

Distance and Distraction Effects in the Apprehension of Spatial Relations

Gordon D. Logan and Brian J. Compton
University of Illinois

Theories of the apprehension of spatial relations differ in the predictions they make about the effects of distance between the arguments of spatial relations and the effects of distracting stimuli presented along with the arguments. One theory predicts no effect of distance, another predicts a monotonic increase in reaction time (RT) with distance, and a third predicts a monotonic decrease. Most theories predict slower RTs but reasonable accuracy when distractors are present, but 1 theory predicts chance-level accuracy. These predictions were tested in 3 sentence–picture comparison experiments, in which subjects searched for targets exemplifying the relations *above* and *below*. Distance had no effect when no distractors were present. Distractors slowed RT but did not reduce accuracy to chance levels. These results suggest modifications to many of the theories of apprehension.

This article concerns the apprehension of spatial relations, such as *above* and *below*. Spatial relations are important in a number of research areas, including psycholinguistics (Clark, 1973; Garnham, 1989; Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1992; Landau & Jackendoff, 1993; Levelt, 1984; Miller & Johnson-Laird, 1976; Talmy, 1983; Vandaloise, 1991), cognitive neuroscience (Farah, Brunn, Wong, Wallace, & Carpenter, 1990; Hellige & Michimata, 1989; Kosslyn, 1987; Kosslyn et al., 1989; Sergent, 1991), and computational vision (Biederman, 1987; Marr & Nishihara, 1978; Ullman, 1984), as well as spatial cognition (Bryant, Tversky, & Franklin, 1992; Franklin & Tversky, 1990; Johnson-Laird, 1983; Logan, 1995; Pick & Acredolo, 1983). The purpose of this article is to distinguish between theories of the process of apprehension, focusing on their ability to account for the effects of distance between the arguments of the relation and the effects of distractors presented along with the arguments. Current theories of apprehension predict different effects of these variables.

The experiments were variants of the sentence–picture comparison procedure of the 1970s (e.g., Clark, Carpenter, & Just, 1973), in which subjects were given sentences describing the spatial relation between two letters (e.g., “B ABOVE D”), followed by pictures that contained an array of letters that did or did not exemplify the relation. Half of the time, sentences and pictures corresponded; half of the time, they did not. The main independent variables were distance and distraction. Distance was varied in four equally spaced steps. To separate absolute distance from relative distance, subjects were tested with two different step sizes,

one twice the size of the other. If absolute distance was the critical variable, the difference between one step and four should be greater when step size was greater; if relative distance mattered, the difference between one step and four should be the same regardless of step size.

Distraction was manipulated by presenting letter pairs alone or in a vertical array of nine letters spaced one step apart (see Figure 1c). Target letters (corresponding to the arguments of the relation) could appear in Positions 2–8. All of the other positions were filled with random consonants. The number of distractors was held constant at seven per display, but the number of distractors between the arguments varied with distance. There were no distractors between arguments when the arguments were one step apart and three between them when they were four steps apart. The two different step sizes allowed us to separate the effects of distance from the effects of the number of intervening distractors. Arguments that were two and four steps apart with the small step sizes were separated by the same amount of absolute space (e.g., degrees of visual angle) as arguments that were one and two steps apart with the large spacing.

The effects of distraction could be predicted, in part, by current theories of visual search, such as Duncan and Humphreys's (1989) and Humphreys and Müller's (1993): Arguments should be hard to find when presented with similar distractors, so reaction time (RT) should be longer with distraction than without it. However, theories of visual search cannot account completely for the present data. They address how subjects find the arguments, but they say nothing about how they compute relations between them (spatial or otherwise). Theories of the apprehension of spatial relations must supplement theories of visual search to provide a complete account of the data.

Theories of Apprehension

Several recent theories have proposed hypotheses about how spatial relations are apprehended. The hypotheses have not always been intended as complete theories of apprehen-

Gordon D. Logan and Brian J. Compton, Department of Psychology, University of Illinois.

This research was supported by Grant BNS 91-09856 from the National Science Foundation. We are grateful to Libby Weber and Sheila Guerrero for helping to test the subjects.

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 via Internet to glogan@s.psych.uiuc.edu.

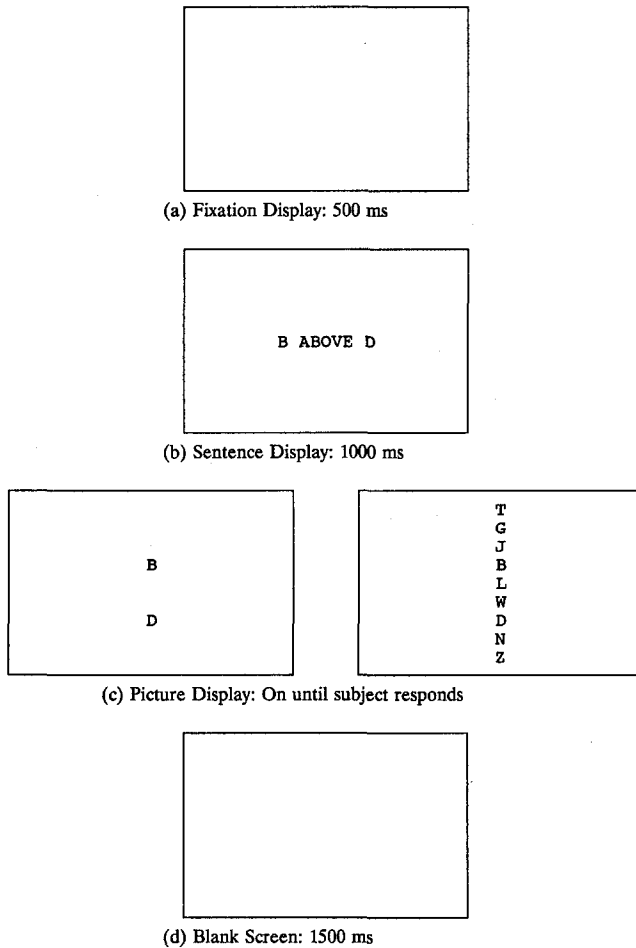


Figure 1. Events on a typical spatial relations trial.

sion. In many cases, the main focus of the theory was on issues other than apprehension, and the hypotheses about apprehension served mostly as theoretical contrasts. Kosslyn (1987), for example, was interested in the difference between "categorical" spatial relations like *above* and *below* that guide cognition and metric or "coordinate" relations that guide action. Hummel and Biederman (1992) hypothesized about apprehension of spatial relations as part of a broader theory of object recognition. Thus, tests of specific hypotheses about apprehension of spatial relations may not bear directly on the theories that spawned them. We consider revisions to the theories in light of the data in the General Discussion. Nevertheless, the specific hypotheses make definite predictions about the process of apprehension, and we designed our experiments to test them.

Logan and Sadler (in press)

Logan and Sadler (in press; see also Logan, 1994, 1995) present a theory of apprehension of spatial relations, derived

from linguistic analyses of spatial prepositions (Clark, 1973; Garnham, 1989; Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1992; Landau & Jackendoff, 1993; Levelt, 1984; Miller & Johnson-Laird, 1976; Talmy, 1983; Vandaloise, 1991), which involved four representations and four processes. The representations included a *perceptual representation*, which is a two- or three-dimensional analog array of objects and surfaces; a *conceptual representation*, which is a predicate expressing the relation between a pair of arguments (e.g., ABOVE [B, D]); a *reference frame*, which maps the conceptual representation onto the perceptual one and gives direction to perceptual space; and a *spatial template*, which represents the regions of perceptual space that contained good, acceptable, and bad examples of the predicated relation. The processes included *spatial indexing*, which establishes correspondence between perceptual objects or surfaces and the arguments of conceptual relations; *reference frame adjustment*, which involves aligning the reference frame with a perceptual object and setting its parameters (i.e., its origin, orientation, direction, and scale); *spatial template alignment*, which involves setting the origin, orientation, direction, and scale of the spatial template in congruence with the parameters of the reference frame; and *computation of goodness of fit*, which determines how well a given object exemplifies the predicated relation with respect to the object on which the reference frame and spatial template was centered. Following linguistic analyses, the theory distinguishes between the arguments of the conceptual relation. One is the *reference object* and the other is the *located object*. The relation specifies the position of the located object with respect to (a reference frame and a spatial template centered on) the reference object (see also Jackendoff & Landau, 1992; Landau & Jackendoff, 1993; Talmy, 1983).

In this theory, the process of apprehension involves (a) spatially indexing the reference object, (b) imposing a reference frame on the reference object, (c) aligning the relevant spatial template with the reference object, (d) spatially indexing the located object, and (e) computing the goodness of fit between the located object and the spatial template. According to this theory, distance between the arguments should have no effect on the time required for apprehension when there are no distractors because none of the steps are sensitive to distance. In particular, computing goodness of fit is assumed to be a spatially parallel process, occurring simultaneously at all points in the visual field. Distractors should have a large effect, however, and an effect of distance should emerge when distractors are present. The distractor effect should occur in Steps a and d, spatially indexing the reference and located objects. Distractors were similar to the arguments, so letters chosen (indexed) at random are more likely to be distractors than targets. Consequently, several iterations will be required to find a target (Duncan & Humphreys, 1989; Humphreys & Müller, 1993). The distance effect should emerge in Step d, spatially indexing the located object, assuming that subjects start at the reference object and move in a direction specified by the relation (up for *above*; down for *below*). Subjects should

iterate through the distractors until they find the located object.

Ullman (1984)

Ullman (1984) proposed a theory that described the transition between “early” and “late” visual processing that consisted of three kinds of representations and the processes that operated on them. The representations included a *perceptual representation*, which is produced by early visual processes (and corresponds to Marr’s, 1982, 2.5-d sketch); a *conceptual representation*, which is a propositional description of the three-dimensional structure of recognized objects; and an *incremental representation*, which is built by processes that operate on the perceptual representation and previously constructed parts of the incremental representation and bridges the gap between perceptual and conceptual representations. There were two kinds of processes, *local parallel processes*, which were obligatory and worked on visual input to produce the perceptual representation, and *serial visual routines*, which were voluntary and worked on the perceptual representation to produce the incremental and conceptual representations. Serial visual routines included *spatial indexing, marking, coloring, and tracing* (with a “mental cursor”).

Ullman (1984) was concerned directly with computing whether an object fell inside a closed contour and whether two objects fell on the same curved line, but his theory can be extended naturally to account for relations like *above* and *below* using serial visual routines, which move a visual marker continuously across space, like moving a mental cursor across a screen or propagating a colored region. *Above* and *below* could be computed by (a) spatially indexing the reference object, (b) moving a mental cursor (or coloring) in the direction specified by the relation (up for *above*; down for *below*) until an object was encountered, and (c) comparing that object with a description of the located object.

The key prediction of this theory is that the time required to compute *above* and *below* should increase monotonically with the distance between the arguments. If the mental cursor moves at a constant rate (Jolicoeur, Ullman, & MacKay, 1986), then the increase should be linear with distance. The increase should occur whether or not distractors are present. Distractors may increase the effect of distance because Step c will have to be iterated (i.e., comparing the indexed object with a description of the located object), but the distance effect should be significant even when no distractors are present.

Kosslyn, Chabris, Marsolek, and Koenig (1992)

Kosslyn, Chabris, Marsolek, and Koenig (1992; see also Kosslyn, 1987; O’Reilly, Kosslyn, Marsolek, & Chabris, 1990; Rueckl, Cave, & Kosslyn, 1989) proposed a connectionist model of apprehension of categorical spatial relations, like *above* and *below*, which they contrasted with apprehension of metric or coordinate spatial relations, like

those used to navigate between specific obstacles, and with apprehension of object identities. The main force of their work was in contrasting different kinds of apprehension to motivate hypotheses about differential localization of the underlying processes in the brain, but they made predictions about the effects of distance nevertheless.

The representations underlying apprehension in the Kosslyn et al. (1992) theory are like the spatial template representation in the Logan and Sadler (in press) theory in that they divide space into regions that contain good and bad examples of the relation. Kosslyn et al. (1992) go beyond Logan and Sadler in building a three-layer connectionist model of the apprehension process, which they trained to discriminate *above* from *below*. The key feature of their model that is relevant to the present article is their prediction that regions close to the reference object are harder to classify as *above* or *below* than regions more remote from it (see Kosslyn et al., 1992, Study 2). According to their theory, the time to apprehend *above* and *below* should decrease monotonically as the distance between the arguments increases, when no distractors are presented. The effect may be mitigated or reversed when distractors are present because subjects will have to iteratively compare distractors until they find the located object. If the time per iteration is less than the reduction per unit distance, the monotonic decrease in RT with distance will be diminished; if the time per iteration equals the reduction per unit distance, the effect of distance will disappear; and if the time per iteration is greater than the reduction per unit distance, a modest increase in RT with distance will appear when distractors are present.

Hummel and Biederman (1992)

Hummel and Biederman (1992) presented a connectionist model of object recognition in which objects were represented as *structured descriptions*, which consisted of elements (*geons*) and relations between them. Hummel and Biederman assumed that relations like *above* and *below* were represented as single-argument predicates, in contrast with a large linguistic literature that treats them as two-argument predicates (Clark, 1973; Garnham, 1989; Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1992; Landau & Jackendoff, 1993; Levelt, 1984; Miller & Johnson-Laird, 1976; Talmy, 1983; Vandaloise, 1991). In Hummel and Biederman’s theory, if a B were presented above a D, then *above* would be associated with (bound to) B, and *below* would be bound to D. The fact that B was above D and not above something else would not be represented explicitly. This scheme would work well if B and D were the only objects in the visual field, but as Hummel and Biederman (1992, p. 513) admitted, it would break down if there were distractors in the scene. B would be *below* one distractor, and D would be *above* another. B would be associated with both *above* and *below* and so would D. There is nothing in their model to resolve the ambiguity. Subjects should not be able to respond accurately when distractors are present, regardless of the distance between the arguments.

Predictions about distance are less clear than predictions about the effects of distractors. Different vertical positions are associated more or less strongly with *above* and *below*, and if position coding were sufficiently coarse, adjacent positions may activate both *above* and *below*. Thus, apprehension may become easier as distance increases, as in the Kosslyn et al. (1992) model. However, this prediction depends on aspects of the model that Hummel and Biederman (1992) did not specify, so it is less critical to the viability of their model than the prediction about distractors, which they did specify.

Previous Research

Several previous researchers varied the distance between elements and required subjects to report the spatial relations between them. Kosslyn et al. (1989) and Hellige and Michimata (1989) varied the distance between reference and located objects in a task that required report of *above* and *below*, but they did not report the effects of distance. Sergent (1991) varied distance in four experiments and found effects of distance in only one of them. That experiment presented a single dot in a circle and required subjects to report whether the dot was *above*, *below*, *left of*, or *right of* the center of the circle. Subjects were slower to report relations when the dots were closer to the center. However, there was no explicit reference object in that experiment. Subjects had to infer the center of the circle. In subsequent experiments, when Sergent included an explicit reference object, there was no effect of distance. Moreover, in the experiment that showed distance effects, dots were presented diagonally *above*, *below*, *left of*, or *right of* the implicit reference point rather than directly *above*, *below*, *left of*, or *right of* it, and diagonal regions are less acceptable examples than vertical or horizontal regions (Logan & Sadler, in press). In the experiments that showed no distance effects with *above* and *below* judgments, located objects and reference objects were aligned vertically and so were better examples of the relations.

Logan and Sadler (in press) reported a sentence-picture verification experiment, in which subjects saw a sentence describing a relation (*above* or *below*) between a dash and a plus (e.g., "DASH ABOVE PLUS") and then a picture that did or did not exemplify that relation. They varied the distance between the dash and the plus and found a non-monotonic relation. The shortest and longest distances were slower than the two intermediate distances.

Two factors limit the interpretation of the previous experiments. First, each of the experiments presented the same one (Sergent, 1991) or two (Hellige & Michimata, 1989; Kosslyn et al., 1989; Logan & Sadler, in press) elements repeatedly, so subjects could have responded to specific configurations of elements rather than computing the relations between them (Heathcote & Mewhort, 1993), or subjects could have responded accurately by noting the location of one of the elements and disregarding the other. Second, the displays in all the experiments were exposed briefly (for less than 200 ms) to prevent eye movements, and that may

have encouraged subjects to respond to configurations or to parts of the displays.

These are important considerations because linguistic analyses suggest that spatial relations depend on *compositional* representations that have internal structure, and their meaning depends jointly on the (separate) meanings of the parts and the relation between the parts (Barsalou, 1993; Fodor & Pylyshyn, 1988). Spatial predicates relate two objects, the located object and the reference object, and the meaning of the predicate depends on the identity of the objects as well as the relation between them. Thus, the representation of "B ABOVE D" has three parts, each with its own meaning. One represents B and refers to a specific percept (perceptual object) that corresponds to B, one represents D and refers to a specific percept that corresponds to D, and one refers to the relation between them (i.e., *above*). B and D must be represented separately because "B ABOVE D" and "D ABOVE B" mean different things. To test theories of apprehension adequately, experiments must require subjects to use compositional representations. The task must require subjects to identify each of the arguments separately and to compute the (spatial) relation between them. We designed our experiments with this constraint in mind.

The Present Experiments

We conducted three experiments using a variant of the sentence-picture comparison paradigm. Subjects saw a sentence describing the relation between specific letters and then a display that did (*true* displays) or did not (*false* displays) depict the relation specified in the sentence. We used the relations *above* and *below* and sampled arguments from a population of 20 letters. Our experiments forced subjects to use compositional representations by presenting different letters to be related on each trial. Thus, a configuration that would be useful on one trial would not be useful on another. Moreover, Experiment 2 used false displays that contained only one of the target arguments from the sentence (e.g., if the sentence was "B ABOVE D" the picture might contain a B above a C or a T above a D) as well as displays in which the letters appeared in the opposite relation (e.g., a D above a B), so that subjects would have to identify both letters to see whether the sentence was true.

The main manipulations in Experiments 1 and 2 were distance between the arguments of the relation and the presence versus absence of distractor letters around and between the arguments. Logan and Sadler's (in press) theory predicts no effect of distance when no distractors are present; the theory of Kosslyn et al. (1992) predicts a reduction in RT as distance increases; and the theory derived from Ullman (1984) predicts an increase in RT as distance increases. Hummel and Biederman (1992) theory predicts good performance when no distractors are present but chance-level performance when distractors appear.

Experiment 1

Method

Subjects. Four groups of 12 students from an introductory psychology class served as subjects. They received course credit for their participation.

Apparatus and stimuli. The stimuli were displayed in white on a dark background on Amdek 722 color monitors controlled by IBM AT computers. There were four computers, each facing a different wall of a large room so several subjects could be tested at once without disturbing each other. The events on a typical trial are depicted in Figure 1. Each trial consisted of (a) a fixation point presented in the center of the screen (row 13, column 39 of the standard IBM 24 × 80 text screen) for 500 ms; (b) a sentence display describing the relation between two letters (e.g., B ABOVE D) presented in row 13, columns 35–43 for 1,000 ms; (c) a picture display containing two or nine letters in column 39 (rows 9–13 for single-spaced displays; 5–21 for double-spaced displays) presented until subjects responded; and (d) a blank screen for a 1,500-ms intertrial interval. The letters were sampled randomly from 20 consonants (excluding A, E, I, O, U, and Y).

Each letter was 5 mm high and 3 mm wide. Viewed at a distance of 60 cm, the visual angle was $0.48^\circ \times 0.29^\circ$. Sentence displays were 5 mm high and 28 mm wide, corresponding to $0.48^\circ \times 2.67^\circ$ of visual angle. Nine-letter (distractor present) picture displays were 6.1 cm from top to bottom when single spaced (5.8° of visual angle) and 11.7 cm when double spaced (11.03°). Measured from center to center, Distances 1–4 were 0.7, 1.4, 2.1, and 2.8 cm, respectively, for single-spaced displays (0.67° , 1.34° , 2.00° , and 2.67° , respectively) and 1.4, 2.8, 4.2, and 5.6 cm, respectively, for double-spaced displays (1.34° , 2.67° , 4.00° , and 5.33° , respectively).

The letters could appear in nine logical positions. Targets never appeared in the top (Position 1) or bottom (Position 9) positions. The positions they occupied depended on the distance between them. There were six possible positions for Distance 1 (i.e., 2–3, 3–4, 4–5, 5–6, 6–7, and 7–8), each of which was sampled 20 times for a total of 120 examples. There were five possible positions for Distance 2 (i.e., 2–4, 3–5, 4–6, 5–7, and 6–8), each of which was sampled 24 times for a total of 120 examples. There were four possible positions for Distance 3 (i.e., 2–5, 3–6, 4–7, and 5–8), sampled 28 times for a total of 112 examples. There were three possible positions for Distance 4 (i.e., 2–6, 3–7, and 4–8), each sampled 40 times for a total of 120 examples. For each distance, half of the examples used the relation *above* and half used *below*, and half were true examples of the relation and half were false. There were 472 trials in total. The order of trials was randomized separately for each subject.

On each trial, 2 letters were sampled randomly from the set of 20 consonants to serve as targets. Assignment of targets to located and reference objects was random. In distractor displays, the remaining seven positions were filled with a random sampling of 7 different letters from the 18 letters that remained after targets were sampled. In no-distractor displays, the remaining seven positions were blank.

Subjects responded by pressing the “z” key or the “/” key, which were the leftmost and rightmost keys on the bottom row of the standard IBM AT keyboard. The display program measured RT in milliseconds and recorded which key was pressed.

Procedure. Subjects were assigned randomly to one of eight between-subject conditions, defined by the factorial combination of two 2-level independent variables (Distractors Present or Absent × Single-Spaced Vs. Double-Spaced Displays) and one

2-level counterbalancing variable (whether subjects pressed “yes” with their left or right hands). Subjects were instructed according to the condition they served in (i.e., they were told whether or not distractors would be present, they were told what the spacing would be, and they were told which keys to press for which responses). All subjects were told the sequence of events that would occur on each trial and were told that their task was to decide as quickly as possible whether the picture matched the sentence and to press the appropriate key to indicate their decision. Subjects were told to rest the index fingers of their left and right hands on the keys at all times during the experiment. They were allowed brief rests every 128 trials, at which time the computer displayed the message “TIME FOR A BREAK; PRESS ANY KEY TO RESUME” in the center of the screen (row 13 beginning at column 19).

Results

Mean RT and percentage correct were computed for each combination of relation (*above* vs. *below*), distance, truth, distractors (present vs. absent), and spacing (single vs. double) and were submitted to analyses of variance (ANOVAs). The mean RTs across subjects and relations are plotted in Figure 2. The accuracy data in each cell of the design are presented in Table 1.

When there were no distractors, RT was not affected by distance in either the single-space or double-space condition. RT was considerably slower when distractors were present and a strong effect of distance emerged. Accuracy, however, remained high. Spacing had little effect when no distractors were present, but it modulated the effect of distance when distractors were present. To see whether distance depended on the number of letters between arguments (relative distance) or the amount of space between them (absolute distance), we computed the slopes of linear regression functions relating RT to distance for true, distractor-present responses. For single-spaced displays, the slope was 54 ms/unit distance (i.e., distance = 1, 2, 3, or 4); for double-spaced displays, the slope was 40 ms/unit distance for absolute distance (distance = 2, 4, 6, or 8) and 79 ms/unit distance for relative distance (distance = 1, 2, 3, or 4). Both the number of letters and the amount of space between the arguments seemed to matter.

These conclusions were confirmed by 4 (distance) × 2 (spacing: single vs. double) × 2 (relation: *above* vs. *below*) × 2 (truth) × 2 (distractors present vs. absent) ANOVAs on the mean RTs and accuracy scores. In the RT ANOVA, there were significant main effects of distance, $F(3, 132) = 7.59, p < .01, MSE = 22,773.83$; relation (with *above* faster than *below*; $M_s = 1,387$ and $1,478$, respectively), $F(1, 44) = 82.07, p < .01, MSE = 19,241.43$; truth, $F(1, 44) = 79.96, p < .01, MSE = 59,286.54$; and distractor, $F(1, 44) = 137.04, p < .01, MSE = 1,823,541.90$. The main effect of spacing was not significant ($F < 1.0$). The (important) interaction between distance and distractors was significant, $F(3, 132) = 13.75, p < .01, MSE = 22,773.83$. Linear trend tests showed no effect of distance when there were no distractors ($F < 1.0$) but significant effects of distance when distractors were present, $F(1, 132) = 30.05, p < .01, MSE = 22,773.83$.

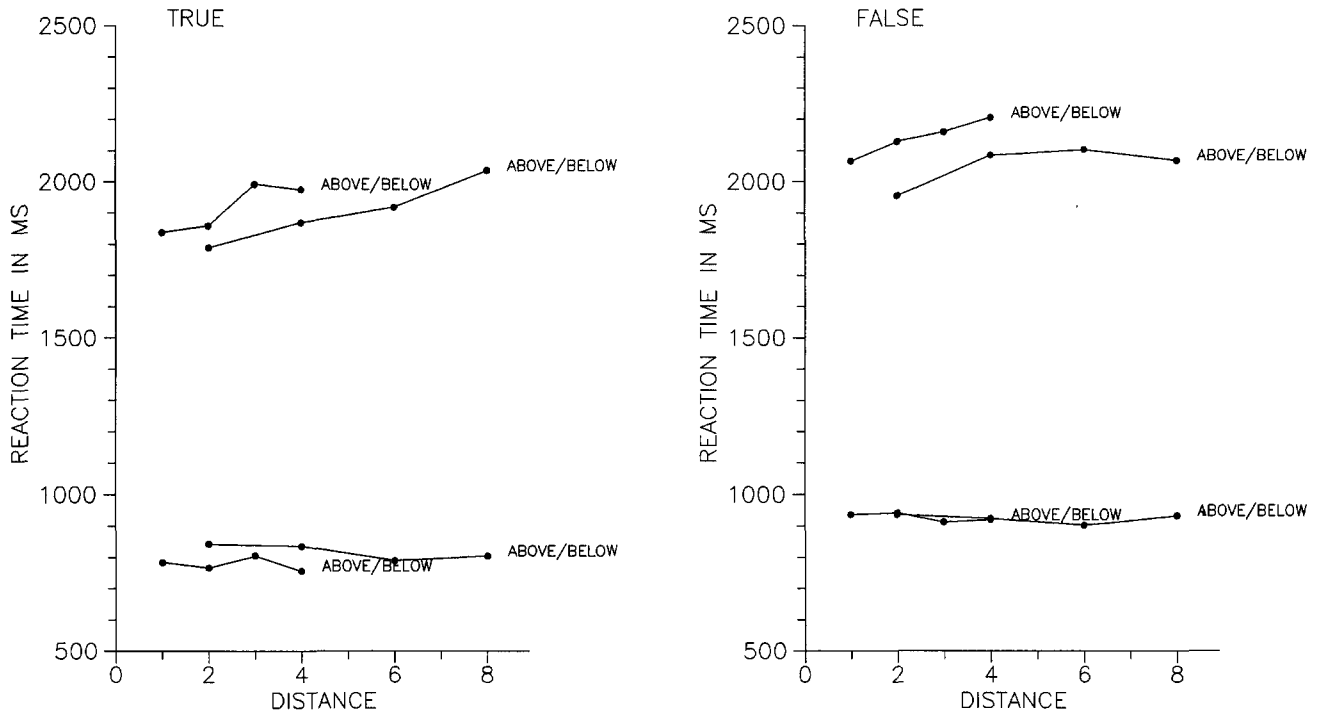


Figure 2. Mean reaction time averaged across relations (*above* and *below*) as a function of absolute distance for true responses (left panel) and false responses (right panel) in Experiment 1. (The top two lines in each panel represent distractors present; the bottom two lines represent no distractors. The long lines represent double spacing; the short lines represent single spacing.)

The distance effect with no distractors is important theoretically, so we analyzed it in more detail. Linear trend tests on the data from true responses showed a significant effect for *above*, double-spaced displays, $F(1, 132) = 6.34, p < .05, MSE = 15,899.02$, but not for *below*, double-spaced displays, $F(1, 132) = 2.03, MSE = 15,899.02$, or for *above* and *below*, single-spaced displays ($F_s < 1.0$). Fisher's least significant difference (LSD) test showed no significant differences ($p < .05$) between any of the distances within any of these conditions. The significant reduction in RT with distance for *above*, double-spaced displays is hard to interpret because there was no effect of distance in single-spaced displays, where distances were smaller and, thus, should have been more likely to show an effect.

In addition to these effects, there were significant interactions between relation and truth, $F(1, 44) = 43.31, p < .01, MSE = 29,092.41$; distractors, relation, and truth, $F(1, 44) = 19.76, p < .01, MSE = 29,092.41$; distractors, spacing, relation, and truth, $F(1, 44) = 9.40, p < .01, MSE = 29,092.41$; relation, truth, and distance, $F(3, 132) = 4.36, p < .01, MSE = 15,899.02$; and distractors, relation, truth, and distance, $F(3, 132) = 3.75, p < .05, MSE = 15,899.02$. These interactions did not compromise the main effects of interest.

In the accuracy ANOVA, there was a significant main effect of relation, $F(1, 44) = 36.69, p < .01, MSE = 40.70$, and significant interactions between distractor and truth,

$F(1, 44) = 9.86, p < .01, MSE = 36.95$, and distractor, truth, and distance, $F(3, 132) = 3.05, p < .05, MSE = 12.91$. These effects did not compromise the interpretation of the RT results.

Table 1
Mean Accuracy (Percentage Correct) in Experiment 1 as a Function of Relation, Truth, Distance, Spacing, and the Presence of Distractors

Spacing	True				False			
	1	2	3	4	1	2	3	4
No distractors								
<i>Above</i>								
Single	98	97	98	98	96	97	97	96
Double	98	98	97	99	96	96	97	93
<i>Below</i>								
Single	93	94	94	94	93	93	91	92
Double	94	94	94	95	95	95	93	94
Distractors present								
<i>Above</i>								
Single	96	93	96	94	94	97	96	96
Double	94	96	96	94	95	97	96	98
<i>Below</i>								
Single	93	91	93	91	94	94	92	94
Double	91	94	92	95	91	95	96	96

Note. 1-4 refer to distance.

Discussion

The data were consistent with Logan and Sadler's (in press) theory of apprehension. RT was not affected by distance when there were no distractors, and it was strongly affected by distance when distractors were present. The data were inconsistent with a theory based on Ullman's (1984) idea of serial visual routines because RT did not increase monotonically with distance when no distractors were present. The data were also inconsistent with the theory of Kosslyn et al. (1992) because RT did not decrease monotonically with distance when no distractors were present. Finally, the data were inconsistent with Hummel and Biederman's (1992) theory in that subjects performed accurately when distractors were present. Distractors slowed their RTs and produced a strong effect of distance, but they did not lead to the confusion that Hummel and Biederman's theory would predict.

Experiment 2

Experiment 1 showed no effect of distance when no distractors were presented and a strong effect of distance when distractors were presented. However, in Experiment 1, displays exemplifying false relations always contained targets in the opposite relation to the one specified in the sentence, and this may have led subjects to adopt special strategies to exploit the contingency (e.g., responding after they find the first letter mentioned in the sentence). Experiment 2 was conducted to control for this possibility by introducing false displays in which only one target was present. In addition, it introduced a new relation, *and* (logical conjunction), that required subjects to detect the presence of two targets without specifying the spatial relation between them. *And* is interesting because it should be simpler to apprehend than *above* and *below*. It requires spatial indexing of the arguments, but it does not require aligning a reference frame and fitting a spatial template. Thus, it should be apprehended more quickly than *above* and *below*, at least when no distractors are present to complicate the picture.

Method

Subjects. The subjects were 64 students from an introductory psychology class who served for course credit. None had served in Experiment 1.

Apparatus and stimuli. Experiment 2 used the same apparatus as Experiment 1; however, the stimuli were slightly different. The same letters and the same positions were used. The combinations were different, however, because a new relation was added (*and*) and false displays for *above* and *below* sometimes contained only one of the two arguments presented in the sentence. As in Experiment 1, there were six, five, four, and three possible positions for Distances 1–4, respectively. There were 36, 30, 24, and 36 examples of each distance (1–4, respectively) for *above* and *below* and 48, 40, 32, and 48 examples of each distance (1–4, respectively) for *and*, for a total of 420 trials. Half of the trials displayed true examples and half displayed false examples.

There were three kinds of false examples for *above* and *below*. One kind presented the arguments in the order opposite to the one described in the sentence, as in Experiment 1 (e.g., if the sentence was "B ABOVE D," the display would contain a B below a D). Another presented the letter corresponding to the reference object but omitted the letter corresponding to the located object, substituting a randomly selected letter from the letters that remained after sampling targets and distractors (e.g., if the sentence was "B ABOVE D," the display might contain a C above a D). The third presented the letter corresponding to the located object but omitted the reference object, substituting a randomly selected letter (e.g., if the sentence was "B ABOVE D," the display might contain a B above a C). Each kind of false example appeared equally often in the experiment.

There were four kinds of false examples for *and*. In the first, the first argument appeared on top and a different letter appeared on the bottom (e.g., if the sentence was "B AND D," the display might contain a B above a C). In the second, the first argument appeared on the bottom and a different letter appeared on top (e.g., if the sentence was "B AND D," the display might contain a B below a C). In the third, the second argument appeared on top and a different letter appeared on the bottom (e.g., if the sentence was "B and D," the display might contain a D above a C). In the fourth, the second argument appeared on the bottom and a different letter appeared on the top (e.g., if the sentence was "B AND D," the display might contain a C above a D). Each kind of false example appeared equally often in the experiment.

There were two kinds of true examples for *and*. In one, the first argument appeared on top and the second on the bottom, and in the other, the first argument appeared on the bottom and the second appeared on top. Each kind of true example appeared equally often.

The sentences for the *and* relation included two spaces between each letter and the word so that the spatial extent of the sentence was the same as for *above* and *below* (i.e., nine characters).

Procedure. The procedure was the same as in Experiment 1, except that there were 420 trials and breaks every 84 trials.

Results

Mean RT and percentage correct were computed for each combination of relation (*above*, *below*, *and*), distance, truth, distractors (present vs. absent), and spacing (single vs. double) and submitted to ANOVAs. The mean RTs across subjects and relations are plotted in Figure 3. The accuracy data in each cell of the design are presented in Table 2.

As in Experiment 1, there was little effect of distance when no distractors were present. When distractors were present, however, RT increased substantially, and a strong effect of distance emerged. Accuracy remained high, as it did in Experiment 1. Again, spacing had little effect on performance when no distractors were present, but it modulated the effect of distance when distractors were present. We fitted regression lines to the distance effects for true distractor-present responses to see whether absolute or relative distance mattered. For single-spaced displays, the slope for *above* and *below* was 87 ms/unit distance (distance = 1, 2, 3, or 4); for double-spaced displays, the slope for *above* and *below* was 80 ms/unit distance for relative distance (distance = 1, 2, 3, or 4) and 40 ms/unit distance for absolute distance (distance = 2, 4, 6, or 8). For single-spaced displays, the slope for *and* was 131 ms/unit distance;

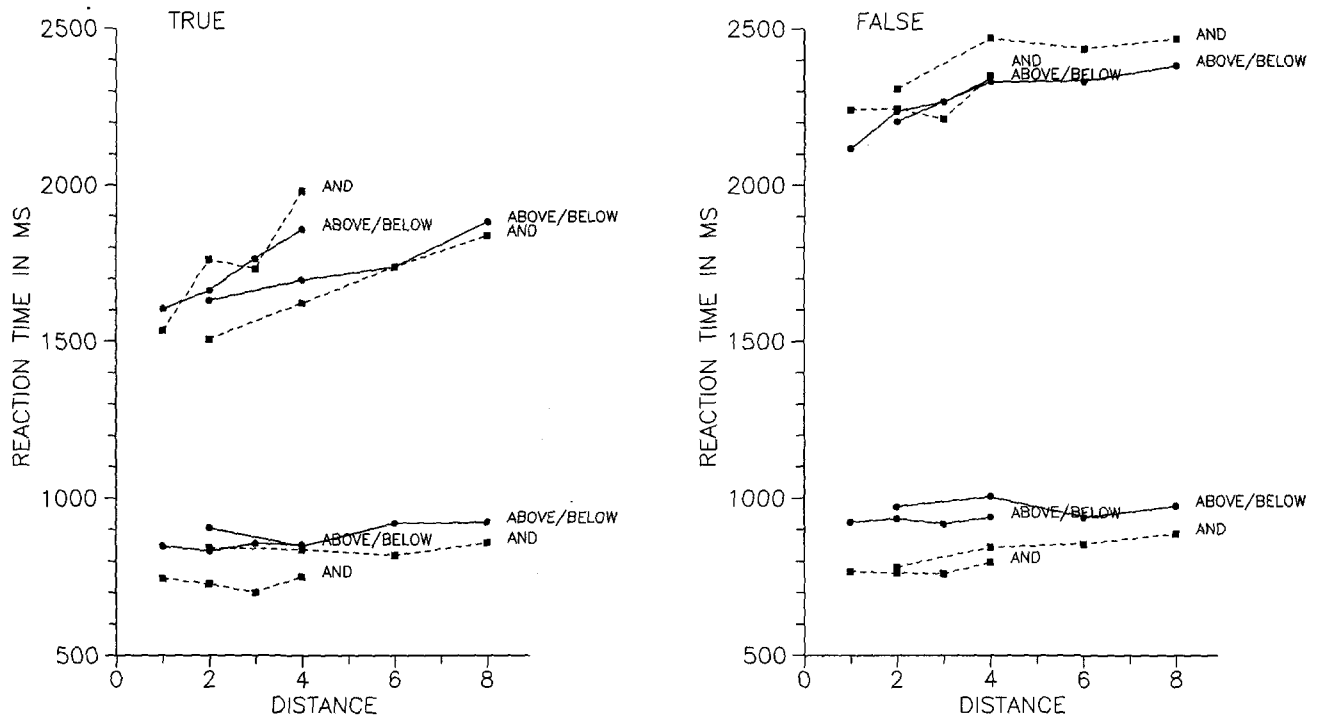


Figure 3. Mean reaction time for *and* (broken lines) and the average of *above* and *below* (solid lines) as a function of absolute distance for true responses (left panel) and false responses (right panel) in Experiment 2. (The top three lines in each panel represent distractors present; the bottom three lines represent no distractors. The long lines represent double spacing; the short lines represent single spacing.)

for double-spaced displays, the slope was 112 ms/unit relative distance and 56 ms/unit absolute distance. In this experiment, the number of letters between the arguments was the critical variable; the amount of space was not.

The difference between *and* on the one hand and *above* and *below* on the other depended on the presence of distractors. When there were no distractors, *and* was faster than *above* and *above* was faster than *below* ($M = 796, 868$, and 956 ms, respectively), which is consistent with the idea that *and* is easier to apprehend. When distractors were present, the differences were diminished and *and* was slower than *above* but faster than *below* ($M = 2,029, 1,955$, and $2,052$ ms, respectively).

These conclusions were supported in 4 (distance) \times 2 (spacing: single vs. double) \times 3 (relation: *above* vs. *below* vs. *and*) \times 2 (truth) \times 2 (distractors present vs. absent) ANOVAs on the mean RTs and accuracy scores. In the RT ANOVA, there were significant main effects of distance, $F(3, 180) = 53.72, p < .01, MSE = 20,475.58$; relation, $F(2, 60) = 27.27, p < .01, MSE = 52,951.23$; truth, $F(1, 60) = 207.85, p < .01, MSE = 185,550.08$; and distractor, $F(1, 60) = 212.67, p < .01, MSE = 2,338,861.50$. The crucial interaction between distance and distractor was significant, $F(3, 180) = 38.58, p < .01, MSE = 20,475.58$. Linear trend tests showed a significant effect of distance when distractors were present, $F(1, 180) = 88.38, p < .01, MSE = 20,475.58$, but no significant effect when no distractors were present ($F < 1.0$).

As in Experiment 1, we analyzed the no-distractor distance effect in more detail, focusing on true responses to *above* and *below* relations. Linear trend tests showed a significant increase in RT with distance for *above*, double-spaced displays, $F(1, 360) = 7.26, p < .01, MSE = 21,329.56$, largely because of an unusually long RT at Distance 4 ($M = 885$ ms) compared with the three shorter distances ($M_s = 813, 817$, and 821 ms, respectively, for Distances 1–3). None of the other linear trends were significant (all $F_s < 1.0$). Fisher's LSD test showed significantly longer RTs for Distance 4 than for Distances 1–3 for *above*, double-spaced displays ($p < .05$) and significantly faster RTs for Distance 2 than for Distances 1, 3, and 4 for *below*, double-spaced displays ($p < .05$). No other within-condition differences were significant.

In addition to these effects, there were significant interactions between relation and truth, $F(2, 120) = 19.96, p < .01, MSE = 28,474.58$; distractors, relation, and truth, $F(2, 120) = 10.32, p < .01, MSE = 28,474.58$; distractors, spacing, relation, and truth, $F(2, 120) = 5.44, p < .01, MSE = 28,474.58$; truth and distance, $F(3, 180) = 8.04, p < .01, MSE = 15,183.18$; truth, distance, and distractors, $F(3, 180) = 5.17, p < .01, MSE = 15,183.18$; and distractors, relation, truth, and distance, $F(6, 360) = 2.27, p < .05, MSE = 21,329.56$. These interactions did not compromise the interpretation of the main results.

In the accuracy ANOVA, there were significant main effects of relation, $F(2, 120) = 22.85, p < .01, MSE =$

Table 2
Mean Accuracy (Percentage Correct) in Experiment 2 as a Function of Relation, Truth, Distance, Spacing, and the Presence of Distractors

Spacing	True				False			
	1	2	3	4	1	2	3	4
No distractors								
<i>Above</i>								
Single	97	98	97	97	95	97	95	95
Double	98	97	97	98	96	98	97	97
<i>Below</i>								
Single	92	93	94	94	93	90	94	93
Double	96	95	96	96	94	92	96	94
<i>And</i>								
Single	97	97	99	97	97	98	98	97
Double	97	98	98	96	98	100	100	98
Distractors present								
<i>Above</i>								
Single	95	96	94	96	96	97	94	97
Double	95	97	95	94	96	96	98	98
<i>Below</i>								
Single	93	90	94	93	95	96	95	96
Double	94	95	95	94	95	97	98	97
<i>And</i>								
Single	95	93	91	91	98	98	98	98
Double	96	93	94	94	98	99	97	97

Note. 1–4 refer to distance.

36.13, and truth, $F(1, 60) = 13.31$, $p < .01$, $MSE = 40.85$, and significant interactions between relation and distractors, $F(2, 120) = 9.39$, $p < .01$, $MSE = 36.13$; truth and distractors, $F(1, 60) = 25.22$, $p < .01$, $MSE = 40.85$; relation and truth, $F(2, 120) = 6.54$, $p < .01$, $MSE = 38.21$; and relation and distractors, $F(6, 360) = 2.45$, $p < .05$, $MSE = 25.96$. These effects did not compromise the interpretation of the RT results.

Finally, we analyzed the false responses in more detail to evaluate the effects of the different kinds of false displays. When no distractors were present, responses to false *above* and *below* relations were faster when the first ($M = 857$ ms) and second ($M = 899$ ms) arguments were missing than when the arguments appeared in the wrong order ($M = 1,111$ ms). This suggests that subjects first found the arguments and then computed the relation between them. When distractors were present, the pattern reversed. Subjects were slower when the first and second arguments were missing than when the arguments appeared in the wrong order ($M_s = 2,407$, $2,342$, and $2,081$ ms, respectively). This also suggests that subjects first searched for the two arguments and then computed the relation between them. However, search for the arguments was faster when both were present because only one to three letters would have to be examined to find the second argument after finding the first, whereas if one argument was missing, as many as eight letters might have to be examined before responding false.

These conclusions were confirmed in 4 (distance) \times 2 (spacing) \times 2 (relation: *above* vs. *below*) \times 3 (type of false display) \times 2 (distractors present vs. absent) ANOVAs on the mean RTs and accuracy scores. In the RT ANOVA, the

interaction between type of false display and distractor was significant, $F(2, 120) = 66.18$, $p < .01$, $MSE = 184,707.19$, reflecting the trend described above. In addition, there were significant main effects of distance, $F(3, 180) = 14.04$, $p < .01$, $MSE = 54,943.86$; relation, $F(1, 60) = 6.05$, $p < .01$, $MSE = 56,748.70$; and distractor presence, $F(1, 60) = 196.23$, $p < .01$, $MSE = 3,413,312.10$. Distance interacted significantly with distractor presence, $F(3, 180) = 13.85$, $p < .01$, $MSE = 54,943.86$; type of false display, $F(6, 360) = 2.04$, $p < .06$, $MSE = 75,824.69$; distractor presence and type of false display, $F(6, 360) = 2.35$, $p < .05$, $MSE = 75,824.69$; and spacing, distractor presence, and type of false display, $F(6, 360) = 2.77$, $p < .05$, $MSE = 75,824.69$, reflecting a distance effect when both arguments were present in the wrong relation but not when either argument was missing. There were also significant interactions between type of false display and relation, $F(2, 120) = 17.23$, $p < .01$, $MSE = 63,741.80$; distractor presence, type, and relation, $F(2, 120) = 5.04$, $p < .01$, $MSE = 63,741.80$; type, relation, and distance, $F(6, 360) = 3.70$, $p < .01$, $MSE = 67,096.00$; distractor presence, type, relation, and distance, $F(6, 360) = 2.87$, $p < .05$, $MSE = 67,096.00$; and distractor presence, spacing, and relation, $F(1, 60) = 4.78$, $p < .05$, $MSE = 56,748.70$.

In the accuracy ANOVA, there was a significant effect of type of false display, $F(2, 120) = 42.44$, $p < .01$, $MSE = 142.39$, reflecting lower accuracy when the arguments were in the wrong order ($M = 92.0$) than when the first or second arguments were missing ($M_s = 98.2$ and 96.6 , respectively). In addition, there was a significant main effect of relation, $F(1, 60) = 11.86$, $p < .01$, $MSE = 98.72$; and significant interactions between relation and distractor presence, $F(1, 60) = 7.48$, $p < .01$, $MSE = 98.72$; type of false display and relation, $F(2, 120) = 9.86$, $p < .01$, $MSE = 58.29$; and distractor presence, type, and relation, $F(2, 120) = 4.89$, $p < .01$, $MSE = 58.29$.

Discussion

The results replicated Experiment 1 and confirmed the conclusions drawn there. The null effect of distance with no distractors confirms Logan and Sadler's (in press) theory and disconfirms the theory of Kosslyn et al. (1992) and the theory derived from Ullman (1984), which predicted monotonic decreases and increases in RT with distance, respectively. When distractors were present, distance had a strong effect on RT but not on accuracy. This confirms Logan and Sadler's theory and disconfirms Hummel and Biederman's (1992) theory, which predicted confusion and near-chance accuracy when distractors were present.

The results are particularly important, in contrast with Experiment 1 and previous experiments in the literature (Hellige & Michimata, 1989; Kosslyn et al., 1989; Logan & Sadler, in press), because the procedure required subjects to identify both of the arguments of the relation before responding. The false trials, on which only one argument was presented, did not allow subjects to respond accurately after identifying only one letter. Moreover, the random sampling

of targets from a set of 20 letters would have precluded responding based on learned configurations, as it would have in Experiment 1 as well (Heathcote & Mewhort, 1993). Thus, the present experiments provide the clearest test so far of the apprehension of compositional spatial relations.

It is interesting as well that *and* was processed faster than *above* and *below* when no distractors were present. This confirms the hypothesis that apprehension of *above* and *below* requires more than simply identifying the arguments of the relation and suggests that the hypothesized additional processes—reference frame alignment, fitting spatial templates, and computing goodness of fit—are psychologically real in that they take measureable time.

Experiment 3

Experiment 3 assessed the effects of attention on distraction. The targets were colored white and the distractors were colored red. The purpose was to see whether theories of visual search could be combined with theories of the apprehension of spatial relations. Theories of visual search predict that distractors should have little effect when they are sufficiently different from the targets (Duncan & Humphreys, 1989; Humphreys & Müller, 1993). If theories of visual search describe the processes by which subjects find the targets and theories of apprehension describe the subsequent processes by which they compute spatial relations, then subjects should be able to filter out the distractors, and their performance should resemble that of the no-distractor conditions of Experiments 1 and 2. However, if theories of visual search do not apply to the apprehension of spatial relations and subjects cannot filter out dissimilar distractors, then their performance should resemble that of the distractor-present conditions of Experiments 1 and 2.

Method

Subjects. The subjects were 32 students from an introductory psychology class who received course credit for participating. They were screened for color blindness with the Ishihara (1987) test.

Apparatus and stimuli. The apparatus and stimuli were the same as those used in Experiment 1. The only difference was that the targets in the picture display remained white (IBM 15), while the distractors were colored red (IBM 12). Distractors were always presented.

Procedure. The procedure was the same as in Experiment 1, except that subjects were told about the color cue. There were 472 trials, and breaks were given every 118 trials. Pictures exemplifying false relations had arguments in the opposite relation to that described in the sentence (e.g., B ABOVE D followed by a B below a D).

Results

Mean RT and percentage correct were computed for each combination of relation (*above* vs. *below*), distance, truth, and spacing (single vs. double), and submitted to ANOVAs.

The mean RTs across subjects and relations are plotted in Figure 4. The accuracy data in each cell of the design are presented in Table 3.

RT was fast and not affected much by distance, like the no-distractor conditions and unlike the distractor-present conditions in Experiments 1 and 2. Mean RT was 891 ms, very close to the mean value for *above* and *below* with no distractors in Experiments 1 and 2 ($M = 862$ and 912 ms, respectively). Subjects had little difficulty screening out the distractors.

These conclusions were confirmed in 4 (distance) \times 2 (relation) \times 2 (truth) \times 2 (spacing) ANOVAs on the mean RTs and accuracy scores. In the RT ANOVA, there were significant main effects of distance, $F(3, 90) = 14.70$, $p < .01$, $MSE = 5,323.26$; relation, $F(1, 30) = 60.89$, $p < .01$, $MSE = 14,044.56$; and truth, $F(1, 30) = 123.63$, $p < .01$, $MSE = 16,337.92$, but not spacing, $F(1, 30) = 1.02$, $MSE = 865,480.13$. A linear trend test performed on the distance effect was significant, $F(1, 90) = 10.40$, $p < .01$, $MSE = 5,323.36$, revealing reduction in RT as distance increased. We tested the linear trend for true responses in each combination of relation and spacing conditions and found a significant reduction in RT with distance in each case, $F_s(1, 90) = 10.21, 5.00, 7.15$, and 10.42 , for *above* single-spaced, *below* single-spaced, *above* double-spaced, and *below* double-spaced, respectively (all p s $< .05$ and all MSE s = $5,659.83$). Fisher's LSD test showed that RTs were significantly longer ($p < .05$) to Distance 1 than to Distances 2–4 for all conditions except below, single-spaced, in which Distance 1 was significantly longer than Distance 4 ($p < .05$).

In addition to these effects, there were significant interactions between spacing and relation, $F(1, 30) = 9.45$, $p < .01$, $MSE = 14,044.56$, and between relation and truth, $F(1, 30) = 18.11$, $p < .01$, $MSE = 7,927.10$. These effects did not compromise the interpretation of the main results.

In the accuracy ANOVA, only the effect of relation was significant, $F(1, 30) = 29.84$, $p < .01$, $MSE = 12.90$.

Discussion

The data resembled the no-distractor conditions of Experiments 1 and 2 more than the distractor-present conditions. Responses were fast and accurate, and RT did not increase with distance. These data demonstrate that an important principle that governs visual search—target-distractor similarity—also governs the apprehension of spatial relations. They motivate the development of theories that combine processes of visual search and visual attention with processes involved in apprehending spatial relations to provide a more complete account of visual cognition.

RTs tended to decrease as distance increased. This is inconsistent with the theory derived from Ullman (1984), which predicts a monotonic increase with distance. It appears to be inconsistent with Logan and Sadler's (in press) theory, which predicts no effect of distance, and consistent with the theory of Kosslyn et al. (1992), which predicts a monotonic decrease with distance. However, the distance

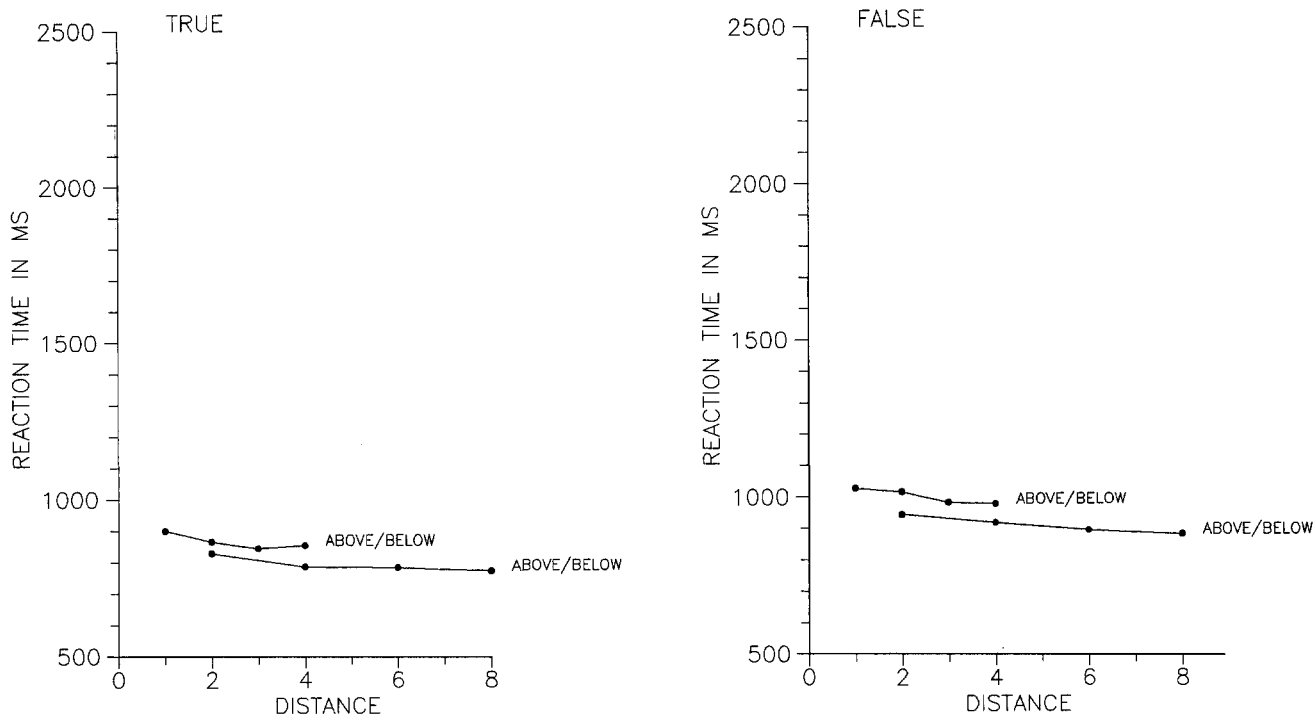


Figure 4. Mean reaction time averaged across relations (*above* and *below*) as a function of absolute distance for true responses (left panel) and false responses (right panel) in Experiment 3. (The long lines represent double spacing; the short lines represent single spacing.)

effects may have to do with processes involved in screening out the distractors rather than the processes involved in computing the relation between the targets. Eriksen and Webb (1989) cued two positions in a multielement display in a task that did not require computing spatial relations between the cued positions and found that RT decreased as the distance between the cues increased. Perhaps it is easier to ignore distractors if several of them can be grouped together (Duncan & Humphreys, 1989; Humphreys & Müller, 1993). Alternatively, subjects may be set to look for two distinct targets, and targets adjacent to each other may be hard to deal with because they appear to be one object.

General Discussion

In Experiments 1 and 2, there were no effects of distance when no distractors were present. This confirms the predictions of Logan and Sadler's (in press) theory and disconfirms the predictions of Kosslyn et al.'s (1992) theory and the theory based on Ullman (1984). In Experiments 1 and 2, distractors slowed RT but did not reduce accuracy to chance levels, disconfirming predictions of Hummel and Biederman's (1992) theory. In Experiment 3, distractor effects disappeared when targets and distractors were colored differently, which suggests that distraction effects are not obligatory but, rather, can be modulated by attention (Duncan & Humphreys, 1989; Humphreys & Müller, 1993).

Do these results disconfirm the theories? Not exactly. They disconfirm hypotheses about how spatial relations are

apprehended that were embedded in the theories, but they need not disconfirm the theories themselves. In the remainder of this section, we discuss modifications to the theories that would account for the results and the aspects of Logan and Sadler's (in press) theory that led to its superior performance.

Ullman (1984)

The serial visual routine model, based on Ullman's (1984) theory, predicted a monotonic increase in RT as distance increased. This prediction stemmed from the sequential, incremental nature of the movement of a mental cursor (e.g., Jolicoeur et al., 1986) or a "coloring" operation.

Table 3
Mean Accuracy (Percentage Correct) in Experiment 3 as a Function of Relation, Truth, Distance, and Spacing

Spacing	True				False			
	1	2	3	4	1	2	3	4
<i>Above</i>								
Single	99	99	98	99	98	98	99	98
Double	98	98	98	98	95	96	96	98
<i>Below</i>								
Single	97	97	95	97	97	97	96	96
Double	96	97	95	96	96	95	96	96

Note. 1-4 refer to distance.

To account for the observed results, these operations would have to be replaced by something that operated in parallel, independent of distance. The spatial templates in Logan and Sadler's (in press) theory would be a reasonable substitute. They are analogous to a coloring operation, where "darker" colors represent better examples of the relation in question, and they are applied in parallel to the whole visual field.

The Ullman-based theory would have to incorporate a reference frame to account for the results, even if the serial nature of the visual routine did not change. The reference frame is necessary because it tells the routine where to start (i.e., at the reference object) and it tells the routine in which direction it should travel (i.e., it gives direction to perceptual space). Reference frames play these roles in the Logan and Sadler (in press) theory.

The idea of spatially indexing the arguments of the relation, which plays a central role in Logan and Sadler's (in press) theory, was borrowed largely from Ullman (1984; see also Pylyshyn, 1989). Spatial indexing is important because (a) it allows the system to distinguish between the reference object and the located object (A ABOVE B is not the same as B ABOVE A), (b) it accounts for the linear effect of distance when distractors were present in Experiments 1 and 2 (i.e., subjects selected and examined distractors one-by-one until the target was found), and (c) it accounts for the absence of distance effects when distractors were present but targets were cued in Experiment 3 (i.e., subjects indexed or marked the cued targets without having to consider the distractors).

Kosslyn et al. (1992)

The Kosslyn et al. (1992) theory predicted a monotonic decrease in RT with distance, which was not observed. This prediction was built into their model by the training regime. The model had a "retina" consisting of 28 input units, and the reference object—a bar—occupied 4 of them. The bar appeared in five different positions: (a) the central four positions, (b) four positions centered two units above or below the central position, and (c) four positions centered four units above or below the central position. The model was trained on input patterns in which "dots" appeared above or below the bar. Because of the variation in the location of the bar, units near the center would have fewer training examples than units remote from it, and some units near the bar would be associated with examples of both *above* and *below*. Both of these effects would make the model less responsive to inputs near the bar than to inputs far from it, and that is what underlies the predicted distance effects.

We suggest that the model would predict no distance effects if the position of the bar was held constant during training. If the bar was constant, all input units above and below it could be trained equally often and no units would be associated with both relations. The model then would resemble the spatial templates in Logan and Sadler's (in press) theory and, therefore, account for the data in the present experiments.

It is not clear why Kosslyn et al. (1992) chose to let the position of the bar vary in their simulations. They might have done so to account for the distance effects that were observed in Sergent's (1991) first experiment, in which the position of the reference object was uncertain (because no reference object was presented). However, that choice leads the model to make the wrong predictions about Sergent's other experiments, in which a salient reference object was present, and in the present experiments, in which a salient reference object was also present.

Strictly speaking, the Kosslyn et al. (1992) model cannot account for the present data (nor the data of Logan, 1995) because the division between *above* and *below* occurs at a (relatively) fixed locus on its retina. In the present experiments, however, some examples of *above* occurred when both arguments appeared near the bottom of the screen and some examples of *below* occurred when both arguments appeared near the top of the screen. The division between *above* and *below* must move flexibly around the visual field. In Logan and Sadler's (in press) theory, the reference frame is responsible for this flexibility. Its origin, orientation, direction, and scale can be adjusted at will, centering the division between *above* and *below* on the reference object wherever it may appear. If reference frame computations and spatial indexing were added to the Kosslyn et al. (1992) theory, it would resemble Logan and Sadler's theory and make the same predictions. There appears to be nothing inherent in the Kosslyn et al. theory to prevent these modifications.

Hummel and Biederman (1992)

Hummel and Biederman's (1992) theory predicts confusion when more than two objects are present in the visual field because it assumes that spatial relations are represented by one-argument predicates. *Above* and *below* get bound to single objects by a bottom-up process. There is nothing to prevent them from being bound to the same object and when that happens, there is ambiguity and confusion. An obvious way to correct this problem would be to represent spatial relations with two- or three-argument predicates, following what everyone does in the linguistic and psycholinguistic literature (e.g., Clark, 1973; Garnham, 1989; Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1992; Landau & Jackendoff, 1993; Levelt, 1984; Miller & Johnson-Laird, 1976; Talmy, 1983; Vandaloise, 1991). It is not clear how this can be done in the Hummel and Biederman model, however, because their model binds properties and relations to objects by synchronizing firing in all of the units that represent them. By definition, things that fire synchronously are part of one object, and things that fire asynchronously are part of separate objects. The problem with two- and three-argument predicates is that the arguments refer to different objects, and the distinction between objects would be lost if they fired synchronously.

Put differently, linguistic analyses assume that spatial relations are compositional (Barsalou, 1993; Fodor & Pylyshyn, 1988), in that their meaning depends on the (separate)

meanings of their parts and on the relations between the parts. The representations in Hummel and Biederman's (1992) theory are not compositional because everything that is bound together by synchronous firing is considered to be a single perceptual object. It is not clear how to make compositional representations with their method of binding.

Another way to fix the problem of ambiguity and resolve the ensuing confusion would be to restrict the focus of attention so that it excludes the distractors. If attention could be focused so that only two objects impinge on the model at one time, then *above* and *below* could be assigned unambiguously. However, it is not clear whether attention can be directed simultaneously to two different locations; the issue is currently controversial (see, e.g., Castillo & Umiltà, 1992; Eriksen & St. James, 1986; Eriksen & Webb, 1989; Eriksen & Yeh, 1985; Posner, 1980; Shaw, 1978). The evidence that attention can be split comes mostly from experiments that cue the locations of single items (e.g., Castillo & Umiltà, 1992). Experiments that present several items simultaneously often suggest that attention cannot be split (e.g., Eriksen & St. James, 1986; Eriksen & Webb, 1989; Eriksen & Yeh, 1985). Unfortunately, displays of several items are the ones that cause ambiguity and confusion in the Hummel and Biederman (1992) theory.

In a way, the criticism of the Hummel and Biederman (1992) theory is unfair. The problems with their theory arise from their assumptions about the mechanism underlying the spatial indexing process, which the other theories simply assume without explaining. On the positive side, their theory does explain how different properties of an object are bound together, and that is a major function of spatial indexing (Kahneman, Treisman, & Gibbs, 1992). Moreover, their theory provides for some forms of reference-frame computation, in that it extracts the major axes of perceptual objects, which correspond to the major axes of object-centered reference frames (Marr & Nishihara, 1978). Finally, their theory was intended to address within-object spatial relations, whereas the other theories were intended to address between-object spatial relations, and the semantics may be different (Jackendoff & Landau, 1992; Landau & Jackendoff, 1993). Between-object and within-object spatial relations differ psychophysically (Baylis & Driver, 1993; Logan, 1975), neurophysiologically (Ungerleider & Mishkin, 1982), and computationally (Jacobs, Jordan, & Barto, 1991; Rueckl et al., 1989). Hummel and Biederman's theory does a lot toward increasing our understanding of object recognition and the within-object spatial relations on which it relies. It may be too much to expect it to explain between-object spatial relations as well.

Conclusions

Experiments 1–3 showed that distance is not an important factor in apprehending spatial relations unless distractors are present. The theoretical attempts to account for the results emphasize the importance of the distinction between reference and located objects, the spatial indexing process that allows the system to focus on objects individually, and

reference-frame computation that allows the system to assign *above* and *below* to arbitrary regions of space. Theories that incorporate these representations and processes do a better job of accounting for the apprehension of spatial relations than theories that do not incorporate them. Finally, the attempt to account for the results emphasizes the importance of parallel processing in deciding whether a given located object is a good example of a specified spatial relation with respect to another, reference object.

References

- Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. C. Collins, S. E. Gathercole, & M. A. Conway (Eds.), *Theories of memories* (pp. 29–101). London: Erlbaum.
- Baylis, G. C., & Driver, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 451–470.
- Biederman, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review*, *94*, 65–96.
- 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.
- Castillo, U., & Umiltà, C. (1992). Splitting focal attention. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 837–848.
- 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). New York: 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). New York: Academic Press.
- Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review*, *96*, 433–458.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception and Psychophysics*, *40*, 225–240.
- Eriksen, C. W., & Webb, J. M. (1989). Shifting of attentional focus within and about a visual display. *Perception and Psychophysics*, *45*, 175–183.
- Eriksen, C. W., & Yeh, Y. Y. (1985). Allocation of attention in the visual field. *Journal of Experimental Psychology: Human Perception and Performance*, *11*, 583–598.
- 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. *Neuropsychologia*, *28*, 335–347.
- Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, *28*, 3–71.
- 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.
- Heathcote, A., & Mewhort, D. J. K. (1993). Representation and selection of relative position. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 488–516.
- Hellige, J. B., & Michimata, C. (1989). Categorization versus

- distance: Hemispheric differences for processing spatial information. *Memory & Cognition*, 17, 770–776.
- Herskovits, A. (1986). *Language and spatial cognition: An interdisciplinary study of the prepositions in English*. Cambridge, England: Cambridge University Press.
- 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). S_Earch via Recursive Rejection (SERR): A connectionist model of visual search. *Cognitive Psychology*, 25, 43–110.
- Ishihara, S. (1987). *Ishihara's tests for colour-blindness*. Tokyo: Kanehara.
- 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. 145–169). Hillsdale, NJ: Erlbaum.
- Jacobs, R. A., Jordan, M. I., & Barto, A. G. (1991). Task decomposition through competition in a modular connectionist architecture: The what and where vision tasks. *Cognitive Science*, 15, 219–250.
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge, MA: Harvard University Press.
- Jolicoeur, P., Ullman, S., & MacKay, L. (1986). Curve tracing: A possible basic operation in the perception of spatial relations. *Memory & Cognition*, 14, 129–140.
- Kahneman, D., Treisman, A., & Gibbs, B. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, 24, 175–219.
- Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres. *Psychological Review*, 94, 148–175.
- Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., & Koenig, O. (1992). Categorical versus coordinate spatial relations: Computational analyses and computer simulations. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 562–577.
- Kosslyn, S. M., Koenig, O., Barrett, A., Cave, C. B., Tang, J., & Gabrieli, J. D. E. (1989). Evidence for two types of spatial representations: Hemispheric specialization for categorical and coordinate relations. *Journal of Experimental Psychology: Human Perception and Performance*, 15, 723–735.
- Landau, B., & Jackendoff, R. (1993). “What” and “where” in spatial cognition and spatial language. *Brain and Behavioral Sciences*, 16, 217–238.
- 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. (1975). On the independence of naming and locating masked targets in visual search. *Canadian Journal of Psychology*, 29, 51–58.
- Logan, G. D. (1994). Spatial attention and the apprehension of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1015–1036.
- Logan, G. D. (1995). Linguistic and conceptual control of visual spatial attention. *Cognitive Psychology*, 28, 103–174.
- Logan, G. D., & Sadler, D. D. (in press). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. Peterson, M. Garrett, & L. Nadel (Eds.), *Language and space*. Cambridge, MA: MIT Press.
- Marr, D. (1982). *Vision*. New York: Freeman.
- 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.
- Miller, G. A., & Johnson-Laird, P. N. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.
- O'Reilly, R. C., Kosslyn, S. M., Marsolek, C. J., & Chabris, C. F. (1990). Receptive field characteristics that allow parietal lobe neurons to encode spatial properties of visual input: A computational analysis. *Journal of Cognitive Neuroscience*, 2, 141–155.
- Pick, H. L., & Acredolo, L. P. (Eds.). (1983). *Spatial orientation: Theory, research, and application*. New York: Plenum Press.
- Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology*, 32, 3–25.
- Pylyshyn, Z. (1989). The role of location indices in spatial perception: A sketch of the FINST spatial-index model. *Cognition*, 32, 65–97.
- Rueckl, J. G., Cave, K. R., & Kosslyn, S. M. (1989). Why are “what” and “where” processed by separate cortical visual systems? A computational investigation. *Journal of Cognitive Neuroscience*, 1, 171–186.
- Sergent, J. (1991). Judgments of relative position and distance on representations of spatial relations. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 762–780.
- Shaw, M. L. (1978). A capacity allocation model for reaction time. *Journal of Experimental Psychology: Human Perception and Performance*, 4, 586–598.
- 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 Press.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97–159.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *The analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- Vandaloise, C. (1991). *Spatial prepositions: A case study from French*. Chicago: University of Chicago Press.

Received August 2, 1994

Revision received November 14, 1994

Accepted December 20, 1994 ■