.g., whether a given X is red). The conceptual representation is a two-argument predicate, in which the arguments are different properties of a single object and the relation is conjunction or spatial coincidence. My experiments addressed pairs of objects. The conceptual representation was also a two-argument predicate, but the arguments were different objects and the relations were *above, below, left,* and *right* (i.e., ways in which objects do not coincide spatially). My goal was to generalize the principles discovered in experiments on spatially coincident conjunctions to a broader class of conceptual representations. The first step was to analyze relations that take two objects as arguments rather than one. My hypothesis was that apprehending two-object relations required the same sort of spatial attention that was required in single-object conjunction search.

My theoretical focus is different from the usual focus of conjunction search experiments. Theories of conjunction search focus on the perceptual representation(s), asking, for example, what constitutes a feature (Treisman & Gormican, 1988) and how perceptual similarity affects search (Cave & Wolfe, 1990; Duncan & Humphreys, 1989; Humphreys & Müller, 1993; Wolfe, Cave, & Franzel, 1989). Researchers look to psychophysics and neurophysiology for constraints on theory. By contrast, my focus is on the conceptual representation and the processes that relate it to the perceptual representation, and I look to linguistics and cognitive science for constraints on my theory. These perspectives are complementary. My goal is to build on current theories of conjunction search, not to test my predictions against theirs.

I conducted seven experiments. The first four addressed the involvement of attention in apprehending spatial relations, and the last three addressed the generality of the results. The first four experiments were conducted twice, once with *above* and *below* as relations and once with *left* and *right*. These relations are important in linguistic (Clark, 1973; Garnham, 1989; Herskovitz, 1986; Jackendoff, 1983; Levelt, 1984; Vandaloise, 1991), psychological (Bryant et al., 1992; Carpenter & Just, 1975; Clark et al., 1973;

Franklin & Tversky, 1990; Logan, in press), and neuropsychological analyses of spatial cognition (Kosslyn, 1987; Farah et al., 1990). The *above–below* experiments used vertical pairs of dashes and pluses and the *left–right* experiments used horizontal pairs. Otherwise, the experiments were identical. Experiments 1 and 2 used search tasks to see whether apprehending spatial relations requires spatial indexing. Experiments 3 and 4 addressed spatial indexing by cuing the position of the target validly and invalidly (Briand & Klein, 1987).

The questions from the sentence displays are presented in Table 1. There were four questions in each experiment, formed by factorially combining elements (dash and plus) and relations (*above* and *below* or *left* and *right*). Data were analyzed separately for each question, and question was treated as a four-level factor in the analyses of variance (ANOVAs).[1] The question effect was analyzed with three orthogonal planned comparisons. The weights for each contrast are given in Table 1.

The relation contrast addressed semantic differences between the relations, comparing *above* with *below* and *left* with *right*. Linguistic analyses suggest that *above* should be easier than *below*. Clark (1969, 1973) argued that perceptual space is normally represented positively, extending upward from the ground and forward from the viewer. Thus, *above* and *in front of* are neutral and unmarked lexically because they represent relations with respect to the viewer and the ground, which serve as natural reference objects. *Below* and *behind* are lexically marked because they refer to regions that cannot be perceived directly in terms of these natural reference objects. They are harder to apprehend because they require an unnatural reference object and an unnatural direction with respect to the reference object. This prediction was readily confirmed in many sentence–picture verification experiments (e.g., Carpenter & Just, 1975; Clark & Chase, 1972; Clark et al., 1973). Clark's analysis does not apply to *left* and *right* because there are no natural reference objects that divide space into positive and negative regions corresponding to *left* and *right*. Consequently, the relation contrast should be significant with *above* and *below* but not with *left* and *right*.

The four questions described two different targets. "Dash above plus" and "plus below dash" refer to the same target; "dash below plus" and "plus above dash" refer to another target. The target type contrast compared the two different targets, addressing the importance of perceptual differences between the targets independent of how they are described conceptually. A significant target type contrast indicates differences in perceptual processing.

---

[1] The 4-level question effect could have been analyzed as two 2-level factors and their interaction. The three contrasts in Table 1 can be interpreted as testing two 2-level factors and their interaction. Any two contrasts could be the main effects; the third would be their interaction. The analyses are equivalent except for the error terms. My approach tested the three contrasts against a common error term; treating questions as two main effects and an interaction would test the contrast against different error terms.

The argument order contrast compared the questions that began with "dash" with those that began with "plus." It addressed the importance of the surface structure of the sentence in the relation search task. The sentence appeared well before the picture, so there was plenty of time to develop a semantic representation from the surface structure to compare with the elements of the display. Thus, there should be no effect of surface structure on search performance. A significant argument order contrast would indicate that surface structure affects search performance.[2]

Experiment 1

The first experiment used a conjunction-type search task to determine whether apprehension of spatial relations requires spatial attention. Subjects searched for targets in displays of distractors that differed only in the spatial relation between their parts. Each target and distractor consisted of a dash (–) above or below a plus (+) or a dash left or right of a plus. If the target was a dash above a plus, the distractors would all be pluses above dashes. If the target was a dash to the right of a plus, the distractors would all be pluses to the right of the dashes.

I chose dashes and pluses because Treisman and colleagues suggested they were elementary features that are easy to discriminate preattentively (Treisman & Gormican, 1988; Treisman & Paterson, 1984; Treisman & Souther, 1985; see also the present Experiment 2). I wanted the difficulty in discriminating targets from distractors to stem from computing the relation, not from discriminating the arguments of the relation.

If spatial attention is required to compute spatial relations, search should be difficult. If targets and distractors cannot be distinguished by preattentive processes, spatial relations will have to be computed between the elements of each of the items in the display until the target is found. The function relating RT to the number of items in the display (display size) should be linear with a substantial slope (around 40–60 ms/item), following Treisman and Gelade (1980). Alternatively, if spatial attention is not required to compute spatial relations, search should be easy. If spatial relations can be computed preattentively, preattentive processes should be able to discriminate the target from the distractors. The target should "pop out" of the display; the function relating RT to display size should have a slope near zero (see Cave & Wolfe, 1990; Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Treisman & Gormican, 1988; Wolfe et al., 1989).

From a formal perspective, this experiment addresses capacity limitations involved in computing spatial relations rather than the serial nature of spatial indexing, imposing reference frames, and locating objects in regions of acceptability. As Townsend has argued for years, a linear function with a steep slope implies capacity limitations rather than serial search (see, e.g., Townsend, 1990; Townsend & Ashby, 1983). A flat slope may imply parallel unlimited-capacity processing, but a steep slope is consistent with parallel limited-capacity processing as well as serial (limited-

capacity) processing. Thus, the experiment asks whether computing spatial relations is a limited-capacity process.[3] Most researchers consider spatial indexing to be a serial process (e.g., Treisman, 1991; Ullman, 1984; Wolfe et al., 1989), but some suggest that spatial indexes may be deployed in parallel (Pylyshyn, 1984, 1989; Pylyshyn & Storm, 1988; but see Yantis, 1992). Everyone agrees that spatial indexing is a demanding process, whether it is parallel or serial, and the present experiment addressed that issue.

Experiments 1 and 2 were conducted at the same time and used the same subjects. There were three sessions, one for Experiment 1 and two for Experiment 2. The order in which subjects performed the experiments was counterbalanced (see below) so that one third of the subjects performed Experiment 1 each session. Experiments 1 and 2 are discussed separately for ease of exposition. The following Method and Results sections concentrate on Experiment 1.

Method

Subjects. The subjects were 24 graduate and undergraduate students who were paid $4 for each of three 1-hr sessions. Half of the subjects served in the *above–below* experiment and half served in the *left–right* experiment.

Apparatus and stimuli. The stimuli were displayed on Amdek model 722 color monitors controlled by IBM–PC XT and AT computers. Each trial consisted of a sequence of five displays: First, the fixation display was exposed for 500 ms. It consisted of a single dot (period) presented in the center of the screen. In the *above–below* experiment, the fixation point was presented in row 13, column 40 of the standard IBM 24 × 80 text screen. In the *left–right* experiment, it appeared in row 12, column 41.

Second, the sentence display was presented for 1,000 ms. It consisted of one of four questions formed by factorially combining the two arguments with the two relations. The questions used in the *above–below* and *left–right* versions of the experiment appear in Table 1. In the sentence display, the question was presented in the center of the screen, beginning on row 13, column 32 in the *above–below* experiment and on row 12, column 32 in the *left–right* experiment. Each question began with a capital letter and ended with a question mark. The other letters were lowercase. Individual letters were 5 mm high × 3 mm wide. Sentences were 5 mm high × 54 mm long. Viewing distance was not constrained,

---

[2] It is tempting to interpret the argument order contrast in terms of the distinction between *topic* and *comment* (e.g., Halliday, 1967): The topic usually corresponds to the first noun phrase in a sentence (e.g., "dash" in "dash above plus?"), and the comment is some thought about the topic (e.g., "it may be above a plus"). In verification tasks, however, subjects are asked whether the display contains a target (e.g., "contains (display, above (dash, plus))"?), so the display is the topic and containing a target (dash above plus) is the comment.

[3] The test for capacity limitations depends on demonstrating that search is self-terminating and focuses on target-present rather than target-absent responses. Parallel unlimited-capacity models predict zero slopes for target-present responses only when search is self-terminating. They predict nonzero slopes for positive responses if search is exhaustive, and they predict nonzero slopes for target-absent responses because search is necessarily exhaustive with target-absent displays (see Townsend & Ashby, 1983).

Table 1

Questions Presented in the Sentence Display for Above–Below and Left–Right Versions of Experiment 1

Question	Above–Below	Left–Right	R	T	A
1	Dash above plus	Dash left of plus	1	1	1
2	Plus below dash	Plus right of dash	−1	1	−1
3	Dash below plus	Dash right of plus	−1	−1	1
4	Plus above dash	Plus left of dash	1	−1	−1

Note. R, T, and A are weights for orthogonal contrasts testing for the effects of different relations (R), different target types (T), and argument order (A) in the main effect of question.

but the computers were situated so that it was approximately 60 cm. At this distance, individual letters subtended .48° × .29° of visual angle and sentences subtended .48° × 5.16° of visual angle.

Third, the fixation display returned for 500 ms.

Fourth, the picture display was presented. It remained on until the subject responded. It consisted of 1, 2, 4, or 8 pairs of dashes and pluses (ASCII Characters 45 and 43, respectively). They appeared in eight positions on the screen, equally spaced around an imaginary circle (within the constraints of the 24 × 80 text screen). There were two positions above the fixation point, two below it, two left of it, and two right of it. For the *above–below* experiment, the text-screen coordinates for Positions 1–8, respectively, were (1) 9,36 and 10,36; (2) 9,44 and 10,44; (3) 11,49 and 12,49; (4) 14,49 and 15,49; (5) 16,44 and 17,44; (6) 16,36 and 17,36; (7) 14,31 and 15,31; and (8) 11,31 and 12,31. For the *left–right* experiment, the coordinates for Positions 1–8, respectively, were (1) 9,36 and 9,38; (2) 9,44 and 9,46; (3) 11,49 and 11,51; (4) 14,49 and 14,51; (5) 16,44 and 16,46; (6) 16,36 and 16,38; (7) 14,31 and 14,33; and (8) 11,31 and 11,33. Targets occurred equally often in each of the eight positions with each display size (1, 2, 4, or 8 items). The positions occupied by distractors were determined randomly for Display Sizes 2 and 4. There were no distractors for Display Size 1, and all nontarget positions contained distractors for Display Size 8. Each dash was 1 mm × 2 mm and each plus was 2 mm × 2 mm in size, corresponding to .10° × .19° and .19° × .19° of visual angle, respectively. A dash–plus pair was 2 mm × 8 mm, or .19° × .76° of visual angle. The arrays were not quite circular because they were programmed on a text screen. Diagonals from Positions 1 to 5 and 2 to 6 were 58 mm, or 5.54° of visual angle. Diagonals from Positions 3 to 7 and 4 to 8 were 63 mm, or 6.02° of visual angle.

Fifth, after the subject responded, the screen went blank for a 1,500-ms intertrial interval.

Subjects responded by pressing the "/" and z keys on the computer keyboards. These were the leftmost and rightmost keys on the bottom row of the AT keyboard and nearly so for the XT keyboard.

Procedure. There were 512 experimental trials consisting of two replications of the basic experimental design. The basic design was the factorial combination of 4 display sizes (1, 2, 4, and 8), 2 types of displays (target present and target absent), 4 questions, and 8 target positions. The order of trials within each replication was randomized separately for each subject. The program paused every 128 trials to allow subjects a brief rest.

The instructions described the events on a typical trial and told subjects how to respond to them. Subjects were told they would see a sentence display followed by a picture display, and that their task was to indicate whether the target defined by the sentence display was or was not present in the picture display by pressing the "/" or z key as quickly as possible. Half of the subjects pressed

the "/" key to indicate that a target was present and the z key to indicate it was not, and half did the opposite.

This experiment was run together with two sessions of Experiment 2; the same subjects were tested throughout, using the same relations each session. The order in which subjects experienced the single session of Experiment 1 and the two sessions of Experiment 2 was balanced. There were six possible orders, and two subjects received each order. One subject in each order was assigned to each of the two response counterbalancing conditions (i.e., z for target presence, etc.).

Results

Above–below. Mean RTs were computed for each subject in each combination of display size, target presence, and question. The means across subjects, collapsed across question, are plotted in Figure 1. RT increased linearly with display size for both target-present and target-absent responses. The slopes were substantial and steeper for target-absent responses (118 ms/item) than for target-present responses (85 ms/item).

These conclusions were supported in a 4 (display size) × 2 (target-present vs. -absent) × 4 (question) ANOVA on the mean RTs. The main effect of display size was significant, $F(3, 33) = 54.99$, $p < .01$, $MS_e = 172,998.07$, as was the main effect of target presence, $F(1, 11) = 38.24$, $p < .01$, $MS_e = 61,677.31$, and the interaction between display size and target presence, $F(3, 33) = 13.78$,
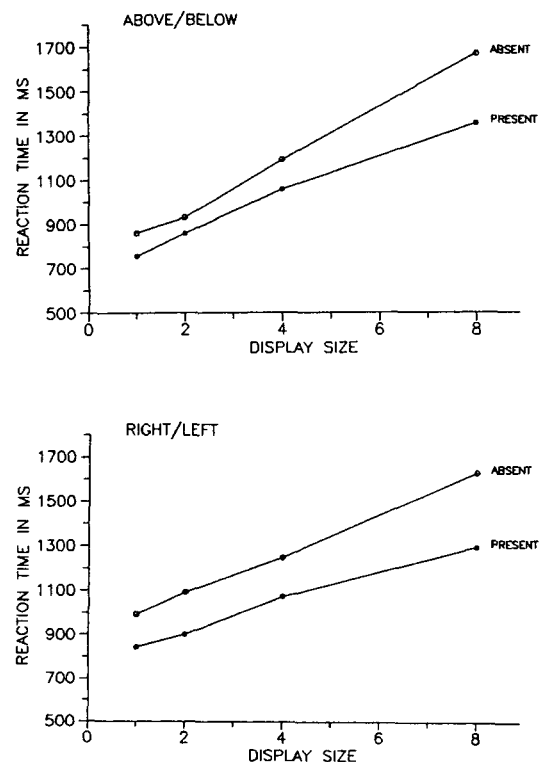

Figure 1. Mean reaction time for target-present and target-absent responses as a function of display size in Experiment 1. (Top: *above–below* condition; bottom: *left–right* condition.)

$p < .01, MS_e = 20,194.98$. A planned comparison revealed a significant linear trend in the main effect of display size, $F(1, 33) = 157.28, p < .01, MS_e = 172,998.07$, that accounted for 95% of the sum squares due to display size. Another planned comparison tested for a difference in linear trend in the interaction between display size and target presence. It was significant, $F(1, 33) = 40.26, p < .01, MS_e = 20,194.98$, and accounted for 97% of the sum squares for the interaction.

The main effect of question was significant, $F(3, 33) = 5.86, p < .01, MS_e = 34,823.85$. It was analyzed with three planned orthogonal comparisons, indicated in Table 1. The relation contrast was highly significant, $F(1, 33) = 12.94, p < .01, MS_e = 34,823.85$, showing faster responses to above than below, consistent with findings in the sentence–picture verification literature (Carpenter & Just, 1975; Clark et al., 1973). The target type contrast was also significant, $F(1, 33) = 4.52, p < .05, MS_e = 34,823.85$; displays with dash-above-plus as target were harder than displays with plus-above-dash. The argument order contrast was not significant ($F < 1$).

Question interacted significantly with display size, $F(9, 99) = 3.22, p < .01, MS_e = 14,960.81$, indicating steeper slopes with below questions (107 ms/item) than with above questions (96 ms/item) and steeper slopes with displays in which plus-above-dash was target (112 ms/item) than with displays in which dash-above-plus was target (91 ms/item). The Question × Truth interaction approached significance, $F(3, 33) = 2.89, p < .06, MS_e = 15,059.41$.

Error rate was computed for each subject for each combination of conditions. Error rate was low, averaging 5.8%, and was not strongly correlated with the mean RTs in the corresponding cells of the design ($r = -.266$). The correlation was not significant, $F(1, 30) = 2.28$.

Left–right. Mean RTs were computed for each subject in each combination of display size, target presence, and question. The means across subjects, collapsed across question, are plotted in Figure 1. RT increased linearly with display size. The slope was 66 ms/item for target present and 91 ms/item for target absent.

These conclusions were supported in a 4 (display size) × 2 (target present vs. target absent) × 4 (question) ANOVA on the mean RTs. The main effect of display size was significant, $F(3, 33) = 47.38, p < .01, MS_e = 119,787.85$, as was the main effect of target presence, $F(1, 11) = 114.48, p < .01, MS_e = 38,042.25$, and the Display Size × Target Presence interaction, $F(3, 33) = 5.78, p < .01, MS_e = 29,202.86$. A planned comparison revealed a significant linear trend in the main effect of display size, $F(1, 33) = 135.57, p < .01, MS_e = 119,787.85$, that accounted for 95% of the sum squares due to display size. Another planned comparison tested for a difference in linear trend in the interaction between display size and target presence. It was significant, $F(1, 33) = 16.33, p < .01, MS_e = 29,202.86$, and accounted for 94% of the sum squares for the interaction.

The main effect of question was not significant, $F(3, 33) = 2.13, MS_e = 65,746.76$, nor were its interactions with display size and target presence. The three planned comparisons (relation, target type, and argument order; see Table 1) were performed, and none of them was significant.

Error rate was computed for each subject for each combination of conditions. Error rate was low, averaging 5%, and was not correlated with RT ($r = .068, F < 1$).

Discussion

The data from both versions of the experiment showed a pattern characteristic of difficult conjunction search rather than preattentive pop-out. RT increased linearly with display size. The slope was large, even by the standards of the conjunction-search literature (see, e.g., Treisman & Gelade, 1980; Treisman & Gormican, 1988), and it was steeper for target-absent displays than for target-present displays, suggesting that search was self-terminating.

The slope differences are suggestive of self-terminating search: Townsend and van Zandt (1990) showed that exhaustive search is unlikely to produce anything but parallel slopes. As a further test, I analyzed the linear trend relating RT variance to the number of items in the display. Serial self-terminating search predicts steeper slopes for target-present responses than for target-absent responses, whereas serial exhaustive search predicts equal slopes (Townsend & Ashby, 1983).[4] For the above–below experiment, the slopes were 35,775 ms²/item and 15,472 ms²/item for target-present and target-absent responses, respectively. For the right–left experiment, the slopes were 19,008 and 12,504, respectively. Both experiments suggest search was self-terminating, not exhaustive.

The conclusion that search was self-terminating is important because it allows us to rule out unlimited-capacity parallel models as explanations of the data. Unlimited-capacity parallel models could account for the increase in RT with display size if search were exhaustive, but if search is self-terminating, they predict flat search functions for target-present responses (Townsend & Ashby, 1983). The steep functions for target-present responses combined with the evidence for self-terminating search allow us to rule out these models. The data are consistent with serial models or with parallel models that assume that capacity is limited. By hypothesis, the limitations in capacity stem from the need to attend to the items in the display individually.

Experiment 2

Experiment 1 suggested that spatial attention was necessary to apprehend spatial relations. Experiment 2 asked

---

[4] According to Townsend and Ashby (1983), the variance of serial self-terminating RTs is $Var_{st} = [(N + 1)/2]Var(T) + [(N - 1)(N + 1)/12][E(T)]^2$, where there are N items in the display and T is duration of the serial comparisons. The variance of serial exhaustive search is simply $Var_{ex} = NVar(T)$. The variance of serial self-terminating search increases faster with N than the variance of serial exhaustive search because there is variability in the number of comparisons performed as well as variability in the time taken to perform them. Parallel limited-capacity models can mimic these predictions.

whether it was sufficient. Subjects searched for targets that differed from distractors in the identity of their elements. The target was a dash above or below a plus or a dash to the right or left of a plus, and the distractors were all pairs of pluses or pairs of dashes. Under these conditions, search should be easy. Targets should attract attention to themselves and pop out of the distractors. The slope of the function relating RT to display size should be close to zero (Cave & Wolfe, 1990; Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Treisman & Gormican, 1988; Wolfe et al., 1989). The question was whether the spatial relation between the elements would pop out with the target. If spatial indexing is necessary and sufficient, it should. But if spatial indexing is necessary but not sufficient, further processing will be required to apprehend the relation.

Subjects were tested in two conditions, a feature condition and a relation condition. Target-present displays were the same in the two conditions. They contained one target and 0, 3, or 7 identical distractors. Target-absent displays were different. In the feature condition, all of the distractors were the same (either pairs of dashes or pairs of pluses). Subjects should be able to discriminate targets from these distractors on the basis of element identity, without having to compute the spatial relation between the target's dash and plus. Search should be fast, and the slope should be near zero.

In the relation condition, all of the distractors in target-absent displays were identical (all pairs of dashes or all pairs of pluses) except for one element, which consisted of a dash and a plus arranged in the spatial relation opposite to the target. If the target was a dash above a plus, this distractor would be a plus above a dash. Targets should attract attention and pop out with these displays, just like in the feature condition, but targets cannot be discriminated from distractors without computing the relation between the elements. If spatially indexing the target is sufficient to compute the relation, then performance in the relation condition should be no different from that in the feature condition. In each condition, attention is attracted to the target and that provides sufficient information to respond.

Logan's (in press) theory suggests that spatial indexing is necessary but not sufficient to compute spatial relations. Spatial indexing is necessary because the reference object and the located object must be distinguished, but spatial indexing does not specify relative location. In order to compute relative location, a reference frame must be imposed on the reference object, and the located object must be compared with the relevant region of acceptability. These processes should take time. Search should be slower in the relation condition than in the feature condition. The slope should be near zero, like the feature condition, but the intercept should be elevated.

The feature condition and the relation condition were run in separate sessions. These sessions were run together with the single session from Experiment 1, for a total of three sessions. The order of feature, relation, and Experiment 1 conditions was balanced across subjects.

Method

Subjects. The subjects were the same 24 who served in Experiment 1 (for pay). Half of them served in the *above–below* experiment and half served in the *left–right* experiment.

Apparatus and stimuli. These were the same as Experiment 1, except for the distractors. There were two kinds of distractors, feature distractors and relation distractors. Feature distractors were pairs of dashes or pairs of pluses. All of the feature distractors were identical in any given display. All were pairs of dashes or all were pairs of pluses. Relation distractors were a dash and a plus arranged in the opposite spatial relation to the target. If the target was a dash above a plus, the relation distractor would be a plus above a dash. Relation distractors appeared only in target-absent displays in the relation condition, and only one of the distractors was a relation distractor. The others were feature distractors.

Procedure. The procedure was the same as Experiment 1. The two sessions of Experiment 2 were conducted together with the single session of Experiment 1 such that the order of conditions was balanced across subjects and orthogonal to the balancing of response mapping conditions.

There were two types of feature distractors—pairs of pluses and pairs of dashes. Type of feature distractor was confounded with argument order in sentences. Sentences beginning with "Dash" (Questions 1 and 3; see Table 1) were followed by displays with pairs of pluses as feature distractors; sentences beginning with "Plus" (Questions 2 and 4) were followed by pairs of dashes. The confounding was justified because Experiment 1 showed no effect of argument order and because the contrast between the types of feature distractors was interesting theoretically. Treisman and colleagues (Treisman & Gormican, 1988; Treisman & Souther, 1985) found an asymmetry between searching for targets defined by feature presence, such as a plus in a display of dashes, and targets defined by feature absence, such as a dash in a display of pluses. Searching for feature presence is easy. It produces flat slopes characteristic of preattentive pop out. Searching for feature absence is hard. It produces steeper slopes more like conjunction search. If Treisman's principles generalize to relation search, slopes should be flatter when the distractors are pairs of dashes (Questions 2 and 4) than when distractors are pairs of pluses (Questions 1 and 3). The argument order contrast (see Table 1) should be interpreted as a test of Treisman's hypothesis. This test is orthogonal to the display and relation contrasts.

Subjects were instructed in the same way as in Experiment 1.

Results

Above–below. Mean RTs were calculated for each subject for each combination of task, display size, target presence, and question. The means collapsed across subjects and questions are plotted in Figure 2. In contrast with Experiment 1, RT was fast and unaffected by display size. The slope of the function relating RT to display size was 12 ms/item in the feature condition and 13 ms/item in the relation condition. These values are typical of easy feature search (Treisman & Gelade, 1980; Treisman & Gormican, 1988).

RTs were 112 ms slower in the relation condition than in the feature condition, suggesting that some extra processing went on. Target presence had different effects in the two conditions. Target-present RTs were 88 ms faster than target-absent RTs in the relation condition, but they were 10
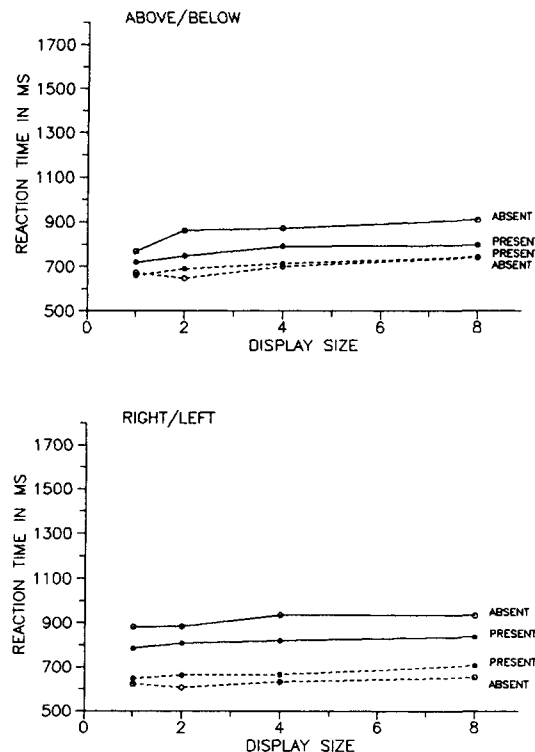
Figure 2. Mean reaction time for target-present (filled circles) and target-absent responses (open circles) as a function of display size in the feature (broken lines) and relation conditions (solid lines) of Experiment 2. (Top: *above–below* condition; bottom: *left–right* condition.)

ms slower than target-absent RTs in the feature condition. Not only did something extra go on in the relation condition; these results suggest something different went on. The interactions with question give further insight into the differences.

The mean RTs were submitted to a 2 (condition: feature vs. relation) × 4 (display size) × 2 (target present vs. target absent) × 4 (question) ANOVA. The main effect of condition approached significance, $F(1, 11) = 4.10, p < .07$, $MS_e = 33,703.18$. There were significant main effects of display size, $F(3, 33) = 18.28, p < .01, MS_e = 17,251.25$, target presence, $F(1, 11) = 6.30, p < .05, MS_e = 45,244.23$, and question, $F(3, 33) = 4.32, p < .05, MS_e = 33,703.18$. Condition interacted significantly with question, $F(3, 33) = 6.42, p < .01, MS_e = 9,515.25$, target presence, $F(1, 11) = 5.02, p < .05, MS_e = 95,749.18$, target presence and question, $F(3, 33) = 3.07, p < .05, MS_e = 6,312.34$, and display size, $F(3, 33) = 5.30, p < .01, MS_e = 5,797.62$. Question interacted significantly with display size, $F(9, 99) = 2.22, p < .05, MS_e = 12,038.65$. No other effects were significant.

The effects of display size were analyzed with planned comparisons. There was a significant linear trend in the main effect of display size, $F(1, 33) = 41.87, p < .01, MS_e = 17,251.25$, that accounted for 76.3% of the sum

squares for the main effect. The slope may have been shallow, but it was significantly greater than zero. A contrast assessing the difference in linear trend in the feature and relation conditions was not significant ($F < 1$). The 1-ms difference between the slopes was not significant.

The planned comparisons listed in Table 1 were used to analyze the Question × Condition interaction. In the relation condition, the relation contrast was significant, $F(1, 33) = 26.76, p < .01, MS_e = 9,515.25$, as was the target-type contrast, $F(1, 33) = 20.89, p < .01, MS_e = 9,515.25$. The argument order contrast, which reflects the difference between pairs of dashes and pairs of pluses as distractors, was not significant, $F(1, 33) = 1.84, MS_e = 9,515.25$. The opposite results obtained in the feature condition. The relation contrast was not significant, $F(1, 33) = 1.22, MS_e = 9,515.25$, nor was the target type contrast ($F < 1$), but the argument-order/distractor-type contrast was significant, $F(1, 33) = 12.36, p < .01, MS_e = 9,515.25$. The semantic difference between *above* and *below* mattered in the relation condition, where *above* was 52 ms faster than *below*, but it did not matter in the feature condition, where *above* was only 11 ms faster than *below*. The perceptual difference between pairs of dashes and pairs of pluses as distractors mattered in the feature condition, where pairs of dashes as distractors were responded to 35 ms faster than pairs of pluses, but not in the relation condition, where pairs of dashes were only 14 ms faster than pairs of pluses.

Error rate was low and positively correlated with RT. In the relation condition, mean percent error was 7.1 and the correlation (r) between RT and error rate was .351, $F(1, 30) = 4.22, p < .05$. In the feature condition, mean percent error was 3.6 and the correlation between RT and error rate was .151 ($F < 1$).

Left–right. Mean RTs were calculated for each subject in each combination of conditions. The means across subjects and questions are plotted in Figure 2. RT was fast and not affected much by display size. The slope of the function relating RT to display size was 7 ms/item in the feature condition and 8 ms/item in the relation condition. RT was 211 ms slower in the relation condition than in the feature condition, and the effects of target presence were different. Target-present responses were 136 ms faster than target-absent responses in the relation condition, but they were 42 ms slower than target-absent responses in the feature condition.

The mean RTs were submitted to a 2 (condition: feature vs. relation) × 4 (display size) × 2 (target present vs. target absent) × 4 (question) ANOVA. The main effects of condition, $F(1, 11) = 4.95, p < .05, MS_e = 1,227,950.96$, and display size, $F(3, 33) = 10.46, p < .01, MS_e = 11,044.34$, were significant, but the main effects of target presence, $F(1, 11) = 3.03, MS_e = 30,167.77$, and question ($F < 1$), were not. Condition interacted significantly with target presence, $F(1, 11) = 26.05, p < .01, MS_e = 47,856.69$. No other effects were significant.

The effects of display size were analyzed with planned comparisons. There was a significant linear trend in the main effect of display size, $F(1, 33) = 22.92, p < .01, MS_e = 11,044.34$, that accounted for 73% of the sum

squares for the main effect. The slope was significantly greater than zero. A contrast assessing the difference in linear trend in the feature and relation conditions, comparing the 1 ms difference in slope, was not significant ($F < 1$).

The interaction between question and condition was analyzed with planned comparisons even though it was not significant. None of the contrasts were significant in the relation condition. Only the argument-order/distractor-type contrast was significant in the feature condition, $F(1, 33) = 5.00$, $p < .05$, $MS_e = 11,534.19$, reflecting a 25-ms advantage for pairs of dashes as distractors over pairs of pluses.

Error rate was low and uncorrelated with RT. In the relation condition, mean percentage of error was 5.1 and the correlation (r) between RT and error rate was $-.053$ ($F < 1$). In the feature condition, mean percentage of error was 2.7 and the correlation between RT and error rate was .205, $F(1, 30) = 1.31$.

Discussion

The data from the *above–below* and *left–right* versions of the experiment showed the characteristics of preattentive pop-out rather than difficult conjunction search. RT was fast and hardly affected by display size. These results suggest that dashes and pluses were easy to discriminate from each other. A single dash popped out of an array of pluses, and a single plus popped out of an array of dashes. But did the relation pop out with the discrepant feature? Apparently not. Spatial indexing may be necessary to apprehend spatial relations, but it is not sufficient. In both versions of the experiment, the relation condition was slower than the feature condition. The discrepant feature may have directed attention to the right location in both conditions, but the relation condition required further processing to determine whether the dash and plus were related appropriately (i.e., imposing a reference frame on the reference object and comparing the located object with the relevant region of acceptability). This further processing took time, which was indexed by the difference between the relation and the feature conditions.

The question effect in the *above–below* experiment provided further evidence that subjects performed differently in the relation and feature conditions. The relation contrast and the target type contrast were significant in the relation condition but not in the feature condition. The relation contrast reflects semantic differences between *above* and *below* (Clark, 1969, 1973). The target type contrast reflects the difference between structured descriptions of targets (dash above or below plus; dash left or right of plus). Together, these contrasts suggest that structured semantic representations of the target were compared to the display in the relation condition but not in the feature condition. This is consistent with the conclusion that subjects computed the relation in the relation condition but not in the feature condition.

The argument order contrast reflects perceptual differences between distractors—whether they were all dashes or all pluses. It was significant in the feature condition but not in the relation condition. Its significance in the feature condition is consistent with Treisman's results with feature search (Treisman & Gormican, 1988; Treisman & Souther, 1985) although the asymmetries between feature present (dash distractors) and feature absent (plus distractors) were less marked in these experiments than in Treisman's. It is interesting that the asymmetries disappeared in the relation conditions. The perceptual differences between targets and distractors remained the same—a dash had to be discriminated from pluses, or vice versa—but the processing that was performed on the target was different. Possibly, the extra processing required to compute the relation in the relation condition absorbed the differences due to distractors. Further research will have to be done to establish the boundary conditions for observing search asymmetry effects.

The flat search functions are suggestive of parallel unlimited-capacity processing (Townsend & Ashby, 1983). The contrast between these functions and the steep functions in Experiment 1 is interesting because it provides converging evidence that apprehension of spatial relations is capacity limited: Experiment 2 used the same display positions and the same range of display sizes as Experiment 1. Only the items were different. Consequently, differences in apparent capacity limitations cannot be due to the number of items processed. Instead, they must be due to differences in how the items were processed. The items in Experiment 1 required focal attention whereas the items in Experiment 2 did not.

Experiment 3

Experiment 3 assessed the role of spatial attention in apprehending spatial relations by manipulating the direction of attention. Subjects were presented with displays that contained five dash–plus pairs. One of them was colored differently from the others, so it should pop out of the display and attract attention to itself. The main manipulation was whether or not the target was in the cued position (cue type). On cue valid trials, the discrepant item was the target. On cue invalid trials, the discrepant item was a nontarget; the target appeared in one of the noncued positions in the same color as the other distractors. If apprehension of spatial relations requires spatial indexing, then performance should be affected strongly by cue type. Performance should be good when attention is directed to the target position and bad when it is directed away from the target position.

The proportion of trials on which the cue was valid versus invalid was manipulated as well (cue validity). In one condition, the cue was valid on 80% of the trials and invalid on 20%. This imbalance in probability in favor of the target should induce subjects to attend strategically to the target's position (Posner, 1980). Cuing effects—the difference between valid and invalid cues—should be strong in this condition. In the other condition, the cue was valid on 20% of the trials and invalid on 80%. In this condition, the target was no more likely to be in the cued position than in any

other position. There was no imbalance in probability to induce attention toward or away from the cued position. Cuing effects should be weak or absent in this condition (cf. Jonides & Yantis, 1988).

In a third, neutral, condition, all of the items in the display were colored the same. There was no discrepancy to attract attention automatically or strategically. This condition served as a baseline to evaluate performance in cue-valid and cue-invalid conditions. Facilitation would be observed if cue-valid RTs were faster than neutral; inhibition would be observed if cue-invalid RTs were slower than neutral (Jonides & Mack, 1984; Posner & Snyder, 1975).

Method

Subjects. Two groups of 12 graduate and undergraduate students served as paid subjects for one 1-hr session. All subjects were screened for red–green color blindness with the Ishihara (1987) test.

Apparatus and stimuli. These were the same as in Experiment 1 except for the construction of the picture displays. The picture displays contained five dash-plus pairs approximately equally spaced on an imaginary circle centered on the fixation point. In the *above–below* experiment, the coordinates of the positions on the IBM text screen were (1) 40,9 and 40,10; (2) 49,11 and 49,12; (3) 45,15 and 45,16; (4) 35,15 and 35,16; and (5) 31,11 and 31,12. In the *left–right* experiment, they were (1) 40,10 and 42,10; (2) 49,12 and 51,12; (3) 45,16 and 47,16; (4) 35,16 and 37,16; and (5) 31,12 and 33,12. The fixation point appeared at 40,13 in the *above–below* experiment and at 41,13 in the *left–right* experiment. The imaginary circle on which the items were located was 60 mm in diameter, which corresponds to 5.73° of visual angle at a viewing distance of 60 cm. Viewing distance was not constrained.

The items were colored differently from the previous experiments. Items were white in the previous experiments; here they were red (IBM 12) and green (IBM 10). In the 20% and 80% valid conditions, one item was red and the rest were green or one item was green and the rest were red. On cue-valid trials, the discrepant item was the target; on cue-invalid trials, the discrepant item was not the target. The discrepant color was assigned randomly ($p = .25$) to each of the four nontarget positions. In the neutral cue condition, all items were red or all items were green. There were 200 neutral trials. The display was red on 80 trials and green on 80 trials. The remaining 40 trials were either all red or all green, depending on a "coin toss" in the program that generated the stimuli. In the long run, the remaining 40 trials would be red for half of the subjects and green for the other half.

Procedure. There were 600 experimental trials, divided into three blocks of 200. Cue validity (i.e., the percentage of trials on which the cue was valid) was manipulated between blocks. The order in which cue validity conditions were presented was balanced across subjects, with 2 subjects receiving each of the six possible orders. Half of the subjects had red cues and green distractors and half had the opposite. One subject in each order received each color assignment.

Each 200 trial block was constructed from five replications of 40 basic trial types, defined by the factorial combination of five target positions, four questions, and target-present versus target-absent displays. In the 80% valid condition, the cue was valid in four replications and invalid in one. In the 20% valid condition, the cue was valid in one replication and invalid in four. In the neutral condition, the cue was not valid in any replication. In all condi-

tions, the 200 trials formed by the five replications were scrambled, so that the various factors would be presented in random order. A different random order was used for each block for each subject.

Subjects were instructed as in the previous experiments, for the most part. They were given separate instructions for each cue validity condition. In the 20% and 80% valid conditions, they were told that a color cue would appear in each display and that the cue would sometimes indicate the position of the target. They were told the percentage of trials on which the cue was valid (80% or 20%). In the 20% valid condition, they were told that the relation between the cue and the target was random.

Results

Above–below. Mean RTs were computed for each subject for each combination of conditions. The means for target present responses, collapsed over questions and subjects, are plotted in Figure 3 as a function of cue validity. RTs for target-absent responses were 1,228 ms, 1,275 ms, and 1,224 ms for the 80% valid, 20% valid, and neutral conditions, respectively. Figure 3 shows that RTs were strongly affected by cue type. RTs to valid cues were much faster than RTs to invalid cues. The effects of cue type were strongly affected by cue validity. The difference between valid and invalid cues was 359 ms when 80% of the cues were valid but only 149 ms when 20% of the cues were valid. Target-present RTs on neutral trials, averaging 1,069
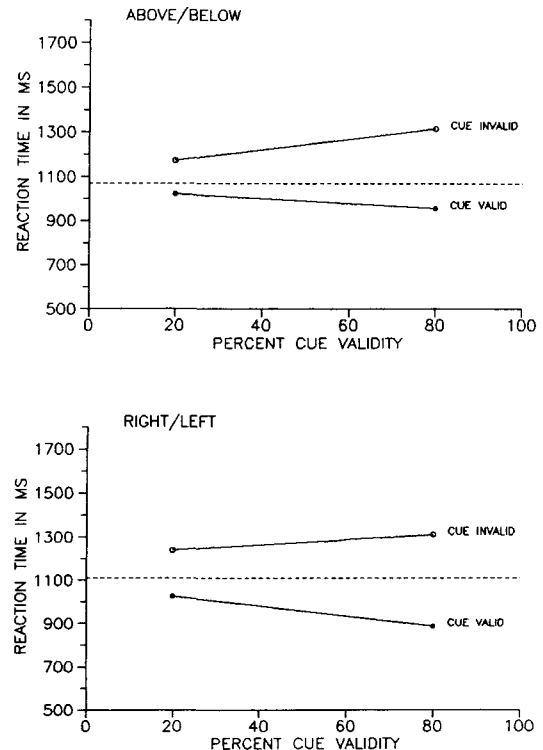
Figure 3. Mean reaction time for target-present responses as a function of percentage of cue validity in Experiment 3. (Top: *above–below* condition; bottom: *left–right* condition.)

ms, fell in between valid cue and invalid cue RTs, indicating that the cue produced both cost and benefit.

These conclusions were supported by a 2 (cue type: valid vs. invalid) $\times$ 2 (percent cue validity: 80 vs. 20) $\times$ 4 (question) ANOVA on mean RTs from target-present displays. The ANOVA focused on target-present responses from cue valid and cue invalid trials because they were more important theoretically than RTs from neutral and target-absent trials.

The main effect of cue type was significant, $F(1, 11) = 29.16$, $p < .01$, $MS_e = 106,467.34$, but the main effect of percent cue validity was not ($F < 1$). The interaction between them, reflecting the larger cue validity effect when the cue was valid 80% of the time, was significant, $F(1, 11) = 25.86$, $p < .01$, $MS_e = 20,436.74$. The main effect of question was significant, $F(3, 33) = 3.42$, $p < .05$, $MS_e = 46,004.39$. Planned comparisons (see Table 1) revealed a significant effect for relation, $F(1, 33) = 9.72$, $p < .01$, $MS_e = 46,004.39$, but not for target or argument order (both $Fs < 1$).

Error rate was low, averaging 5.6%, and positively correlated with RT, $r = .442$, $F(1, 30) = 7.30$, $p < .05$.

Left–right. Mean RTs were computed for each subject for each combination of conditions. The means for target-present responses, presented in Figure 3, replicate the findings with *above–below*. RT was strongly affected by cue type and cue validity. The difference between valid cue and invalid cue RTs was 426 ms when 80% of cues were valid and 213 ms when 20% of cues were valid. Neutral target-present RTs averaged 1,109 ms, in between valid-cue and invalid-cue RTs. The cue produced both cost and benefit. RTs to target-absent displays averaged 1,200 ms, 1,225 ms, and 1,196 ms for the 80%, 20%, and neutral conditions, respectively.

These conclusions were supported by a 2 (cue type: valid vs. invalid) $\times$ 2 (percent cue validity: 80 vs. 20) $\times$ 4 (question) ANOVA on mean RTs from target-present displays. The main effect of cue type was significant, $F(1, 11) = 38.20$, $p < .01$, $MS_e = 128,503.56$, but the main effect of percent cue validity was not ($F < 1$). The interaction between them, reflecting the larger cue validity effect in the 80% condition, was significant, $F(1, 11) = 10.33$, $p < .01$, $MS_e = 52,854.20$. The main effect of question was not significant, $F(3, 33) = 1.01$, $MS_e = 43,098.23$. None of the planned comparisons from Table 1 were significant.

Error rate was low, averaging 7.9%, and correlated positively with RT, $r = .466$, $F(1, 30) = 8.30$, $p < .01$.

Discussion

In both versions of the experiment, performance was strongly affected by the color cue. RT was faster when the target was cued and slower when a distractor was cued, as if the cued item drew attention to itself. The difference between valid and invalid cues was strongly affected by cue validity, as if attention were directed to the cue more often or more intensely when the target was more likely to be cued. The effects of color cuing and their modulation by cue

validity replicate results in the literature that were interpreted similarly, in terms of the cue attracting attention in proportion to cue validity (e.g., Jonides, 1981; Jonides & Yantis, 1988; Posner, 1980).

These results are consistent with the hypothesis that apprehension of spatial relations requires spatial attention. Spatial relations between the elements of the target were easier to compute when attention was directed to the target than when attention was directed away from the target (cf. Briand & Klein, 1987).

Experiment 4

The costs and benefits observed in Experiment 3 may reflect the effects of acuity rather than attention. The displays remained on until subjects responded, and eye movements directed toward the discrepant item would place the discrepant item in the fovea and the distractors outside it. Targets may be processed faster in the fovea than in the periphery—peripheral targets may require a further eye movement—and that difference may explain the benefits and costs of cuing. Experiment 2 suggests that acuity may not limit performance with the present displays, because subjects had no trouble detecting targets that differed from distractors in elementary features (e.g., a dash–plus pair in dash–dash pairs). Nevertheless, Experiment 4 was conducted to determine whether eye movements could account for the costs and benefits observed in Experiment 3. The procedure was the same except that the stimuli were exposed too briefly to allow eye movements (200 ms). If differences in acuity produced by eye movements were responsible for the costs and benefits of cuing in Experiment 3, then there should be no costs or benefits with the brief exposures in Experiment 4. If attention (spatial indexing) was responsible for (part of) the costs and benefits in Experiment 3, then there should be costs and benefits in Experiment 4 as well. The costs and benefits may appear in error rate instead of RT because of the brief exposure, but the pattern of effects should be essentially the same.

Method

Subjects. Two groups of 12 graduate and undergraduate students served as paid subjects for one 1-hr session. All subjects were screened for red–green color blindness with the Ishihara (1987) test.

Apparatus and stimuli. These were the same as in Experiment 3, except that the stimuli were exposed for 200 ms during the experimental trials.

Procedure. The procedure was the same as in Experiment 3, except that there were 720 trials instead of 600. The experiment was divided into three sets of 240 trials, each set with a different percentage of cue valid trials (0%, 20%, or 80%). The first 40 trials in each set were practice. The displays were exposed for 1,000 ms during the first 20 practice trials. Exposure duration was reduced to 500 ms for the last 20 practice trials, and then to 200 ms for the 200 experimental trials. The practice trials were intended to ease subjects into each condition gradually, starting with displays that were exposed long enough that accuracy should not be a problem. The 200 experimental trials were constructed in the same way as

in Experiment 3 (i.e., representing 5 replications of the 40 basic trial types).

Results

Above–below. Mean RTs were computed for each subject for each combination of conditions. The means for target-present responses, collapsed over questions and subjects, are plotted in Figure 4 as a function of cue validity. RTs for target-absent responses were 916 ms, 950 ms, and 968 ms for the 80% valid, 20% valid, and neutral conditions, respectively. Figure 4 shows that RTs were affected by cue type, although less strongly than in Experiment 3. RTs were faster with valid cues than with invalid cues, but they were not much slower with invalid cues than with neutral displays. The effects of cue type were modulated by cue validity, but not as strongly as in Experiment 3. The difference between valid and invalid cues was 123 ms when 80% of the cues were valid but only 9 ms when 20% of the cues were valid. RTs with valid and invalid cues were not much different from target-present RTs on neutral trials. Only the valid cue, 80% cue validity condition diverged substantially from the rest.

These trends did not receive much support in a 2 (cue type: valid vs. invalid) × 2 (percent cue validity: 80 vs. 20) × 4 (question) ANOVA on mean RTs from target-present displays. Three subjects had no correct responses in

the invalid cue condition when cue validity was 80%. Their RTs for that cell were estimated following Winer's procedure (1962, pp. 281–283). The only significant effect was the main effect of question, $F(3, 24) = 3.50$, $p < .05$, $MS_e = 37,558.37$. Planned comparisons (see Table 1) revealed a significant effect for relation, $F(1, 24) = 9.56$, $p < .01$, $MS_e = 37,558.37$, but not for target and argument order. The main effect of cue type was not significant, $F(1, 8) = 3.19$, $MS_e = 73,482.43$. Neither was the main effect of percent cue validity ($F < 1$) or the interaction between them, $F(1, 8) = 3.57$, $MS_e = 49,794.14$.

Error rate was much higher in this experiment than in previous ones, because of the brief exposures, averaging 18.2% over all conditions. Consequently, the error data were analyzed in detail, in the same way as the RTs. The mean error rates for target-present responses are plotted in Figure 4 as a function of cue validity. The mean error rate for target-absent responses was 14.8%, 16.5%, and 16.3% for the 80% valid, 20% valid, and neutral conditions, respectively.

Error rate for target-present responses was strongly affected by cue type, averaging 10.6% with valid cues and 28.6% with invalid ones. The difference between valid and invalid cues was strongly modulated by cue validity, averaging 26.8% in the 80% valid condition versus 9.3% in the 20% valid condition. Error rate was 19.5% in the neutral condition, between valid and invalid cues, suggesting that cuing produced both costs and benefits.

These conclusions were supported by a 2 (cue type: valid vs. invalid) × 2 (percent cue validity: 80 vs. 20) × 4 (question) ANOVA on mean percentage of error scores from target-present displays. The main effect of cue type was significant, $F(1, 11) = 29.36$, $p < .01$, $MS_e = 530.92$, and so was the main effect of percent cue validity, $F(1, 11) = 4.94$, $p < .05$, $MS_e = 367.28$. The interaction between them, reflecting the larger cue type effect when the cue was valid 80% of the time, was significant as well, $F(1, 11) = 13.55$, $p < .01$, $MS_e = 271.31$. The main effect of question was not significant, $F(3, 33) = 2.48$, $MS_e = 265.34$. Planned comparisons (see Table 1) revealed a significant effect for target, $F(1, 33) = 11.96$, $p < .01$, $MS_e = 265.34$, but not for relation or argument order (both $Fs < 1$). Pluses above dashes were easier to find than dashes above pluses.

Left–right. Mean RTs were computed for each subject for each combination of conditions. The means for target-present responses, presented in Figure 5, replicate the findings with *above–below.* RT was affected by cue type and cue validity, but less strongly than in Experiment 3. The difference between valid cue and invalid cue RTs was 282 ms when 80% of cues were valid and 50 ms when 20% of cues were valid. Neutral target-present RTs averaged 1,033 ms, in between valid-cue and invalid-cue RTs. The cue produced cost with 20% and 80% cue validity and benefit with 80% cue validity. RTs to target-absent displays averaged 1,152 ms, 1,227 ms, and 1,190 ms for the 80%, 20%, and neutral conditions, respectively.

These conclusions were supported by a 2 (cue type: valid vs. invalid) × 2 (percent cue validity: 80 vs. 20) × 4
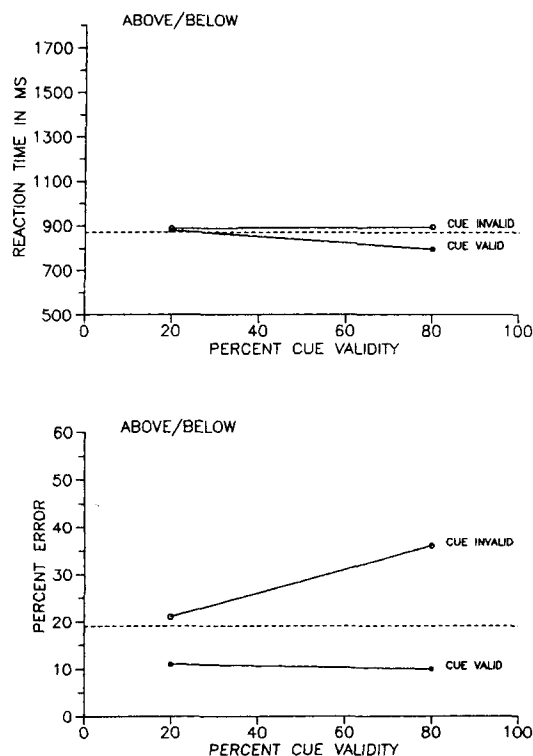
Figure 4. Mean reaction time (top) and error rate (bottom) for target-present responses as a function of percentage of cue validity in the *above–below* condition of Experiment 4.
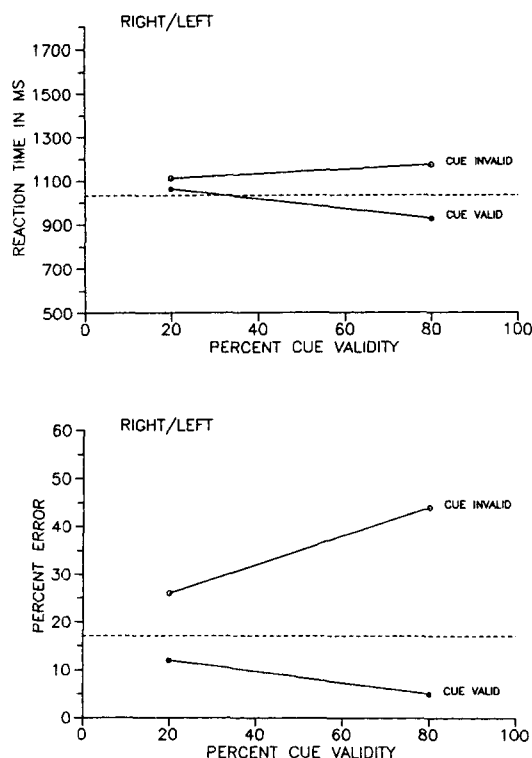
Figure 5. Mean reaction time (top) and error rate (bottom) for target-present responses as a function of percentage of cue validity in the *right–left* condition of Experiment 4.

(question) ANOVA on mean RTs from target-present displays. Two subjects had no correct responses in the invalid-cue 80% cue-validity condition. Their RTs for that cell were estimated following Winer's procedure (1962, pp. 281–283). The main effect of cue type was significant, $F(1, 9) = 20.97, p < .01, MS_e = 73,507.34$, but the main effect of percent cue validity was not $(F < 1)$, nor was the interaction between them, $F(1, 9) = 2.86, MS_e = 224,997.35$. The main effect of question was not significant, $F(3, 27) = 1.57, MS_e = 73,507.34$. None of the planned comparisons from Table 1 were significant.

Error rate was high because of the brief exposures, averaging 19.0% over all conditions, so the error data were analyzed in detail. The mean error rates for target-present responses are plotted in Figure 5 as a function of cue validity. The mean error rate for target-absent responses was 12.8%, 14.0%, and 19.5% for the 80% valid, 20% valid, and neutral conditions, respectively.

Cue type and cue validity had strong effects on target-present responses. Error rate averaged 9.0% with valid cues and 35.3% with invalid ones. The average difference between valid and invalid cues was 38.8% in the 80% valid condition and 13.8% in the 20% valid condition. Error rate was 17.5% in the neutral condition, suggesting that cuing produced both costs and benefits.

These conclusions were supported by a 2 (cue type: valid vs. invalid) × 2 (percent cue validity: 80 vs. 20) × 4

(question) ANOVA on mean error rates from target-present displays. The main effect of cue type was significant, $F(1, 11) = 20.53, p < .01, MS_e = 1,598.25$, and so was the main effect of percent cue validity, $F(1, 11) = 5.46, p < .05, MS_e = 355.07$, and the interaction between them, $F(1, 11) = 12.73, p < .01, MS_e = 579.50$. The main effect of question was not significant $(F < 1)$. Planned comparisons (see Table 1) revealed no significant effects. The Question × Cue Validity interaction was significant, $F(3, 33) = 3.82, p < .05, MS_e = 95.72$.

Discussion

In both versions of the experiment, performance was strongly affected by the color cues. Performance was better when the target was cued and worse when a distractor was cued, and the difference was larger when the cue was valid 80% (vs. 20%) of the time. The effects appeared in error rate more clearly than in RT because the stimuli were exposed briefly, but the pattern was the same as the one observed in Experiment 3. Cuing effects are easier to interpret here than in Experiment 3 because the brief exposure rules out explanations based on eye movements and acuity: The cues manipulated attention.

The results are consistent with the hypothesis that apprehension of spatial relations requires spatial attention. The spatial relation between the elements of the target was easier to apprehend when attention was directed toward the target than when attention was directed away from it.

Experiment 5

Experiments 1–4 suggest that spatial attention is necessary to apprehend spatial relations between objects. The remaining experiments tested the generality of that conclusion. Experiment 5 tested the role of linguistic representation in producing the pattern of results seen in Experiment 1: Was search difficult because the target was described linguistically? Logan's (in press) theory suggests that attention is necessary to map conceptual representations of relations onto perceptual representations, so the results should replicate even if the target is not described linguistically.

Experiment 5 was a replication of Experiment 1, except that the sentence was replaced by a picture of the target. The sentence display contained a single dash above or below a single plus or a single dash right or left of a single plus. The subjects' task was to indicate whether or not that target appeared in the subsequent picture display. If relation search was hard in Experiment 1 because subjects compared a linguistic representation of the target with the display, then presenting the target pictorially rather than linguistically should change the pattern of results. The search function may become flatter, resembling preattentive feature search. Indeed, MacLeod, Hunt, and Mathews (1978; also see Mathews, Hunt, & MacLeod, 1980) performed a standard single-item sentence–picture verification task and found that subjects who formed a pictorial representation of the sentence to compare with the picture produced a different

pattern of results than subjects who formed a linguistic representation of the sentence. Their data were less well characterized by the constituent comparison model that describes processing with the linguistic representation (Carpenter & Just, 1975).

Alternatively, Logan's (in press) theory would suggest that relation search was hard in Experiment 1 because the target had to be represented by a structured description (i.e., a proposition representing the dash, the plus, and the relation between them; see Biederman, 1987; Marr & Nishihara, 1978) and that structured description had to be compared against the display. The structured description may be extracted from the sentence through linguistic representations, but it may not depend on linguistic representations after it is extracted. If relation search was hard because subjects had to compare a structured description against the display, then presenting the target pictorially should not change the pattern of results. Subjects will have to encode the picture as a structured description and compare that structured description against the display. The search function should remain linear with a steep slope, just like Experiment 1.

Method

Subjects. The subjects were 24 graduate and undergraduate students who were paid to serve in a single 1-hr session. Half of them performed the *above–below* experiment and half performed the *left–right* experiment.

Apparatus and stimuli. These were the same as in Experiment 1, except that a picture of the target rather than a sentence (fragment) describing it appeared in the sentence display. There were only two pictures of the target—a dash above a plus and a plus above a dash in the *above–below* experiment and a dash to the left of a plus and a plus to the left of a dash in the *left–right* experiment.

Procedure. The procedure was the same as in Experiment 1. Subjects were told to search the picture display for the target presented in the sentence display.

Results

Above–below. Mean RTs were calculated in each combination of conditions. The means across subjects for target-present and target-absent responses are plotted as a function of display size in Figure 6. RT increased linearly with display size for both target-present and target-absent responses. The slope was steeper for target-absent than for target-present responses (123 ms/item vs. 86 ms/item). RT variance was regressed onto display size, and the slope was steeper for target-present responses (23,075 ms²/item) than for target-absent responses (6,556 ms²/item). These two effects are characteristic of self-terminating search (Townsend & Ashby, 1983; Townsend & van Zandt, 1990).

These conclusions were confirmed in a 4 (display size) × 4 (question) × 2 (target present vs. target absent) ANOVA on the mean RTs. The main effect of display size was significant, $F(3, 33) = 100.00$, $p < .01$, $MS_e = 100,147.27$, as was the main effect of target presence, $F(1,$
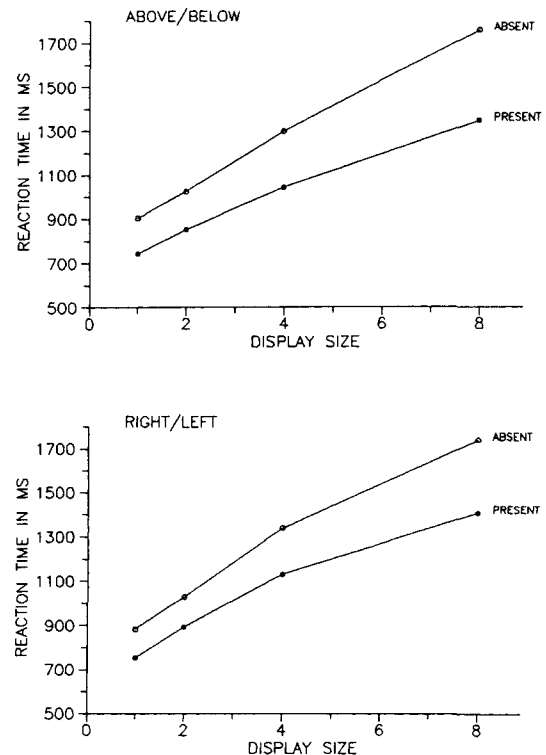
Figure 6. Mean reaction time for target-present (filled circles) and target-absent (open circles) responses as a function of display size in Experiment 5. (Top: *above–below* condition; bottom: *left–right* condition.)

$11) = 29.68$, $p < .01$, $MS_e = 203,001.75$, and the interaction between display size and target presence, $F(3, 33) = 12.05$, $p < .01$, $MS_e = 26,717.40$. Planned comparisons revealed a significant linear trend in the main effect of display size, $F(1, 33) = 282.62$, $p < .01$, $MS_e = 100,147.27$, which accounted for 94.2% of the sum squares due to display size, and a significant difference in linear trend in the interaction between display size and target presence, $F(1, 33) = 35.40$, $p < .01$, $MS_e = 26,718.40$, which accounted for 97.9% of the sum squares for the interaction.

There were four levels in the main effect of question, replicating the distinction drawn in Experiment 1, but there were only two different types of displays, corresponding to the two target types described in Table 1. Two levels reflected one target type and two reflected the other. The main effect of question was not significant $(F < 1)$ but question interacted significantly with display size, $F(9, 99) = 3.18$, $p < .01$, $MS_e = 15,813.69$. This interaction was relatively minor and did not compromise the main results.

Error rate was low, averaging 4.5%, and uncorrelated with RT, $r = -.077$ $(F < 1)$.

Left–right. Mean RTs from the *left–right* experiment are plotted in Figure 6. RT increased linearly with display size with a steeper slope for target-absent responses (123 ms/item) than for target-present responses (93 ms/item). These

conclusions were confirmed in a 4 (display size) × 4 (question) × 2 (target present vs. target absent) ANOVA on the mean RTs. The main effect of display size was significant, $F(3, 33) = 113.51, p < .01, MS_e = 95,113.52$, as was the main effect of target presence, $F(1, 11) = 59.17, p < .01, MS_e = 66,250.99$, and the interaction between display size and target presence, $F(3, 33) = 11.73, p < .01, MS_e = 18,713.27$. Planned comparisons revealed a significant linear trend in the main effect of display size, $F(1, 33) = 305.39, p < .01, MS_e = 95,113.52$, which accounted for 89.7% of the sum squares due to display size, and a significant difference in linear trend in the interaction between display size and target presence, $F(1, 33) = 33.89, p < .01, MS_e = 18,713.27$, which accounted for 96.3% of the sum squares for the interaction.

The main effect of question was not significant ($F < 1$) but question interacted significantly with display size, $F(9, 99) = 5.96, p < .01, MS_e = 14,140.78$, and with target presence and array size, $F(3, 33) = 2.47, p < .05, MS_e = 9,901.93$. These interactions were minor and did not compromise the major results.

RT variance was regressed onto display size to test for self-terminating search. The slope was steeper for target-present responses (47,646 ms²/item) than for target-absent responses (18,560 ms²/item), characteristic of self-terminating search (Townsend & Ashby, 1983).

Error rate was low, averaging 7.2%, and uncorrelated with RT, $r = .114$ ($F < 1$).

Discussion

The data from *above–below* and *left–right* versions of the experiment showed the same pattern of results, characteristic of difficult conjunction search, that was observed in Experiment 1. RT increased linearly with display size. The slopes were substantial and greater for target-absent than for target-present responses. These results, together with the results of Experiment 1, suggest that the format in which the target is presented does not have much effect on the search process. Search was difficult (i.e., capacity limited) whether the target was presented as a sentence or a picture. This suggests that in both cases, the target was represented as a structured description (i.e., a proposition), and the structured description was compared against the display.

Experiment 6

Experiments 1–5 used relatively low amounts of practice—500 to 600 trials. The results may change with greater amounts of practice (see e.g., Shiffrin & Schneider, 1977; Treisman, Vierra, & Hayes, 1992; but see Czerwinski, Lightfoot, & Shiffrin, 1992; Shiffrin, Dumais, & Schneider, 1981; Treisman & Gelade, 1980). Heathcote and Mewhort (1993) found that a moderate amount of practice (1,600 trials) produced a marked change in search for targets that differed from distractors only in the spatial arrangement of their components. The slope of the function relating RT to

display size was steep initially but flattened to near zero by the end of practice.

Experiment 6 was designed to determine whether extended practice would change the pattern of performance seen with the present displays. Experiment 6 focused on *above* and *below* rather than *left* and *right* to take advantage of the effect of the semantic difference between *above* and *below* on performance (see Experiments 1–4; also see Carpenter & Just, 1975; Clark & Chase, 1972; Clark et al., 1973). With extended practice, subjects may be able to learn to respond to the displays without computing spatial relations (see, e.g., Donnelly, Humphreys, & Riddoch, 1991; Duncan & Humphreys, 1989; Humphreys, Quinlan, & Riddoch, 1989; Wolfe et al., 1989; but see Wolfe et al., 1990). Logan's (in press) theory assumes that subjects apprehend spatial relations by establishing correspondence between perceptual and conceptual representations of the items and performing certain computations. Subjects justify their assertions about target presence or absence by performing these computations (Logan, in press). The (relation) contrast between *above* and *below* provides a way to test this assumption. It should be significant if subjects justify assertions by computing relations between elements; it should not be significant if subjects find some way to justify assertions that does not involve computing relations.

Method

Three subjects served in the *above–below* version of Experiment 1 for twelve 512-trial sessions.

Results

The slopes for target-present and target-absent displays, averaged across the three subjects, are plotted as a function of practice session in Figure 7. The first session replicated Experiments 1 and 5: The slope was 65 ms/item for target-present responses and 94 ms/item for target-absent responses. The slopes stayed pretty much the same over practice, averaging 59 ms/item for target-present responses
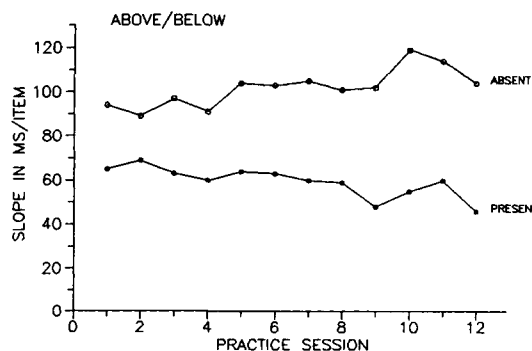
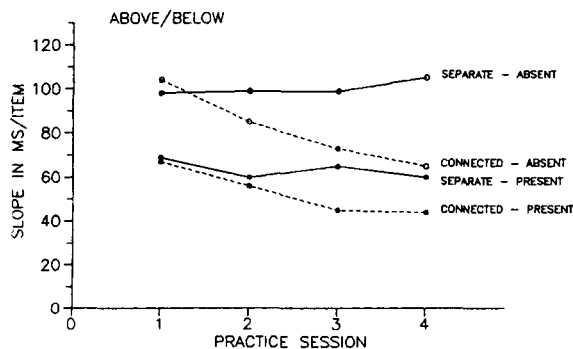*Figure 7.* Mean slope of the linear function relating reaction time to display size for target-present and target-absent responses as a function of practice session in extended practice searching for *above–below* in Experiment 6.

and 102 ms/item for target-absent responses. The slopes were substantial even on the twelfth session, averaging 46 ms/item for target-present responses and 104 ms/item for target-absent responses. Targets did not pop out, even after 6,144 practice trials.

Notice that the slopes for target-present and target-absent responses diverged a little with practice. The ratio of target-absent to target-present slopes was 1.44 on the first session, near the value for Experiment 1 (1.39). The ratio increased with practice, approximating a 2:1 ratio by Session 9 (the ratios for Sessions 9–12 were 2.13, 2.18, 1.93, and 2.25, respectively). It is tempting to interpret this effect as suggesting that subjects became more efficient at self-terminating search over practice, because many researchers interpret a 2:1 slope ratio as characteristic of self-terminating search (e.g., Humphreys & Müller, 1993; Treisman & Gelade, 1980; Treisman & Gormican, 1988). However, Townsend and others (e.g., Townsend & Ashby, 1983; Townsend & van Zandt, 1990) have shown that self-terminating search can produce a variety of slope ratios, depending on one's assumptions about the rates of processing targets and distractors, so it is best not to yield to that temptation.

These conclusions were supported by 2 (target present vs. target absent) $\times$ 4 (question) $\times$ 12 (session) ANOVAs on the slopes and intercepts of linear functions fitted to each subject's RTs. In the slope analysis, there was a significant main effect of target presence, $F(1, 2) = 32.70, p < .05, MS_e = 3,990.11$, a marginal main effect of question, $F(3, 6) = 4.37, p < .06, MS_e = 1,096.30$, and a significant interaction between session and target presence, $F(11, 22) = 4.94, p < .01, MS_e = 214.39$. The main effect of session was not significant ($F < 1$).

The main effect of question was analyzed with the planned comparisons in Table 1. The analysis revealed a significant effect for relation, $F(1, 6) = 11.97, p < .05, MS_e = 1,096.30$, showing that slopes were shallower for *above* than for *below* (71 ms/item vs. 84 ms/item, respectively), and a significant effect for target, $F(1, 6) = 33.25, p < .01, MS_e = 1,096.30$.

The intercept ANOVA revealed significant main effects of session, $F(11, 22) = 3.00, p < .05, MS_e = 11,017.12$, and question, $F(3, 6) = 8.94, p < .05, MS_e = 3,157.61$. Planned comparisons (see Table 1) revealed a significant effect for relation, $F(1, 6) = 18.52, p < .01, MS_e = 3,157.61$, showing lower intercepts for *above* than for *below* (455 ms vs. 484 ms, respectively), and a significant effect for target, $F(1, 6) = 32.07, p < .01, MS_e = 3,157.61$.

Error rate was low, averaging 5.4% over the whole experiment. It was generally uncorrelated with RT. Correlations between RT and error rate calculated within each session averaged −.199.

Discussion

This experiment showed that search remained difficult over 12 sessions. The spatial relation between the elements of an item did not pop out despite 6,144 trials of practice. The significance of the relation contrast in the analyses of

slopes and intercepts and the null interactions between question and sessions suggest that subjects computed relations between the elements of the items throughout practice. Thus, the pattern of performance in Experiments 1 and 5 was not due to the low levels of practice involved.

Experiment 7

The final experiment examined practice effects in searching for relations between and within objects. Several experiments in the literature reported pop-out for targets that differed from distractors in the spatial arrangements of their parts, even at low levels of practice (Donnelly et al., 1991; Duncan & Humphreys, 1989; Enns & Resnick, 1990a, 1990b, 1991; Humphreys et al., 1989; but see Wolfe et al., 1990). Those experiments involved within-object relations, whereas the present experiments involved between-object relations. Some recent evidence suggests that within-object relations may be easier to process than between-object relations (Baylis & Driver, 1993; Elder & Zucker, 1993), so Experiment 7 compared search for targets defined by spatial relations within and between objects.

There were two groups of subjects, a between-object group and a within-object group. All subjects saw sentences describing relations (*above* or *below*) between dashes and pluses. Subjects in the between-object group saw displays containing separate dashes and pluses, much like Experiment 1. Subjects in the within-object group saw displays in which the dashes and pluses were connected to form a single object that looked like a cross extending upward from a baseline or a *T* with a bar across its stem. All subjects were trained for four 512-trial sessions in case the tendency to pop out developed with practice.

As in Experiment 6, this experiment focused on *above* and *below* in order to exploit the performance difference between them as an index that subjects computed spatial relations between the elements (or parts) of the items. A significant relation contrast was interpreted as evidence of semantic processing.

Method

Subjects. Two groups of six graduate and undergraduate students served as paid subjects for four 1-hr sessions.

Apparatus and stimuli. The apparatus was the same as in the previous experiments but the picture displays were different. Whereas the previous experiments displayed pictures in text mode, Experiment 7 displayed them in graphics mode. Pairs of dashes and pluses were drawn with a graphics program. Two sets were constructed, a between-object set and a within-object set. In the between-object set, the dashes and pluses were separated; in the within-object set, the "tail" of the plus was extended until it intersected the dash to form a single object. Each dash–plus pair was 10 mm high and 4 mm wide. The gap between the dash and the plus in the separated displays was 3 mm. Viewed at a typical (but unconstrained) distance of 60 cm, each pair subtended .95° $\times$ .38° of visual angle and the gap between the dash and the plus in the separated displays subtended .29°.

As in the previous experiments, each picture display contained 1, 2, 4, or 8 dash–plus pairs arrayed in an imaginary circle.

However, this experiment used a different set of alternative positions, corresponding to 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315° from the top of the circle. The diameter of the circle was 6 cm or 5.71° of visual angle. Targets occurred equally often in each of the eight positions with each display size. Distractor positions were determined randomly for Display Sizes 2 and 4.

Time parameters were the same as in Experiment 1.

Procedure. The procedure was the same as in Experiment 1, except that each subject served for four 512-trial sessions.

Results

The slopes of linear functions relating RT to display size are plotted in Figure 8 as a function of practice session. For both between- and within-object relations, slopes were steeper for target-absent displays than for target-present displays, suggesting self-terminating search. Slopes were stable over practice for between-object displays, averaging 84 ms/item on the first session and 82 ms/item on the fourth. Slopes diminished a little with practice for within-object relations, averaging 86 ms/item on the first session and 55 ms/item on the fourth. Even though the slopes diminished, there was no tendency for targets to pop out of within-object displays; slopes were well above zero even after 2,048 trials of practice (cf. Heathcote & Mewhort, 1993).

Target-absent slopes were steeper than target-present slopes throughout practice for both between- and within-object relations. The ratio of target-absent to target-present slopes was 1.58 for between-object relations and 1.54 for within-object relations, suggestive of self-terminating search (Townsend & van Zandt, 1990). The variance of RT was regressed onto display size as a further test for self-terminating search. For between-object relations, the slope was 16,888 ms^2/item for target-present responses and 6,189 ms^2/item for target-absent responses. For within-object relations, the slope was 11,527 ms^2/item for target-present responses and 3,876 ms^2/item for target-absent responses, suggestive of self-terminating search (Townsend & Ashby, 1983).

Mean RTs were analyzed in a 2 (between-object vs. within-object relations) × 4 (display size) × 2 (target present vs. target absent) × 4 (question) × 4 (session) ANOVAs. There were significant main effects for display size, $F(3, 30) = 135.01, p < .01, MS_e = 151,618.39$, target presence, $F(1, 10) = 126.70, p < .01, MS_e = 68,047.53$, question, $F(3, 30) = 11.54, p < .01, MS_e = 39,673.74$, and session, $F(3, 30) = 19.85, p < .01, MS_e = 219,753.11$. There were significant two-way interactions between question and target presence, $F(3, 30) = 11.09, p < .01, MS_e = 14,933.46$, question and session, $F(9, 90) = 4.31, p < .01, MS_e = 5,335.47$, question and display size, $F(3, 30) = 9.01, p < .01, MS_e = 12,358.06$, and target presence and session, $F(3, 30) = 8.40, p < .01, MS_e = 13,981.23$. There was a significant interaction between target presence and display size, $F(3, 30) = 24.97, p < .01, MS_e = 41,891.65$, reflecting the steeper slopes with target-absent responses. There were significant interactions between session and display size, $F(9, 90) = 5.45, p < .01, MS_e = 113,837.92$, and between session, display size, and between- versus within-object relations, $F(9, 90) = 3.49, p < .01, MS_e = 113,837.92$, reflecting the (slight) reduction in slope with practice that was greater for within-object relations than for between-object relations. The main effect of between- versus within-object relations was not significant ($F < 1$). It interacted significantly with question and target presence, $F(3, 30) = 3.35, p < .05, MS_e = 14,933.46$, question and display size, $F(9, 90) = 2.57, p < .05, MS_e = 12,358.06$, and sessions and display size, $F(9, 90) = 3.49, p < .01, MS_e = 13,837.92$.

The effect of question was analyzed with the planned comparisons in Table 1. Between- and within-object relations were analyzed separately, using the error term for the interaction between question and relations ($MS_e = 39,673.73$). For between-object relations, there was a significant effect for relation, $F(1, 30) = 10.50, p < .01$, indicating faster RTs for *above* than for *below* (1,043 ms vs. 1,093 ms, respectively). The effect of target was significant, $F(1, 30) = 6.39, p < .05$, but the effect of argument order was not ($F < 1$). For within-object relations, the effect of relation was significant as well, $F(1, 30) = 7.55, p < .05$, indicating faster RTs for *above* than for *below* (919 ms vs. 958 ms, respectively). The effects of target, $F(1, 30) = 6.81, p < .05$, and argument order, $F(1, 30) = 7.55, p < .05$, were also significant.

The contrast between *above* and *below* was significant in each session for between-object relations, though its magnitude was greater in the first session than in subsequent sessions. For Sessions 1–4, the Fs were 54.03, 9.50, 21.16, and 14.76, respectively, with 1 and 90 degrees of freedom and $MS_e = 5,335.47$. The pattern was similar for within-object relations. The contrast between *above* and *below* was large in Session 1, significant in Sessions 2 and 3, but not significant in Session 4. For Sessions 1–4, the Fs were 51.28, 6.56, 16.25, and 1.52, respectively, with 1 and 90 degrees of freedom and $MS_e = 5,335.47$.

Error rate was low, averaging 6.8% for between-object relations and 4.4% for within-object relations. It was uncorrelated with RT for between-object relations, $r = -.147$,

Figure 8. Mean slope of the linear function relating reaction time (RT) to display size for target-present (filled circles) and target-absent responses (open circles) for separate (solid lines) and connected (broken lines) objects in Experiment 7.

$F(1, 30) < 1$, and for within-object relations, $r = -.280$, $F(1, 30) = 2.58$.

Discussion

The results from between- and within-object search largely replicated the results of the previous experiments. Search was difficult. RT increased linearly with display size and the slopes were steep and steeper with target-absent displays than for target-present displays. There was some suggestion that within-object search was affected more by practice than between-object search. The slope flattened somewhat and by the fourth session, the (relation) contrast between *above* and *below* was no longer significant, which suggests that subjects may have discovered some way to perform the task without computing the spatial relations. However, even on the fourth session, the within-object slopes were substantially greater than zero, much larger than the 0 to 10 ms/item slopes that are characteristic of preattentive pop-out (Duncan & Humphreys, 1989; Treisman & Gelade, 1980; Wolfe et al., 1989).

The results suggest that the difference between the present experiments and the ones in the literature in which targets defined by spatial relations pop out of distractors (e.g., Donnelly et al., 1991; Duncan & Humphreys, 1989; Enns & Resnick, 1990a, 1990b, 1991; Humphreys et al., 1989) is not that between-object relations are harder than within-object relations (see also Wolfe et al., 1990). I suspect the difference lies in the nature of the relations. The relations studied in the present experiments (*above, below, left,* and *right*) take two objects as arguments. They apply naturally to separate objects, but they can apply easily to parts of a single object; the single objects are parsed into parts and the parts are treated as separate objects. It is not clear what relations exist between the parts of the single objects studied in the experiments that produced pop-out. Those authors said little about the semantics of the relations they studied, focusing instead on judged similarity of targets and distractors and homogeneity versus heterogeneity of distractors. By contrast, the semantics of the relations I used are clear, having been studied by linguists for two decades at least (Clark, 1973; Garnham, 1989; Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1992; Levelt, 1984; Miller & Johnson-Laird, 1976; Olson & Bialystock, 1983; Talmy, 1983; Vandaloise, 1991). Further investigation of the semantic differences between within- and between-object relations may clarify the issue.[5]

General Discussion

Does the apprehension of spatial relations require spatial attention? Three lines of evidence converge on the conclusion that it does. First, search for targets that differed from distractors only in the spatial relation between their elements was very difficult in Experiments 1, 5, 6, and 7. The data were characteristic of difficult conjunction search: RT increased linearly with display size with a substantial slope, and the slope was steeper for target-absent arrays than for target-present arrays. Data like these are interpreted as evidence consistent with serial self-terminating processing in the conjunction search literature, and the serial processing involves moving attention from item to item (e.g., Treisman & Gelade, 1980). Thus, the data suggest that apprehension of targets defined by the spatial relations of their elements requires spatial attention.

From a more formal perspective, the data are consistent with parallel limited-capacity processing as well as serial processing: The strong evidence for self-terminating search together with the steep slopes on target-present trials rules out parallel unlimited-capacity models. Thus, the data suggest that search for targets defined by the spatial relations of their elements is capacity limited. I suggest that search was capacity limited because subjects had to spatially index items in the display, impose a reference frame on the reference objects, and compare the located objects with the relevant region of acceptability.

Second, in Experiment 2, search for targets that differed from distractors in their elements was easy, characteristic of easy feature search. Targets popped out of the display; slopes were close to zero, around 10 ms/item. However, the relation between the elements did not pop out with the target. Extra processing was required in the relation condition, in which the item that popped out was sometimes a distractor with elements in the opposite spatial relation. Spatial indexing may be necessary to apprehend spatial relations, but it was not sufficient. Imposing a reference frame and comparing the located object with a region of acceptability were necessary as well (Logan, in press; Logan & Sadler, in press).

The data from Experiment 2 are consistent with parallel unlimited-capacity processing, and that converges on the conclusion that the capacity limitations seen in Experiments 1, 5, 6, and 7 were due to the requirement to attend to the display items individually. Experiment 2 used the same display positions and display elements but did not require subjects to compute the spatial relations between the elements to find the target.

Third, in Experiments 3 and 4, search for targets was strongly affected by cuing attention. Performance was better when the cues pulled attention toward the target than when they pulled it away. These data are consistent with the hypothesis that apprehension of spatial relations requires spatial attention (see Briand & Klein, 1987). The data could be consistent with a parallel limited-capacity model, like

---

[5] Duncan and Humphreys (1989) and Humphreys et al. (1989) showed that steep slopes sometimes occur when the ratio of item size to retinal eccentricity is low, which is essentially an acuity effect: If the items are small and far from the center of the retina, they will be hard to discriminate in peripheral vision. Size:eccentricity ratio cannot explain the steep slopes in the present experiment. I ran an experiment with the same display program, in which subjects searched for upside-down *T*s in uniform arrays of upright *T*s, and found shallow slopes. The *T*s were the same size as the dash–plus pairs in this experiment, and the diameter of the display was the same (so the size:eccentricity ratio was the same), but the slopes were 13 ms/item for target-present responses and 28 ms/item for target-absent responses.

for such targets was difficult, and Experiments 3 and 4 showed that search was easier when attention was directed toward such targets and harder when attention was directed away from them. Experiment 2 showed that spatial indexing may be necessary for computing spatial relations, but it is not sufficient.

References

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

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization.* (pp. 213–263). Hillsdale, NJ: Erlbaum.

Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review, 94,* 65–96.

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.

Bryant, D. J., Tversky, B., & Franklin, N. (1992). Internal and external spatial frameworks for representing described scenes. *Journal of Memory and Language, 31,* 74–98.

Carpenter, P. A., & Just, M. A. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review, 82,* 45–73.

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

Clark, H. H. (1969). Linguistic processes in deductive reasoning. *Psychological Review, 76,* 387–404.

Clark, H. H. (1973). Space, time, semantics, and the child. In T. E. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 27–63). San Diego, CA: Academic Press.

Clark, H. H., Carpenter, P. A., & Just, M. A. (1973). On the meeting of semantics and perception. In W. G. Chase (Ed.), *Visual information processing* (pp. 311–381). San Diego, CA: Academic Press.

Clark, H. H., & Chase, W. G. (1972). On the process of comparing sentences against pictures. *Cognitive Psychology, 3,* 472–517.

Czerwinski, M., Lightfoot, N., & Shiffrin, R. M. (1992). Automatization and training in visual search. *American Journal of Psychology, 105,* 271–315.

Donnelly, 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.

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

Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research, 33,* 981–991.

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

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

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

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

Farah, M. J., Brunn, J. L., Wong, A. B., Wallace, M. A., & Carpenter, P. A. (1990). Frames of reference for allocating attention to space: Evidence from the neglect syndrome. *Neuropsychologica, 28,* 335–347.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General, 119,* 63–76.

Garnham, A. (1989). A unified theory of the meaning of some spatial relational terms. *Cognition, 31,* 45–60.

Greenspan, S. L., & Segal, E. M. (1984). Reference and comprehension: A topic-comment analysis of sentence–picture verification. *Cognitive Psychology, 16,* 556–606.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English: II. *Journal of Linguistics, 3,* 199–244.

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

Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English.* Cambridge, England: Cambridge University Press.

Hinton, G. E., & Parsons, L. M. (1981). Frames of reference and mental imagery. In J. Long & A. D. Baddeley (Eds.), *Attention and performance* (Vol. 9, pp. 261–277). Hillsdale, NJ: Erlbaum.

Hummel, J. E., & Biederman, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review, 99,* 480–517.

Humphreys, G. W., & Müller, H. J. (1993). SEarch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology, 25,* 43–110.

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.

Ishihara, S. (1987). *Ishihara's tests for colour-blindness.* Tokyo: Kanehara & Co.

Jackendoff, R. (1983). *Semantics and cognition.* Cambridge, MA: MIT Press.

Jackendoff, R., & Landau, B. (1992). Spatial language and spatial cognition. In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman* (pp. 145–170). Hillsdale, NJ: Erlbaum.

Johnson-Laird, P. N. (1983). *Mental models.* Cambridge, MA: Harvard University Press.

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye movement. In J. Long & A. D. Baddeley (Eds.), *Attention and performance* (Vol. 9, pp. 187–203). Hillsdale, NJ: Erlbaum.

Jonides, J., & Mack, R. (1984). The cost and benefit of costs and benefits. *Psychological Bulletin, 96,* 29–44.

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

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres. *Psychological Review, 94,* 148–175.

LaBerge, D., & Brown, V. (1989). Theory of attentional operations in shape identification. *Psychological Review, 96,* 101–124.

Lassaline, M. L., & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 561–581.

Levelt, W. J. M. (1984). Some perceptual limitations on talking about space. In A. J. van Doorn, W. A. de Grind, & J. J. Koenderink (Eds.), *Limits on perception* (pp. 323–358). Utrecht, The Netherlands: VNU Science Press.

Logan, G. D. (1980). Attention and automaticity in Stroop and

priming tasks: Theory and data. *Cognitive Psychology, 12,* 32–63.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology, 22,* 1–35.

Logan, G. D. (in press). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology.*

Logan, G. D., & Sadler, D. D. (in press). A computational analysis of the apprehension of spatial relations. In M. Peterson & P. Bloom (Eds.), *Language and space.* Cambridge, MA: MIT Press.

MacLeod, C. M., Hunt, E. B., & Mathews, N. N. (1978). Individual differences in the verification of sentence–picture relationships. *Journal of Verbal Learning and Verbal Behavior, 17,* 493–507.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, 200,* 269–294.

Mathews, N. N., Hunt, E. B., & MacLeod, C. M. (1980). Strategy choice and strategy training in sentence-picture verification. *Journal of Verbal Learning and Verbal Behavior, 19,* 531–548.

Miller, G. A., & Johnston-Laird, P.N. (1976). *Language and perception.* Cambridge, MA: Harvard University Press.

Nicoletti, R., & Umiltà, C. (1989). Splitting visual space with attention. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 164–169.

Olson, D., & Bialystock, E. (1983). *Spatial cognition.* Hillsdale, NJ: Erlbaum.

Palmer, S. E. (1989). Reference frames in the perception of shape and orientation. In B. E. Shepp & S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 121–163). Hillsdale, NJ: Erlbaum.

Pick, H. L., & Acredolo, L. P. (Eds.). (1983). *Spatial orientation: Theory, research, and application.* New York: Plenum.

Pinker, S. (1984). Visual cognition: An introduction. *Cognition, 18,* 1–63.

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

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium* (pp. 55–85). Hillsdale, NJ: Erlbaum.

Pylyshyn, Z. (1984). *Computation and cognition.* Cambridge, MA: MIT Press.

Pylyshyn, Z. (1989). The role of location indices in spatial perception: A sketch of the FINST spatial-index model. *Cognition, 32,* 65–97.

Pylyshyn, Z., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3,* 179–197.

Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 4,* 586–598.

Shiffrin, R. M., Dumais, S. T., & Schneider, W. (1981). Characteristics of automatism. In J. Long & A. D. Baddeley (Eds.), *Attention and performance* (Vol. 9, pp. 223–238). Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: Part 2. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Simon, J. R. (1969). Reactions toward the source of stimulation. *Journal of Experimental Psychology, 81,* 174–176.

Talmy, L. (1983). How language structures space. In H. L. Pick & L. P. Acredolo (Eds.), *Spatial orientation: Theory, research, and application* (pp. 225–282). New York: Plenum.

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

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes.* Cambridge, England: Cambridge University Press.

Townsend, J. T., & van Zandt, T. (1990). New theoretical results on testing self-terminating vs. exhaustive processing in rapid search experiments. In H. Geissler (Ed.), *Psychophysical explorations of mental structures* (pp. 469–489). Göttingen, Germany: Hogrefe & Huber.

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

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,* 14–48.

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

Treisman, A., & Schmidt, 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., Vieira, A., & Hayes, A. (1992). Automaticity and preattentive processing. *American Journal of Psychology, 105,* 341–362.

Ullman, S. (1984). Visual routines. *Cognition, 18,* 97–159.

Umiltà, C., & Nicoletti, R. (1985). Attention and coding effects in S–R compatibility due to irrelevant spatial cues. In M. I. Posner & O. S. M. Marin (Eds.), *Attention and performance* (Vol. 11, pp. 457–471). Hillsdale, NJ: Erlbaum.

Vandaloise, C. (1991). *Spatial prepositions: A case study from French.* Chicago: University of Chicago Press.

Winer, B. J. (1962). *Statistical principles in experimental design.* New York: McGraw-Hill.

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

Wolfe, J. M., 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.