COGNITIVE PSYCHOLOGY 28, 103–174 (1995)

Linguistic and Conceptual Control of Visual Spatial Attention

GORDON D. LOGAN

University of Illinois

A theory of voluntary, top-down control of visual spatial attention is presented that explains how linguistic cues like "above," "below," "left," and "right" are used to direct attention from one object to another. The theory distinguishes between perceptual and conceptual representations of space and views attention as a set of mechanisms that establish correspondences between the representations. Spatial reference frames play an important part in this analysis. The theory interprets reference frames as mechanisms of attention, similar to spatial indices but with more computational power. The theory was tested in 11 experiments that assessed the importance of linguistic distinctions between classes of spatial relations (*basic*, *deictic*, and *intrinsic*) and examined the flexibility with which subjects manipulated spatial reference frames.

INTRODUCTION

How is attention directed to objects in space? How is attention directed from one object to another? These questions are important because they address the top-down control of attention by language and thought. They are important in the psychological literature on visual spatial attention, where they are addressed in experiments on visual search and attentional cuing: Subjects are asked to find a target in a display of distractors. Sometimes a cue, such as a bar marker, indicates the position of the target. Costs and benefits of distractors and cues are explained in terms of attentional processes, such as spotlights and spatial indices, moving from object to object.

Questions about directing attention are also important in the linguistic literature on spatial deixis and spatial representation. Linguists study sentences like "what is that?" and "is that your glove?" to discover how language directs attention to objects and sentences like "what is under the

This research was supported by National Science Foundation Grant BNS 88-11026 and BNS 91-09856. The final draft was written while I was a Visiting Professor in the Faculty of Psychology at the University of Amsterdam. I thank Jane Zbrodoff for invaluable discussion at all stages of the project, and Julie Delheimer for testing the subjects and analyzing the data. I thank Steve Palmer, Greg Murphy, and three anonymous reviewers for helpful comments on the manuscript. Correspondence and reprint requests should be addressed to Gordon D. Logan, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820. E-mail: glogan@s.psych.uiuc.edu.

0010-0285/95 $9.00

table?" and "is that Kay beside Dave?" to discover how language directs attention from one object to another. Linguistic theory provides a different but complementary perspective on the top-down control of attention. It focuses on representation more than process, distinguishing between classes of spatial representation that underlie top-down control, whereas attention theory focuses on process more than representation, distinguishing between top-down and bottom-up control.

The purpose of this article is to integrate the attention literature and the linguistic literature, putting attentional processes together with linguistic representations to develop a theory of how attention is directed around space. The main contribution of the theory is to explain how attention is directed from one object to another. I argue that this involves different representations and processes than directing attention to a single object. It involves a two-argument conceptual representation that defines the direction from one object to another in terms of a reference frame and a reference frame that defines direction in perceptual space. The current attention literature does an adequate job of explaining how attention is directed to a single object but the representations and processes involved (a perceptual representation of locations and a spatial indexing process) cannot explain how attention is directed from one object to another. I borrowed representations and processes from both literatures to build a theory that explains how a cue can direct attention to a target.

The article begins by reviewing the attention and linguistic literatures and analyzing the answers they offer to the questions about directing attention in space. Then a theory of linguistic and conceptual direction of attention is proposed, contrasting three classes of spatial relations (*basic*, *deictic*, and *intrinsic*) and emphasizing the role of spatial reference frames in directing attention from one object to another. Then 11 experiments on attentional cuing are reported that manipulate the relation between the cue and the target to test the importance of the distinction between classes of spatial relations to test the hypothesis that spatial reference frames act as mechanisms of attention.

ATTENTION AND SPACE

Visual spatial attention has been the dominant paradigm in attention research for the last decade at least. Eriksen's work on the spotlight model, begun in the 1970s (Eriksen & Eriksen, 1974; Eriksen & Hoffman, 1972), blossomed in the 1980s. The spotlight theory developed into zoom-lens (Eriksen & St. James, 1986) and gradient (LaBerge & Brown, 1989) theories and spawned a host of experiments contrasting space-based attention with object-based attention (Duncan, 1984; Kahneman & Henik, 1981; Kahneman, Treisman & Gibbs, 1992; Kramer & Jacobson, 1991). Posner's work on attentional cuing in detection tasks (e.g., Posner, 1980;

Posner & Cohen, 1984) led to an astoundingly successful search for underlying neural structures (Posner, 1988; Posner & Petersen, 1990; Posner & Presti, 1987) as well as an important distinction between voluntary and automatic attentional cuing (Jonides, 1981; Müller & Rabbitt, 1989; Yantis & Jonides, 1990). Treisman's work on feature integration theory (Treisman & Gelade, 1980; Treisman & Schmidt, 1982) introduced the binding problem to the attention literature (Pylyshyn, 1989; Ullman, 1984), renewed interest in visual search (e.g., Duncan & Humphreys, 1989; Treisman, 1991; Wolfe, Cave, & Frankel, 1989), and spawned important experiments on preattentive processes and the nature of perceptual features (Treisman, 1988; Treisman & Gormican, 1988).

A clear message from a decade of research is that space is important. Location is special. There is almost universal agreement that location is the primary attribute on which visual selection is based. Selection on the basis of other visual attributes, such as color and form, depends on selection by location (Nissen, 1985). The major theories of visual selection, except Bundesen's (1990), assume that the selective mechanism addresses locations (e.g., Eriksen & St. James, 1986; LaBerge & Brown, 1989; Posner & Cohen, 1984; Treisman & Gelade, 1980; Treisman, 1991; van der Heijden, 1992).

Representing Space

Location may be special, but the representation of space has not been an important theoretical issue in the attention literature. Very little has been said explicitly about how space is represented, perhaps because it seems that little needs to be said: Objects are arrayed in space in the world; optics preserve that spatial arrangement on the retinae; and the connections from retinae to cortex preserve that spatial arrangement in visual cortex. Most researchers assume, implicitly at least, that attention works with retinotopic representations in cortex. Space is represented perceptually as a two- or three-dimensional map and locations are represented as coordinates in that space. Treisman's idea of a feature map relies on representations like these (e.g., Treisman, 1988; Treisman & Sato, 1990; also see Cave & Wolfe, 1990; Wolfe et al., 1989). Researchers in Eriksen's and Posner's traditions appear to make the further assumption that the space has a Euclidean metric; distance, defined in degrees of visual angle, is an important variable in many studies (e.g., Egly & Homa, 1991; Tsal, 1983). These ideas pervade space-based approaches to attention; spotlights (Treisman & Gormican, 1988), zoom lenses (Eriksen & St. James, 1986), and gradients (LaBerge & Brown, 1989) are defined in terms of degrees of visual angle.

The representation of space is not much more explicit in object-based approaches to attention. Object-based approaches differ from space-based approaches because they assume that attention is directed to objects or groups of objects rather than to regions of visual space. They assume that space is represented perceptually as a two- or three-dimensional array of objects and surfaces that is produced automatically by data-driven bottom-up processes (see e.g., Marr, 1982; Pylyshyn, 1984; Ullman, 1984). Object-based approaches gain support from demonstrations that perceptual grouping effects modulate or override distance effects (Duncan, 1984; Driver & Baylis, 1989; Kahneman & Henik, 1981; Kahneman et al., 1992; Kramer & Jacobson, 1991; Treisman, Kahneman, & Burkell, 1983). Those results call into question the assumption that perceptual distance is strictly Euclidean, but they do not provide an alternative metric.

Processing Space

Location is special because the mechanisms of visual selection address space. The literature discusses (and contrasts) two mechanisms of spatial selection: spotlights and spatial indices. The idea of a spotlight (Posner, 1980; Treisman & Gormican, 1988) or zoom lens (Eriksen & St. James, 1986) or gradient (LaBerge & Brown, 1989) is common in space-based approaches to attention. The spotlight represents a region of space where processing is enhanced. Formally, the spotlight acts like a gain control, amplifying signals from the selected (i.e., the area within the beam) or dampening signals from nonselected regions (i.e., the area outside the beam) or both. It serves as a selection mechanism by increasing the activation of stimuli and responses within the beam relative to stimuli and responses outside the beam; the strongest response comes from attended stimuli. Its ability to select is limited, however. It can select a single object falling within the beam, but it cannot select one or two (or more) objects that fall within the beam or select different properties of a single object (van der Heijden, 1992).

The idea of spatial indexing is common in object-based approaches to attention (Kahneman & Treisman, 1984; Kahneman et al., 1992; Pylyshyn, 1989; Ullman, 1984) and a central part of Treisman's feature integration theory (Treisman & Gelade, 1980; Treisman & Schmidt, 1982; also see Wolfe et al., 1989, but see Duncan & Humphreys, 1989). A spatial index is a symbol that is attached to a location or an object in a perceptual representation, like a mental cursor. The process that attaches the index is called binding or spatial indexing. A spatial index establishes correspondence between conceptual and perceptual representations. It provides conceptual processes access to the perceptual representation of the selected object, serving as a symbolic address for the object.

Spatial indices are mechanisms of attention because they are mechanisms of selection: A specific index is attached to a specific object, distinguishing it from other objects. Spatial indices are limited in their selective ability. They select among alternative objects, but they do not select different properties of a single object. They provide access to that object so that other mechanisms can select among properties.

Spotlights and spatial indices are often contrasted in attempts to distinguish space-based and object-based approaches to attention, as if they are mutually exclusively alternatives. However, they are not incompatible. Empirically, the data support spotlight predictions (distance matters) as well as spatial index predictions (objects matter; e.g., Kramer & Jacobson, 1991). Theoretically, spotlights may play a role in spatial indexing (Briand & Klein, 1987). For example, Treisman explicitly used a spotlight for spatial indexing (e.g., Treisman, 1991; Treisman & Sato, 1990) and van der Heijden (1992) used a gain-control mechanism like the one underlying the spotlight as part of a mechanism for spatial indexing.

How Is Attention Directed to Objects?

The representations and processes discussed so far provide a sufficient account of how attention is directed to single objects: The perceptual representation locates objects in a two- or three-dimensional space and the spatial indexing process marks the location of the selected object. The process is clear when there is only one stimulus. There is a single location active in the perceptual representation, and the spatial indexing process marks it. The process is more complicated when there is more than one stimulus: the selected object must be chosen somehow.

The choice of an object to select—the direction of attention—is strongly influenced by parallel preattentive processes that detect differences in homogeneous displays. Attention is attracted automatically to items that stand out from the others, such as a red letter in a field of green ones or an X among Os, and to items that appear suddenly or move suddenly. Young adults can detect these differences quickly, independent of the number of homogeneous distractors in the display (Treisman, 1988; Treisman & Gelade, 1980). Young adults can find target items much faster if they stand out from the distractors; they take much longer to find them if they do not stand out or if a nontarget item stands out and draws attention away from the target (Jonides & Yantis, 1988; Theeuwes, 1991, 1992; Yantis & Jonides, 1984).

The direction of attention is influenced by preattentive processes but is controlled by top-down processes that allow people to select what they intend to select (environment willing). People generally succeed at search tasks, if the display is on long enough, even if the target does not stand out from the distractors and even if attention is first attracted to a nontarget. Search is rapid and efficient as well as successful. Difficult conjunction search takes 60 to 80 ms per item (e.g., Treisman & Galade, 1980). Subjects adopt a self-terminating strategy (stopping when they find the target), which is more efficient than the alternative, exhaustive search strategy (examining all the items all the time; see e.g., Treisman & Gelade, 1980).

Exactly how the control is accomplished is a mystery. Subjects must examine the items until they find a target and they must keep track of the items they have and have not examined; they should examine each item only once. Several algorithms could accomplish these computational goals, but few have been proposed and fewer have been tested empirically (guided search theory is a notable exception; see Cave & Wolfe, 1990; Wolfe et al., 1989). In principle, it should be possible to propose a control algorithm that relies on the same representations and processes involved in directing attention to a single object (i.e., a spatial indexing process and a perceptual representation of locations). For example, the algorithm could select a location at random from a set of alternatives, process the stimulus at that location, terminate if it is a target, and select another location if it is not a target. The next location could be selected by applying the same algorithm, recursively. Selected locations could be eliminated from the set of alternatives by marking them with spatial indices and restricting the choice of the next alternative to the set of locations that have not been indexed.

How Is Attention Directed from One Object to Another?

The representations and processes that are sufficient to direct attention to a single object are not sufficient to direct attention from one object to another. A spatial index represents the location of a single object, not the location of one object relative to another. A two- or three-dimensional array represents the locations of individual objects explicitly and directly, but relative location is implicit and requires more computation and more computational machinery to make it explicit. The X–Y coordinates of two objects in two-dimensional space do not specify the distance between the objects, but they provide information that could be used to specify the distance, given some further computation (e.g., distance $= [(X_1 - X_2)^2 + (Y_1 - Y_2)^2]^{1/2}$).

Directing attention from one object to another involves choosing a specific object to move to. The choice may not matter much in search tasks, as long as attention moves to a new location, but in cuing tasks, attention is supposed to go to a particular location—one that satisfies the relation between the cue and the target that is specified in the instruction to direct attention. The alternatives can be ordered in priority according to how well they satisfy the relation and attention can be directed to them in order of priority. The choice process requires information, like proximity, direction, or relative position, that is not directly available in the perceptual representation and cannot be produced simply by spatially indexing different items. Further computation with different computational machinery (i.e., a spatial reference frame) is necessary before the choice can be made.

This conclusion is surprising because it means that current theories of attention cannot explain how attention is directed from a cue to a target. A great deal of the last decade's research on visual spatial attention addressed cuing attention, so someone should have explained this important step. Most investigators were interested in other issues, such as the time course of attentional cuing (Colegate, Hoffman, & Eriksen, 1973; Posner & Cohen, 1984), the spatial extent of spatial cuing (Eriksen & Hoffman, 1972; Juola, Bouwhuis, Cooper, & Warner, 1991; LaBerge & Brown, 1989), the processing in uncued locations (Eriksen & Eriksen, 1974; Miller, 1991), and the discrete versus continuous nature of the movement from cue to target (Eriksen & Murphy, 1987; Remington & Pierce, 1984; Shulman, Remington, & McLean, 1979; Tsal, 1983). Their investigations required them to assume only that attention was directed from the cue to the target. How it knew how to get there was not important.

Push Cues and Pull Cues

The conclusion that the attention literature cannot account for how attention is directed from the cue to the target is surprising because a substantial portion of the past decade's research was addressed to mechanisms that underlie cuing effects. This research distinguished between two classes of cues, variously called *central* versus *peripheral* cues (Jonides, 1981), *endogenous* versus *exogenous* cues (Posner, 1980), and *push* versus *pull* cues (Kahneman et al., 1992). In each case, the second kind of cue is thought to attract attention automatically, to "pull" it from its current location to the location of the cue. The first kind of cue in each case is thought to require some sort of cognitive interpretation before attention can be moved or "pushed" from its current location.

Many studies suggest this distinction is important: Push cues take longer to use than pull cues (Eriksen & Collins, 1969). Push cues move attention only when they are valid (i.e., indicate a target's position with greater than chance accuracy), whereas pull cues attract attention whether or not they are valid (Jonides, 1981; Yantis & Jonides, 1990). Push cues are easily preempted by pull cues, whereas pull cues are hard to preempt (Müller & Rabbitt, 1989). Push cues are used less when resources are taxed by a concurrent task, whereas pull cues are utilized despite resource limitations (Jonides, 1981).

Most of the research has been directed toward establishing the characteristics of pull cues: Single items that differ from homogeneous distractors tend to "pop out," pulling attention toward them (Treisman &

Gelade, 1980; Treisman & Gormican, 1988). Discrepant items attract attention automatically (Theeuwes, 1991, 1992). Items that appear suddenly in the periphery are especially powerful (Jonides & Yantis, 1988; Yantis & Jonides, 1984, 1990; also see Miller, 1989). Research has focused on characteristics that distinguish push cues from pull cues, not on characteristics that distinguish push cues from each other. Consequently, little is known about push cues themselves (but see Eriksen & Collins, 1969).

The attention literature provides an adequate account of pull cues: Pull cues affect performance because they attract attention to a single item, pulling it toward or away from the target. Performance is better when attention is pulled toward the target than away from it. Typically, pull cues are not logically necessary to perform the task. In Posner-type experiments, for example, subjects' task is to detect a brief dot and the cue is a box that surrounds it or appears in the other field. Subjects can detect the dot and respond appropriately whether or not they see the cue. The cue directs attention to itself, and that increases sensitivity in its neighborhood. A target falling nearby will benefit from the increased sensitivity; a target falling far away will not. It may require a further shift of attention before it is processed. All that is required to account for these effects is a two- or three-dimensional map of locations and a spatial indexing process.

Push cues require a different explanation. Push cues require two shifts of attention, one to the cue and one from the cue to the target, and the second shift must be directed somehow. A two- or three-dimensional map of locations and a spatial indexing process are not sufficient to specify direction. Thus, the literature on visual spatial attention does not explain how attention is directed voluntarily from one object to another and it does not have the computational machinery necessary to provide an explanation (also see van der Heijden, 1992). To see what else is needed, we need to look beyond the attention literature to the linguistic literature on deixis and spatial relations.

LANGUAGE AND SPACE

Linguists are interested in space because space is central in language. Every known language represents space, providing speakers with means to direct listeners' attention to objects in space and from one object to another. There is considerable interest in *deixis,* the process by which discourse is grounded in the spatial context in which it occurs, and considerable interest in the representation of space per se. Linguistic analyses of space have important implications for (attentional) processing: The spatial concepts that languages express have survived a long process of linguistic evolution. They survive because they express useful distinctions between representations of space, distinctions that the perceptual and attentional systems must be able to support. Linguistic representations specify computational goals for attentional algorithms (Clark, 1973; Jackendoff & Landau, 1991).

An important body of linguistic and psycholinguistic research focuses on elementary spatial relations like *above, below, left,* and *right,* beginning in the 1970s (Clark, 1973; Miller & Johnson-Laird, 1976) and blossoming in the 1980s and 1990s (Garnham, 1989; Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1991; Langacker, 1986; Levelt, 1984; Talmy, 1983; Vandaloise, 1991). Elementary spatial relations are important because they represent the relative positions of objects in space, characterizing the location of one object with respect to the location of another. This kind of representation is what is missing in the attention literature. It is the kind of representation that can direct attention from one object to another, from cue to target. Elementary spatial relations are important as well because their meaning is defined with respect to a reference frame imposed on space. The reference frame defines direction in space and the relation defines direction with respect to the reference frame. This kind of representation is also missing in the attention literature. It is necessary to direct attention from one object to another.

Representing Space

Linguistic representations of space are conceptual. They are propositions that assert predicates about objects and the relations between them. Spatial relations are categorical (Herskovits, 1986; Jackendoff, 1983; Jackendoff & Landau, 1991; Miller & Johnson-Laird, 1976; Talmy, 1983; Vandeloise, 1991; also see Kosslyn, 1987). They do not describe positions precisely, as a Cartesian coordinate system would. Instead, they refer to large regions of space. "Above the keyboard," for example, refers to a region projected directly upward from the keyboard, including my hands, the ceiling, the roof, the clouds, and the stars. Spatial relations are categorical in that they treat objects that occupy discriminably different positions as equally good examples of the category. My hands, the ceiling, the roof, the clouds, and the stars all occupy different positions, but they are equally good examples of "above the keyboard" (Logan & Sadler, 1995).

Three Classes of Spatial Relations

Linguists and psycholinguists distinguish three classes of spatial relations: *basic, deictic,* and *intrinsic* (Garnham, 1989; Herskovits, 1986;

Levelt, 1984; Miller & Johnson-Laird, 1976).[1] All three classes of relations relate objects to a spatial reference frame. They differ in the number and nature of the objects they relate and in the nature of the reference frame they relate them to. These differences impose different computational goals and have implications for attentional processing.

Basic relations. Basic relations take one argument. They specify the location of a single object with respect to the reference frame of the viewer. Basic relations do not indicate where objects are with respect to each other. Knowing that my coffee cup is *there* and my disk is *there* (both basic relations) does not tell me where the coffee cup is with respect to the disk. Basic relations are like spatial indices, pointing to the location of the object they refer to.

Deictic relations. Deictic relations take two or more arguments. They specify the position of one object with respect to another in terms of the reference frame of the viewer, projected onto the other object: The lamp is *left of* the table because it would be to my left if I were in the same position as the table (e.g., if I were to walk there).

The arguments of deictic relations are ordered. One argument represents the *located object* (or *target* or *figure*) and the other(s) represent the *reference object(s)* (or *landmark(s)* or *ground(s)*; Jackendoff, 1983; Talmy, 1983). The located object is the focus of the relation, the object that attention is directed to. The relation specifies the position of the located object with respect to the reference object (i.e., in terms of a reference frame aligned with the reference object). The reference object is chosen because it is a useful landmark. Its position is known in advance or is easy to see in the display.

The distinction between located object and reference object is crucial in noncommunative relations like *above*, *below*, *left*, and *right*, where the truth of the relation depends on the order of the arguments; "the horse is in front of the cart" is not the same as "the cart is in front of the horse." The distinction is important pragmatically even with commutative relations; "the disk is beside the cup" does not mean the same as "the cup is beside the disk." The first case focuses on the location of the disk, the second on the location of the cup.

Intrinsic relations. Intrinsic relations take two or more arguments that are ordered, like deictic relations, but they use a different reference frame. Intrinsic relations specify the position of a located object with respect to the *intrinsic axes* of a reference object. Intrinsic relations are more restricted than deictic relations because they apply only to sets of

---

[1] Most linguists distinguish between deictic and intrinsic spatial relations. Garnham (1989) introduced the idea of basic spatial relations.

objects in which the reference object has intrinsic axes, whereas deictic relations can be applied to any set of objects.

Objects with intrinsic axes have fronts and backs, tops and bottoms, and left and right sides. Some objects, like people, animals, vehicles, and buildings, have all three intrinsic axes. They can serve as reference objects in any intrinsic relation. Objects like trees have tops and bottoms but no fronts and backs. They can support intrinsic *above* and *below* relations but not intrinsic *front* and *back* relations. Arrows and bullets have fronts and backs but no tops and bottoms and no left and right sides. They can support intrinsic *front* and *back* relations but not intrinsic *above* and *below* relations. Objects like balls have no intrinsic axes. The top of a ball is the part that happens to coincide with an extension of the gravitational vertical axis projected upward from the center of the ball. The part of the ball that is the top changes as the ball is rotated. Thus balls and objects like them cannot serve as reference objects in intrinsic relations.

Deictic and intrinsic relations have the representational power necessary to support direction of attention from one object to another, the power necessary to support attentional cuing. They take two objects as arguments and they specify how to get from one to the other. Basic relations lack the power to support attentional cuing. They take one argument rather than two, and they do not specify direction. This difference gives some insight into the difficulty the attention literature has in accounting for attentional cuing: The representations and processes in current theories of attention support basic relations but not deictic and intrinsic relations. The problem is that cuing requires deictic and intrinsic relations. A theory of attentional cuing must explain how deictic and intrinsic relations are computed.

Processing Space

Linguistic theories focus more on representation than process, so detailed theories of processes and their implementation must await further investigation. Nevertheless, the conceptual representations of space impose clear computational goals on processes that would use them. The computational goals require three processing operations: spatial indexing, aligning spatial reference frames, and computing the relation with respect to the reference frame.

Spatial indexing. Spatial indexing is made necessary because the arguments of spatial relations are highly schematized (Clark, 1973; Herskovits, 1986; Jackendoff & Landau, 1991; Talmy, 1983). Spatial relations refer to geometric, volumetric, or topological properties of objects. They treat objects as points, lines, regions, volumes. For example, "the bird is in the tree" schematizes the bird as a point and the tree as a three-dimensional volume. "The tree is on (the side of) the road" schematizes the tree as a point and the road as a line. Spatial relations generalize over metric properties like shape and size. Balls and boxes can be *on* a table, flies and clouds can be *over* one's head. The schematization of arguments makes spatial relations very broad in scope. A given spatial relation can apply to an indefinitely large number of objects. Consequently, the specific objects to be related in a specific case must be chosen by an act of attention (i.e., by spatial indexing).

Spatial indexing is also made necessary by the ordering of arguments of intrinsic and deictic spatial relations. The distinction between located object and reference object is essential to the meaning of spatial relations. Spatial indexing is necessary to keep track of which argument is which.

Aligning spatial reference frames. Spatial relations express location relative to reference frames (Clark, 1973; Garnham, 1989; Levelt, 1984). What is *above* an object depends on what one considers to be *up*, and that choice of direction amounts to a choice of a reference frame. A reference frame is a set of coordinate axes that define a three-dimensional space. A reference frame has four parameters—an origin, an orientation, a direction, and a scale—that can be set according to the demands of the task.

Aligning spatial reference frames is a necessary step in computing deictic and intrinsic relations. Spatial indexing may indicate the location of the target and the landmark, but the relation between them is defined in terms of the reference frame. The reference frame must be projected onto (deictic relations) or extracted from (intrinsic relations) the reference object before the relation can be computed. The origin of the reference frame must be translated so it coincides with the (center of) the reference object, the orientation of the reference frame must be rotated so it coincides with the orientation of the viewer (deictic relations) or the reference object (intrinsic relations), and the scale must be set with respect to the located object, the reference object, or some other standard (Morrow & Clark, 1988). Spatial reference frames are very flexible mechanisms.

Reference frames give an orientation and a scale to the space they are projected on and they mark the location of their origin. They serve as a map between representations, establishing correspondence between conceptual and perceptual representations of space. Reference frame parameters can be adjusted flexibly; reference frames can be moved around space and aligned with any object. Reference frames orient conceptual processes to space, just as spatial indexing processes orient conceptual processes to objects.

Reference frames are powerful as well as flexible. Gain controls (spotlights) may facilitate processing, but reference frames enable it. Spatial relations cannot be computed without a reference frame. Their truth cannot be defined until a reference frame is specified. This power and flexibility suggests that reference frames may be important mechanisms of attention.

Computing the specified relation. The specified relation represents a computation to be performed on the perceptual representation in alignment with the reference frame. The computation specifies the position of the located object with respect to the reference object and the reference frame. The computation could involve imposing a spatial template that defines a region of acceptability, such that objects in that region are good examples of the relation and objects outside it are not (e.g., Herskovits, 1986; Langacker, 1986). For example, *near* to might be represented by a gradient that peaks in the middle (where it is aligned with the reference object) and tapers off toward the edges. Objects *near* to the reference object have high values on this gradient; objects farther from it will have lower values (Logan & Sadler, 1995). Alternatively, the computation could involve applying "serial visual routines" (Ullman, 1984) such as those involved in mental curve tracing (Jolicoeur, Ullman, & MacKay, 1986, 1991). *Near to* could be computed by a "coloring" operation that propagated outward from the reference object. Near objects would be colored before far objects; order of coloring specifies the spatial order of distances.

How Is Attention Directed to Objects?

Language directs attention to objects by naming them, by describing their perceptual properties, or by referring to their (basic-relation) locations. A speaker may direct a listener's attention to a red ball by saying "look at the ball" or "look at the red thing," or by pointing and saying "look at that." This requires the representations and processes necessary for spatial indexing and a dictionary (or a language comprehension system) that can retrieve perceptual descriptions of concepts like "ball" and "red."

The number of objects that language can direct attention to is indefinitely large. Language directs attention to objects by describing them in ways that distinguish them, and language has the capacity to generate an indefinitely large number of descriptions. Young adults know 30,000 object names (approximately; see Biederman, 1987; Jackendoff & Landau, 1991); they have 30,000 distinctions lexicalized. They can distinguish between even more objects (e.g., objects with the same name or objects with no names) by generating sentences.

How Is Attention Directed from One Object to Another?

Language directs attention from one object to another by specifying the location of one relative to the other with a deictic or intrinsic relation. The linguistic distinction between located and reference objects specifies a direction for attention to move—from the reference object to the located object. The relation itself specifies the direction further. It specifies the computation necessary to get from the reference object to the located object. A sentence like "is that Tom beside Denise?" focuses the listener's attention on Tom, via Denise. It tells the listener to find Denise, extract her intrinsic axes, and look to her sides to find Tom. The details of the processing remain to be specified.

The number of ways that language can direct attention from one object to another is indefinitely large. Language directs attention from one object to another by describing the spatial relation between them, and language has the capacity to generate indefinitely many descriptions of spatial relations. English has 70–80 spatial relations lexicalized (Jackendoff & Landau, 1991). Relations that are not lexicalized can be described in sentences by combining elementary relations.

TOWARD A THEORY OF CONCEPTUAL CUING

The parts of a theory of linguistic and conceptual control of attention have already been described. Now they can be put together: The theory assumes two different representations of space, a perceptual one borrowed from the attention literature and a conceptual one borrowed from the linguistic literature. The perceptual representation is a two- or three-dimensional array of objects and surfaces and the conceptual representation is a proposition, a predicate expressing a spatial relation. The theory assumes two basic attentional mechanisms, spatial indices and spatial reference frames, that establish correspondence between perceptual and conceptual representations. Spatial indices and spatial reference frames can be moved freely and assigned to or aligned with any perceptual object. The theory has little to add to current theories of how attention is directed to single objects. Spatial indices, basic relations, and simple perceptual representations appear to be sufficient for the required computation.[2] The contribution of the theory lies in its explanation of how attention is directed from one object to another. The theory explains how attentional cuing is possible.

The theory assumes that attention is directed from one object to another by computing a deictic or intrinsic spatial relation between them. The conceptual representation sets computational goals for the basic attention mechanisms. It specifies the cue as the reference object and the target as the located object and it specifies a deictic or intrinsic relation to be computed between them. The theory assumes there are three major computational goals: The first is to locate the cue, the second is to locate

---

[2] Basic relations may not be sufficient to identify the objects whose locations they specify. Many objects are made of parts and take their identity from the arrangement of their parts (Biederman, 1987; Marr & Nishihara, 1978). The arrangement of parts may be described by deictic and intrinsic relations and so require reference-frame computation to be apprehended.

the target with respect to the cue, and the third is to do what is required with the target.

In the first computation, the location of the cue can be specified by basic, deictic, or intrinsic relations. If basic relations are sufficient, the position of the cue must be indexed in the perceptual representation and that index must be bound to the reference-object argument of the corresponding conceptual relation. Basic relations are sufficient in most attention experiments. Only one cue is presented, and it often appears in a blank field before the display appears, so there is nothing to confuse it with or relate it to deictically or intrinsically. The cue is simply *there*. Basic relations may be sufficient in most conversations; people choose reference objects that are perceptually conspicuous or whose location is already known (Talmy, 1983).

Cue position could be specified deictically or intrinsically. In "look under the chair beside the tree," the cue is "chair" and its position is specified deictically with respect to the observer's reference frame imposed on the tree. In "look under the bench in front of the piano," the position of the cue, "bench," is specified intrinsically with respect to the piano's reference frame. These cases involve the same sort of computation that is required in the second stage, getting from the cue to the target. They can be dealt with by applying second-stage processes.

In the second computation, the location of the target with respect to the cue is specified. This involves imposing (deictic relations) or aligning (intrinsic relations) a reference frame on the cue and computing the relation between the cue and the target with respect to the reference frame. The relation is computed by orienting a template (Herskovits, 1986; Langacker, 1986) or a visual routine (Ullman, 1984) with respect to the reference frame and applying it to the perceptual representation. The procedure continues until the first object is found (or none are found), or it continues until several objects have been found and the one that fits best is selected.

The procedures for computing relations can be used for several purposes besides attentional cuing. For example, they could be used to verify the relation between a pair of objects, as in picture-sentence verification tasks (e.g., Clark, Carpenter, & Just, 1973): The two objects could be located (with spatial indices) and the relation between them computed (with templates or visual routines). The relation computed from the display could be compared with the one in the (sentence) question. The flexibility of the procedures for computing relations is an important property. It makes them very broad in scope.

The theory assumes that the parameters of the reference frame can be adjusted voluntarily. The capacity for voluntary adjustment is the one of the main reasons for considering the reference frame to be a mechanism of visual–spatial attention; a mechanism of spatial attention ought to have this flexibility. The kinds of computations that reference frames support is another: Reference frames support computations that cannot be done by local, bottom-up processing (see Ullman, 1984). These are likely to be attentional computations.

The third computation, processing the target in accord with the task set, is not directly relevant to the analysis of spatial cuing. Nevertheless, it can be understood in much the same way, in terms of operations on a perceptual and a conceptual representation. Task sets could be interpreted as requiring people to create certain conceptual representations of the information in the perceptual representation and report aspects of those conceptual representations. For example, the requirement to identify the letter beside the cue could result in a proposition about the identity of the letter being bound to the perceptual representation of the letter, a proposition that could be expressed in a verbal report or a key press (see Logan, 1990). However, developing a theory that explains the details of that process is well beyond the scope of this article.

THE EXPERIMENTS

The theory draws a distinction between directing attention to a single object and directing attention from one object to another. The former requires basic relations while the latter requires deictic or intrinsic relations. The theory assumes that deictic and intrinsic relations require subjects to impose or extract a reference frame before computing the relation, but basic relations do not. Thus, the theory predicts reference frame effects in directing attention from one object to another but not in directing attention to single objects. The experiments tested this prediction.

The theory assumes that reference frames are mechanisms of spatial attention, like spotlights and spatial indices. Spotlights and spatial indices orient attention to objects; reference frames orient attention to space. The theory assumes that reference frames are mechanisms of attention because they are powerful and flexible. They can be moved about space, oriented in any direction, and set to any scale. The experiments tested this prediction as well.

There were 11 experiments altogether. Experiments 1–3 examined cuing effects in displays much like those in standard attention experiments, showing that reference frame effects occur under conditions already studied in the attention literature. The remaining experiments focus more sharply on reference frames and the distinction between basic versus deictic and intrinsic relations. Experiments 4–6 looked for reference frame effects in simple displays and found them with deictic but not basic cues. Experiments 7 and 9 examined the ability to rotate reference frames in alignment with deictic cues. Experiment 8 examined the ability to translate reference frames with deictic cues. Experiments 10 and 11 examined the ability to rotate and translate reference frames with intrinsic cues.

In each experiment, the cue was like an instruction that specified which target to report, as in Eriksen's experiments (e.g., Eriksen & Colegate, 1969; Eriksen & Hoffman, 1972; also see Eriksen & Eriksen, 1974). The display consisted of (roughly) equal numbers of red and green forms (except for Experiment 2), and the task was to report the color of the form indicated by the cue. There was no way to respond correctly (except by guessing) without processing the cue. This procedure contrasts with experiments like Posner's (e.g., Posner, 1980; Posner & Cohen, 1984; also see Jonides, 1981), in which the cue is more of a suggestion than an instruction. The cue indicates the location of the target (with some validity) but in principle, subjects can find the target without it. Posner's procedure is advantageous because attention can be defined operationally in terms of the costs and benefits of cuing (van der Heijden, 1992), whereas Eriksen's procedure provides no baseline condition to assess benefits and no invalid cuing condition to assess costs. Attention is defined operationally as the ability to select the correct response. Selecting the correct response implies selecting the correct target, which implies using the cue.

Eriksen's procedure is advantageous because it requires subjects to use the cue on every trial, so every trial provides usable, interpretable data. Posner's procedure makes it hard to know exactly when subjects are using the cue. Observed performance is a mixture of performance based on the cue (when subjects found the target by finding the cue and directing attention from cue to target) and performance without the cue (when subjects found the target directly), and mixtures are hard to disentangle. Subjects may mix strategies differently with different cue types, relying on easy cues and ignoring difficult ones. I was more interested in how subjects used the cue than whether they used it, so I adopted Eriksen's procedure rather than Posner's.

MEASURING REFERENCE FRAMES

Reference frames must be defined operationally, in terms of experimental manipulations and their effects on performance. In this article, reference frames are defined in terms of accessibility: Some regions of space are easy to access from the reference frame and others are not. This difference defines the presence, the position, and the orientation of the reference frame: Objects cued on the above-below axis should be more accessible than objects cued on the front-back axis, which in turn, should be more accessible than objects cued on the left-right axis. I will call this the *conceptual frame* hypothesis and contrast it with the *equal availability* hypothesis, which says that all parts of space are equally accessible.

The experiments will use the conceptual frame hypothesis to identify reference frames. I will infer that subjects used a reference frame when the conceptual frame hypothesis is supported and that they did not use a reference frame when the equal availability hypothesis is supported. Thus, the conceptual frame hypothesis should be supported with deictic and intrinsic relations but not with basic relations. The conceptual frame hypothesis will be used to infer the origin and orientation of the reference frame. If the reference frame is rotated 90 degrees, for example, the advantage of *above* and *below* over *left* and *right* should rotate with the reference frame. The objective left-right axis should show an advantage over the objective above-below axis because it is aligned with the subjective above-below axis.

The operational definition of the reference frame capitalizes on experiments by Franklin and Tversky (1990) and Bryant, Tversky and Franklin (1992). They had subjects report objects in imagined environments. The objects were arrayed in three dimensional (imaginary) space, and the one to be reported was cued with a deictic (Franklin & Tversky, 1990; Bryant et al., 1992) or an intrinsic relation (Bryant et al., 1992). Subjects reported the objects fastest with *above-below* as cues. They were intermediate with *front-back* and slowest with *left-right* as cues.

These differences can be understood in terms of the support the different relations receive from the environment (Clark, 1973; Levelt, 1984). Relations like *above-below* and *up-down,* that involve the vertical axis, are easy to compute because they are well supported by the environment. They are nearly always consistent with gravity, they are consistent over rotation about the vertical axis and over horizontal translations of axes (e.g., the ceiling remains *above* me when I turn around and walk across the room), and they are supported by bodily asymmetries—heads are different from feet and dorsal surfaces are different from ventral surfaces.

Relations such as *front-back* and *ahead-behind* that express the front-back axis are harder to compute. They receive no support from gravity; they are orthogonal to the gravitational vector. They change with rotation about the vertical axis and with horizontal translations of axes (e.g., what is in front of me changes as I turn around and walk past it). But front-back relations are supported by bodily asymmetries: Fronts are usually different from backs. Perceptual apparati usually appear on animals' fronts rather than backs. And animals usually move frontward (foreward).

Relations such as *left-right* and *starboard-port* that express the left-right axis are hardest. They receive no support from gravity, they change with rotation about the vertical axis and with horizontal translation of axes, and they receive no support from bodily asymmetries. Higher animals are typically left-right symmetrical. Many artifacts, like cars and boats, are nearly left-right symmetrical. Often, these relations may be computed with reference to up-down and front-back (Clark, 1973).

These analyses place great importance on the gravitational upright. There may be no advantage for a person's intrinsic or egocentric *above* and *below* when the person's upright departs from the gravitational upright (Levelt, 1984). Franklin and Tversky (1990) found that the advantage of *above-below* over *front-back* disappeared when subjects imagined themselves lying on their sides (there was still an advantage of *above-below* over *left-right*, however). Whether these results will replicate when subjects search visible displays and whether consistency with the gravitational upright is essential are empirical questions addressed in the experiments.

EXPERIMENTS 1–3: COMPARING PUSH CUES

Experiments 1–3 looked for reference frame effects in cuing situations much like those in the standard attention literature. Subjects saw a diamond-shaped array of colored circles (Experiments 1 and 3) or I's and l's (Experiment 2) and reported the color or identity of the form indicated by a bar marker cue presented outside the diamond (see Fig. 1). There were four cuing relations: *Next-to, Opposite, Clockwise,* and *Counterclockwise*. In the Next-to condition, the target was the form nearest to the cue. In Fig. 1, for example, subjects in the Next-to condition would report "dark." In the Opposite condition, the target was the form diametrically opposed to the cue. In Fig. 1, subjects would report "light." In the Clockwise and Counterclockwise conditions, the target was the form adjacent to the one nearest to the cue. Its position was defined in terms of motion around the diamond shape. In the Clockwise condition, the target was the form clockwise from the form nearest to the cue ("dark" in Fig. 1), and in the Counterclockwise condition, the target was the form counterclockwise from the form nearest to the cue ("light" in Fig. 1).

Many studies in the attention literature use *next-to* as the cuing relation (e.g., Eriksen & Hoffman, 1972; Eriksen & St. James, 1986; Jonides,

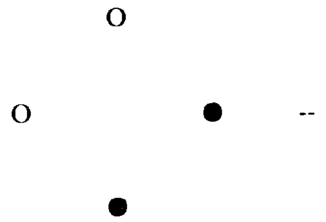

FIG. 1. Sample display from Experiments 1–3 (not to scale).

1981; Yantis & Johnston, 1990). A few studies compared *next-to* with *opposite* and found *opposite* harder, but confounded distance between cue and target with cuing relation (Eriksen & Collins, 1969; Pashler, 1991; Warner, Juola, & Koshino, 1990). Experiments 1–3 used the same displays and the same cues in each cuing condition. Differences observed here cannot be due to stimulus conditions. Few studies, if any, investigated *clockwise* and *counterclockwise*. They were included because they were logical possibilities given the cues and displays and because they involved computations analogous to *left* and *right*.

Each experiment varied the distance between cues and targets in addition to the cuing relation. The cue appeared at one of four different eccentricities relative to the nearest form in the array. The eccentricities were chosen so the distances between cues and targets in the four cuing conditions overlapped (e.g., the nearest cue-target distance in the Opposite condition was smaller than the farthest cue-target distance in the Next-to condition).

The effects of distance within each cuing condition bear on the processes that compute the relation between the cue and the target. Serial visual routines, such as spotlights or spatial indices that move across space continuously, predict that reaction time will increase monotonically with distance (Ullman, 1984; also see Jolicoeur et al., 1986, 1991; Kosslyn, 1980; Tsal, 1983). By contrast, spatial templates (Herskovits, 1984; Langacker, 1986; Logan & Sadler, 1995) can be applied to all the positions simultaneously; spatial parallel processing predicts no distance effect.

Serial visual routines may account for differences between cuing conditions as well: The average distance between cue and target varied with cuing relation, so differences between cuing conditions may be entirely due to distance. Cue-target distances were shortest with *next-to* and longest with *opposite*. *Clockwise* and *counterclockwise* were intermediate in Euclidean distance (i.e., if attention moved in a straight line from the cue to the target) but equal to opposite in city-block distance (i.e., going from the cue to the center of the array and then out to the target). Serial visual routines predict that reaction time should be fastest on average in the *next-to* condition, second fastest in the *clockwise* and *counterclockwise* conditions, and longest in the *opposite* condition. They predict that there should be no differences in reaction time between cuing conditions when distance is held constant.

If the average reaction times do not vary between cuing conditions as predicted and if there are differences between cuing conditions when distance is equal, then serial visual routines cannot account for the differences between cuing conditions. The differences must be due to spatial indexing, reference frame computations, or computing relations with spatial templates. There is no reason to expect differences in computation time for different spatial templates.

Spatial indexing and reference frame effects are hard to separate with these relations: Spatial indexing would predict differences between *next-to,* which takes two arguments and so requires two indexing operations, and *opposite, clockwise,* and *counterclockwise,* which take three arguments (i.e., the cue, the target, and the item the target is *opposite, clockwise* or *counterclockwise* from) and require three indexing operations. Reference frame computation would predict differences between *clockwise* and *counterclockwise,* which require specification of the right-left axis, and *next-to* and *opposite,* which do not. The hypotheses agree on their predictions about *next-to, clockwise,* and *counterclockwise,* disagreeing only on *opposite.* However, reference frame computation might predict that *opposite* should be harder than *next-to* because *opposite* requires an origin and an axis to be specified, whereas *next-to* only requires an origin and a scale. Then the hypotheses would make very similar predictions.

Experiment 3 was a replication of Experiment 1 with brief displays and with cues that preceded the target displays. In Experiment 1 (and 2), cues appeared simultaneously with the target displays and the target displays were exposed until subjects responded. In Experiment 3, cues preceded the target displays by 100 ms and the target displays were exposed for only 100 ms. Cues preceded the displays to allow the transients associated with them to pull attention to the cue (cf. Yantis & Jonides, 1984). The displays were exposed briefly (cue plus target duration was only 200 ms) to prevent eye movements.

The experiments were very similar in design and procedure, so they will be described in one Method section.

Method

Subjects. Each experiment used a separate group of eight subjects recruited from the university population. Each subject served in four sessions and was paid $4 per session. All subjects reported normal or corrected vision. Subjects in Experiments 1 and 3 were screened for red–green color blindness with the Ishihara (1987) test at the beginning of the session.

Apparatus and stimuli. Stimuli were displayed on IBM 8513 VGA monitors controlled by IBM PS/2 Model 50 computers. Viewing distance was held constant by head rests at 45 cm. The stimuli were diamond-shaped arrays of four characters. In Experiments 1 and 3, the characters were capital O's colored red (IBM 12) or green (IBM 10). In Experiment 2, the characters were capital I's or lower-case l's, which were identical except for the small cross-bar at the top. The array was centered on the screen and spanned 2.7 × 2.7 cm or 3.43 × 3.43 degrees of visual angle. Defined in terms of the IBM text screen, which is a 24-row × 80-column matrix, the characters appeared in positions 11,40; 13,35; 13,45; and 15,40.

The cues were two adjacent horizontal dashes (ASCII 45) or one vertical dash (ASCII 124) presented in white (IBM 15). The two horizontal dashes appeared as long as the one vertical dash. Horizontal cues appeared 0.8, 1.8, 2.8, and 3.8 cm above the top target character or below the bottom target character. Vertical cues appeared 0.7, 1.9, 3.1, and 4.2 cm to the right of the rightmost target character or to the left of the leftmost target character. Averaging over vertical and horizontal cues, the distances were .95, 2.23, 3.75, and 5.08 deg of visual angle in the next-to condition and 4.38, 5.66, 7.18, and 8.51 in the opposite condition. As described above, distance is ambiguous in the clockwise and counterclockwise condition, depending on how subjects move attention. The greatest distance would be a city-block measure, in which subjects move attention from the cue to the center of the array and then at a right angle to the target. In that case, the distances would be the same as in the opposite condition. The shortest distance would be Euclidean, in which subjects move attention along a straight line from the cue to the target. In that case, the distances would be 3.19, 4.30, 5.73, and 7.01 degrees of visual angle. Regardless of how distance is measured, the shortest distances for the opposite, clockwise, and counterclockwise conditions are shorter than the longest distance for the next-to condition.

Target characters were assigned to positions randomly with the constraint that each array contained two of each type of character (i.e., 2 red and 2 green O's or 2 I's and 2 l's) and that each character was cued equally often from each distance (i.e., red was cued as often as green; I as often as l). The distractor that was across the array from the target always had a value that was opposite the target value (e.g., if the target was red, the distractor would be green). Two targets, 4 cue positions and 4 distances were combined factorially to produce 32 different trial types. There were 18 replications of each trial type each session for a total of 576 trials. The first 64 trials involved two replications of the 32 basic trial types in random order, to serve as practice. The remaining 512 trials were divided into two sets of 256 in which each trial type occurred exactly eight times in random order.

Subjects responded on numerical keypad on the right hand side of the computer keyboard. They pressed either the "8" and "2" key (top center and bottom center) or the "4" and "6" key (middle left and middle right) to register their responses. These keys were chosen to manipulate the spatial compatibility of stimuli and responses. Analyses of the data showed no evidence of compatibility effects in these experiments. The computer timed response latency in ms and synchronized stimulus presentation with the resetting of the raster scan.

Procedure. Each trial began with a fixation point exposed in the center of the screen (position 13,40) for 500 ms. Then, in Experiments 1 and 2, the cue and array display appeared and remained on until the subject responded. After the response, the screen went blank for a 1500-ms intertrial interval. In Experiment 3, the fixation point was followed by a 100-ms exposure of the cue and then a 100-ms exposure of the cue and the array simultaneously. Then the display terminated and the screen went blank. When the subject responded, a 1500-ms intertrial interval began, during which the screen remained blank.

Cuing conditions were blocked by session; each subject received a different cuing condition each day. The order of cuing conditions over days was determined by a balanced 4 × 4 Latin Square, with 2 subjects receiving each order. Half of the subjects responded by pressing the "8" and "2" keys (vertical mapping) and half responded by pressing the "4" and "6" keys (horizontal mapping). Assignment to mapping conditions was orthogonal to assignment to orders of cuing conditions.

Subjects were told that their task would be to report the color (Experiments 1 and 3) or identity (Experiment 2) of forms displayed on the computer screen. They were shown a picture of a typical display with a cue and told how to respond to it according to the current cuing condition. They were told that position of the cue and the target would vary from trial to trial and so would the distance between them. They were told to select the target that stood in the instructed relation regardless of position and distance. Then the response mapping conditions were described. Subjects were allowed to move the keyboard to a comfortable position (most moved it to the left so the numeric keypad was under the screen). They were told to maintain its orientation with respect to the computer, to keep the vertical mapping vertical and the horizontal mapping horizontal. They were told to rest the index fingers of their right and left hands lightly on the keys at all times. In the vertical mapping condition, they rested the right index finger on the "8" key and the left index finger on the "2." Finally, they were told to rest their foreheads lightly against the headrest to maintain a constant viewing distance throughout the experiment. Once the instructions were understood, the trials began. The computer paused every 96 trials to allow subjects a brief rest.

On subsequent sessions, the general instructions were reviewed briefly and the cuing condition for that session was described in detail, using the picture of a typical display to explain what to do.

In each experiment, mean reaction time was computed for each subject for each distance in each cuing condition. The means were analyzed in a 2 (group: horizontal vs. vertical mapping) × 4 (target position) × 4 (distance) × 4 (cue type) analysis of variance (ANOVA).

Results

Experiment 1: Color. Mean reaction times across subjects appear in Fig. 2. The figure reveals strong effects of cue type. Next-to was fastest (mean = 573 ms), followed by Opposite (753 ms), with clockwise (783) and Counterclockwise (810) slowest. These effects are contrary to the distance hypothesis: Subjects were faster with the longest and shortest distances (i.e., Opposite and Next-to conditions) than with intermediate distances (i.e., Clockwise and Counterclockwise conditions). Moreover, there were strong effects of cuing when distance was equivalent (i.e., the longest distance in Next-to vs the shortest in Opposite vs the second shortest in Clockwise and Counterclockwise).

There were weak effects of distance within each cue type. Reaction

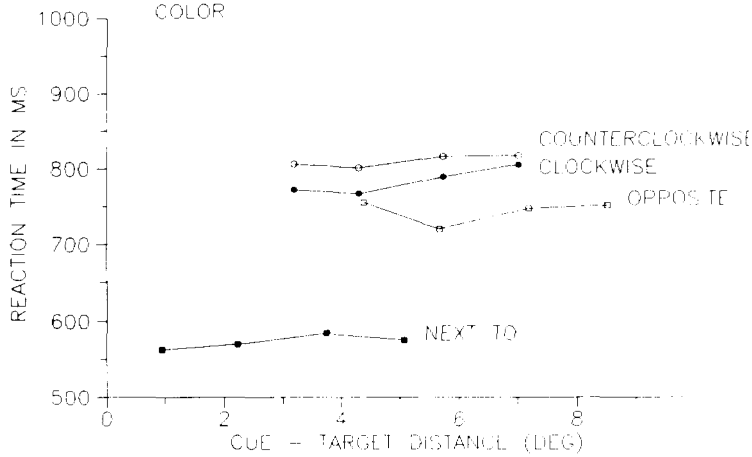

FIG. 2. Mean reaction time for each cue condition as a function of distance between cue and target in Experiment 1.

time tended to increase with distance. This may reflect acuity factors; if subjects were centrally fixated, the farthest cue would appear 5 degrees off the center of the fovea.

These effects were confirmed in ANOVA: The main effect of cue type was highly significant, $F(3,18) = 15.93$, $p < .01$. Fisher's Least Significant Difference (LSD) test revealed significant differences ($p < .05$) between Next-to and the other conditions but no significant differences between Opposite, Clockwise, and Counterclockwise. The main effect of distance was not significant, $F(3,18) = 2.31$, nor was the interaction between cue type and distance, $F(9,54) < 1$.

Cue type interacted significantly with target position, $F(9,54) = 4.55$, $p < .05$, probably reflecting differences in the perceptibility of cues in the horizontal versus vertical axis. Vertical cues extended further into the periphery than horizontal cues (5.3 vs 4.8 degrees of visual angle), so they were probably less perceptible than horizontal cues (because the retina's acuity gradient is roughly circular; Anstis, 1974). The data were consistent with this hypothesis: In the Next-to and Opposite conditions, subjects were faster with targets on the left and right, which were cued from the horizontal periphery, than the targets on the top and bottom, which were cued from the vertical periphery (mean difference = 37 ms). In the Clockwise and Counterclockwise conditions, subjects were faster with targets in the top and bottom positions, which were cued from the horizontal periphery, than with targets in the left and right positions, which were cued from the vertical periphery (mean difference = 22 ms).

Accuracy was high, averaging 92%. Accuracy varied between cuing conditions, averaging 95% in Next-to, 87% in Opposite, 92% in Clockwise, and 92% in Counterclockwise. Accuracy decreased slightly with distance, averaging 92, 91, 91, and 91% from the shortest to the longest distance.

Experiment 2: I versus l. Mean reaction times for each combination of cue type and distance are plotted in Fig. 3. The reaction times were longer than those in Experiment 1, reflecting the more difficult discrimination, but the pattern was essentially the same. Cue type produced strong effects. Next-to was fastest (mean = 750 ms), followed by Opposite (885 ms). Clockwise (968 ms) and Counterclockwise (976 ms) were slowest and not very different from each other. The data contradicted the distance hypothesis, in that intermediate distances (Clockwise and Counterclockwise conditions) produced longer reaction times than the shortest (Next-to) and longest (Opposite) distances. There was a small effect of distance within each cuing condition, probably reflecting acuity factors.

These effects were confirmed in ANOVA: The main effect of cue type was highly significant, $F(3,18) = 19.05$, $p < .01$. Fisher's LSD test revealed significant differences ($p < .05$) between Next-to and the other

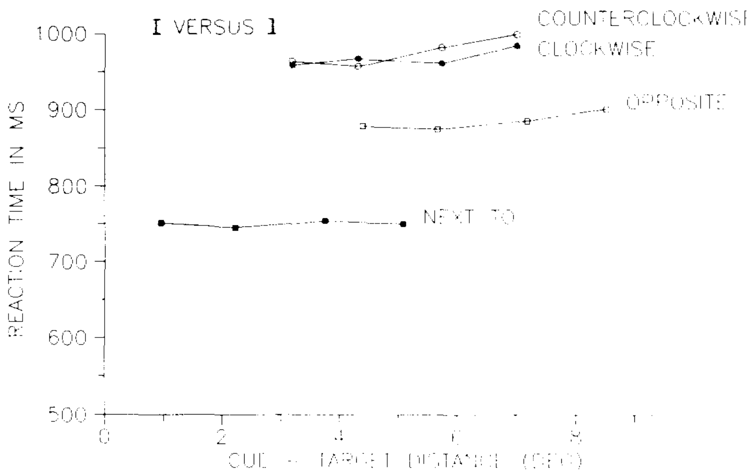

FIG. 3. Mean reaction time for each cue condition as a function of distance between cue and target in Experiment 2.

conditions and between Opposite and Clockwise and Counterclockwise. Clockwise and Counterclockwise did not differ significantly from each other. The main effect of distance was also significant, $F(3,18) = 5.20$, $p < .01$, as was the interaction between cue type and distance, $F(9,54) = 2.53$, $p < .05$. As in Experiment 1, cues presented above and below the array produced slower reaction times than cues presented left and right of the array, probably because of distance and acuity effects.

Accuracy was high, averaging 94%. Accuracy varied between cuing conditions, averaging 96% in Next-to, 91% in Opposite, 95% in Clockwise, and 94% in Counterclockwise. Accuracy did not vary with distance, averaging 94% at each distance.

Experiment 3: Color with brief exposure. Mean reaction times in each combination of cue type and distance are plotted in Fig. 4. The pattern of data is essentially the same as in Experiments 1 and 2: Cue type had strong effects and distance had only weak effects. Next-to was fastest (mean = 597 ms), followed by Opposite (713 ms), and then Clockwise (818 ms) and Counterclockwise (814 ms). The data were inconsistent with the distance hypothesis because reaction time did not increase monotonically with distance.

These effects were confirmed in ANOVA: The main effect of cue type was highly significant, $F(3,18) = 13.06$, $p < .01$. Fisher's LSD test revealed significant differences ($p < .05$) between Next-to and the other conditions and between Opposite and Clockwise and Counterclockwise. Clockwise and Counterclockwise did not differ significantly from each other. The main effect of distance approached significance, $F(3,18) =$

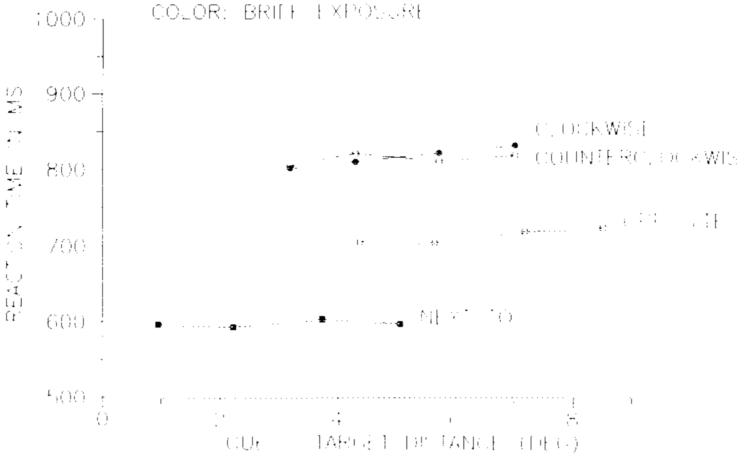

FIG. 4. Mean reaction time for each cue condition as a function of distance between cue and target in Experiment 3.

2.67, $p < .10$. The interaction between cue type and position was significant, $F(9,54) = 4.23$, $p < .01$, reflecting the asymmetry of the cue positions as in Experiments 1 and 2. No other effects were significant.

Accuracy was high, averaging 94%. It varied slightly between cuing conditions, averaging 95% in Next-to, 94% in Opposite, 94% in Clockwise, and 92% in Counterclockwise. Accuracy was relatively unaffected by distance, averaging 94, 94, 94, and 93% from the shortest to the longest distance.

Discussion

The results were the same in all three experiments: Reaction time varied substantially with cue type and only a little with distance. The results refute the hypothesis derived from serial visual routines, that distance is the only thing that matters. In each experiment, intermediate distances (in Clockwise and Counterclockwise conditions) produced longer reaction times than the longest (Opposite) and shortest (Next-to) distances. Moreover, in each experiment, reaction time was strongly affected by cue type when distance was controlled. Clearly, the effects of cue type cannot be explained by distance alone.

The effects of cue type are important because they demonstrate differences between push cues that cannot be attributed to differences in stimulus conditions. The cues and targets were the same in all conditions. Only the cuing relation changed. The results must be explained in terms of the representations and processes that computed the relation.

The effects of cue type may reflect reference-frame computations:

Next-to required only the origin and the scale of the reference frame to be specified, *opposite* required the origin and one axis, and *clockwise* and *counterclockwise* required the origin and two axes. The effects of cue type may also reflect spatial index computations: Next-to required the cue and the target be indexed, whereas *opposite, clockwise,* and *counterclockwise* required the cue, the target, and the array. Most likely, both computations contributed to the differences between cuing conditions.

This uncertainty illustrates the problems that result from working within one literature: Experiments 1–3 used cues that were typical in the attention literature—bar markers outside the array—and used cuing relations that were either typical (next to, opposite) or possible (*clockwise, counterclockwise*) with those cues, with no regard for their semantics. Consequently, their semantics were complicated and confounded. The subsequent experiments used cuing relations typical of the linguistic literature with more straightforward semantics. The cues were objects in the center of the display, not peripheral bar markers, and the cuing relations were (usually) two-argument deictic or intrinsic relations (i.e., *above, below, left,* and *right*). These relations manipulate reference frame computation while keeping spatial indexing the same.

In Experiments 1–3, there were weak distance effects within each cuing relation. The slight increase in reaction time with eccentricity probably reflected the corresponding decrease in acuity (Anstis, 1974; Eriksen & Murphy, 1987). The results suggest that the relation between the cue and the target was computed with (parallel) spatial templates (Herskovits, 1986; Langacker, 1986; Logan & Sadler, 1995) rather than serial visual routines (Ullman, 1984).

EXPERIMENTS 4–6: CUING WITH BASIC AND DEICTIC RELATIONS

Experiments 4–6 tested for reference frame effects with spatial indexing controlled and tested the psychological validity of the linguistic distinction between basic and deictic relations in attentional cuing. Subjects were presented with the four-item diamond-shaped arrays used in Experiments 1–3 but the cue was a word or a symbol that appeared in the center of the display rather than a bar marker, specifying the position of the target with a deictic or a basic relation.

Experiments 4 and 5 used the deictic relations, *above, below, left,* and *right*. Each relation takes two arguments and requires two spatial indexing operations. Moreover, each relation refers to a similar region of space, oriented differently with respect to the reference frame (Logan & Sadler, 1995; Miller & Johnson-Laird, 1976). Differences between relations are attributable to reference frame computations, not spatial indexing or computing the relation with respect to the reference frame.

In Experiment 4, the cues were the words ABOVE, BELOW, LEFT, and RIGHT that named the relations. In Experiment 5, the cues were the letters A, B, L, and R that were associated with the relations. The conceptual-frame hypothesis should be supported in these experiments. Parts of space addressed by *above* and *below* should be more accessible than parts addressed by *left* and *right* for reasons described earlier (also see Bryant et al., 1992; Clark, 1973; Franklin & Tversky, 1990).

Experiment 6 tested the distinction between deictic and basic relations, cuing positions with digits that were mapped arbitrarily onto space (i.e., 1 was the top position, 2 was the right position, 3 was the bottom position, and 4 was the left position; for a similar cuing manipulation, see Eriksen and Collins, 1969). The digits refer to positions independently—each digit represents a different *there*. Thus, they represent position in terms of basic relations, allowing subjects to access one position without reference to the others. Basic relations do not require a reference frame to be specified, so the conceptual frame hypothesis should not be confirmed in this experiment.

The three experiments involved very similar designs and procedures so they will be described in one Method section.

Method

Subjects. Each experiment used a separate group of 8 subjects who were recruited from the Introductory Psychology subject pool or from the general university population. Introductory Psychology subjects received course credit for participation; the others received $3.50. Each subject served in a single session. All subjects had normal or corrected vision and no subjects were red-green color blind, as assessed by the Ishihara (1987) color-blindness test.

Apparatus and stimuli. The stimuli were presented on Amdek Model 720 color monitors controlled by IBM PC/XT or AT computers. Viewing distance was not constrained, varying between 40 and 60 cm. The stimuli were diamond-shaped arrays of four capital O's colored red (IBM 12) or green (IBM 10) centered on the screen. The characters appeared in positions 9,40; 13,31; 13,49; and 17;40 of the 24-row × 80-column IBM text screen.

The cues were presented in white (IBM 15), varying in form between experiments. In Experiment 4, the cues were the words ABOVE, BELOW, LEFT, and RIGHT presented in the center of the diamond-shaped array, beginning at position 13,38. In Experiment 5, they were the letters A, B, L, or R, representing *above, below, left,* and *right,* respectively. The cue letters appeared in the center of the diamond-shaped array, in position 13,40. Experiment 6 used the digits 1–4, where 1 referred to the top position, 2 to the right position, 3 to the bottom position, and 4 to the left position. The digit cue appeared in the center of the array, in position 13,40.

Target characters were assigned to positions randomly with the constraint that each array would contain two of each type of character (i.e., 2 red and 2 green O's) and that each character was cued equally often. The distractors across the array from the target were not constrained. Two targets and 4 cue positions were combined factorially to produce 8 different trial types, which were replicated 72 times in a session to produce 576 trials. The 576 trials were divided into two sets of 288 in which each trial type occurred exactly 36 times in random order.

Subjects responded on the numerical keypad, pressing either the "8" and "2" key or the "4" or "6" key to register their responses.

Procedure. Each trial began with a fixation point exposed in the center of the screen (position 13,40) for 500 ms. Then the cue and array display appeared and remained on until the subject responded. After the response, the screen went blank for a 1500-ms intertrial interval.

Cuing conditions varied between experiments. Experiments 4 and 5 used deictic relations and Experiment 6 basic relations. Colors were mapped onto responses using the horizontal and vertical rules from Experiments 1–3. Mapping varied within and between subjects. Each subject performed half of the experiment with a vertical mapping and half with a horizontal mapping. Half of the subjects had vertical mapping first and horizontal second, and half had the opposite. There were two between-subject mapping conditions: Half of the subjects had red on the top and green on the bottom (vertical) and red on the left and green on the right (horizontal), and half of the subjects had red on the bottom and green on the top (vertical) and red on the right and green on the left (horizontal). There were two subjects in each combination of mapping condition and order.

Subjects were told that their task would be to report the color of forms displayed on the computer screen. They were shown a picture of a typical display and cue and told how to respond to it. They were told that the cue and (consequently) the position of the character to be reported would vary from trial to trial. Then the first response mapping condition was described (horizontal or vertical). Subjects were told to rest the index fingers of their right and left hands lightly on the keys at all times. Once the instructions were understood, the trials began. The computer paused every 96 trials to allow subjects a brief rest. After the 288th trial, the computer displayed a message asking the subject to call the experimenter and the second mapping condition was described (vertical or horizontal).

In each experiment, mean reaction times were computed for each combination of conditions and subjected to a 2 (horizontal vs vertical mapping) × 2 (target color) × 4 (target position) ANOVA.

Results

Experiment 4: Above, below, left, and right. Mean reaction times for each position, averaged across target color and mapping condition, appear in Fig. 5. Reaction times were strongly influenced by position. Reaction times for *above* and *below* were 232 ms faster than reaction times for *left* and *right,* on average, confirming Clark's (1973) hypothesis and replicating Franklin and Tversky's (1990) and Bryant et al.'s (1992) results with visual displays instead of imagined displays.

The only significant ANOVA effect was the main effect of position, $F(3,21) = 10.19, p < .01$. A contrast comparing *above* and *below* with *left* and *right* was highly significant, $F(1,21) = 29.23, p < .01$.

The accuracy data corroborated the reaction time effects. Mean percent correct was 97, 94, 96, and 95% for the top, right, bottom, and left positions, respectively.

Experiment 5: A, B, L, and R. Experiment 5 was conducted to see whether the results of Experiment 4 would replicate when subjects did not have to read words to identify the cue. The cues were the letters A, B, L, and R, representing the relations *above, below, left,* and *right.* Mean reaction times averaged across target color and mapping condition appear in Fig. 5. Reaction times were slightly longer than those in Experiment 4

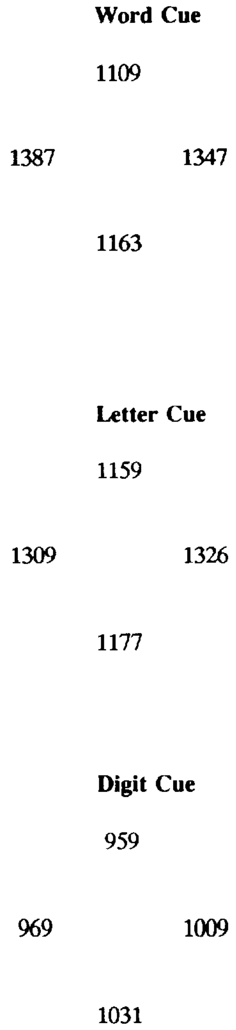

FIG. 5. Mean reaction time for cued position in Experiment 4 (Word Cue), Experiment 5 (Letter Cue), and Experiment 6 (Digit Cue).

but the pattern was essentially the same. Subjects were strongly influenced by position, responding to A and B 150 ms faster than to L and R.

The only significant effect was the main effect of position, $F(3,21) = 5.17$, $p < .01$. A contrast comparing A and B with L and R was highly significant, $F(1,21) = 15.24$, $p < .01$.

The accuracy data corroborated the reaction time effects. Mean percent correct was 93, 90, 94, and 92% for the top, right, bottom, and left positions, respectively.

Experiment 6: Arbitrary digits. In this experiment, the cues were digits referring to positions in an arbitrary but sensible manner. Subjects could perform the task using basic relations, without imposing a reference frame on the display and computing the target position with respect to the reference frame. They could go directly to the position specified by the cue (Garnham, 1989; Pylyshyn, 1989; Pylyshyn & Storm, 1989).

Mean reaction times, averaged across target color and mapping condition, appear in Fig. 5. The results differed from the previous experiments. Reaction time varied with position. The top and left positions (positions 1 and 4) were responded to faster than the right and bottom positions (positions 2 and 3), but there was no advantage to the above-below axis. Mean reaction times for the top and bottom positions were 6 ms slower than those for the left and right positions.

These results were confirmed by ANOVA: The main effect of position was barely significant, $F(3,21) = 3.07$, $p = .05$. A contrast comparing positions 1 and 3 with 2 and 4 (top and bottom with left and right) was not significant, $F(1,21) < 1$. The interaction between position and target color was significant, $F(3,21) = 5.35$, $p < .01$, but not easily interpretable. There was no advantage for top and bottom positions (over left and right positions) for either target color.

The accuracy data corroborated the reaction time effects. Mean percent correct was 96% for all four positions.

Discussion

Experiments 4–6 were designed to test for reference frame effects with spatial indexing controlled and to test the validity of the distinction between basic and deictic relations in attentional cuing. Experiments 4 and 5 cued target position deictically and found a large advantage for the vertical axis, soundly rejecting the equal availability hypothesis. The advantage of *above* and *below* over *left* and *right* confirms the conceptual frame hypothesis (Clark, 1973) and replicates the results of Franklin and Tversky (1990) and Bryant et al. (1992) with visual rather than imaginal displays. Experiment 6 cued target position with basic relations, arbitrarily labeling positions in the display, and found no advantage of the vertical axis over the horizontal, confirming the hypothesized difference between basic and deictic relations. The equal-availability hypothesis should be violated with deictic relations but not with basic relations.

EXPERIMENT 7: CONCEPTUAL VERSUS ENVIRONMENTAL REFERENCE FRAMES

The advantage of the vertical over the horizontal axis in Experiments 4 and 5 supports the conceptual frame hypothesis but is not the strongest test. The conceptual frame hypothesis predicts that the same part of space can be easy or hard to access depending on its relation to the reference frame. Experiments 4 and 5 confounded relations with specific locations (e.g., *above* always referred to the top position) and so did not test this aspect of the hypothesis. A stronger test would move the reference frame around so that the same parts of space could be accessed by different relations. Experiment 7 was designed to provide such as test.

In Experiment 7, subjects were presented with diamond-shaped displays of colored O's with a word cue (ABOVE, BELOW, LEFT, or RIGHT) in the center. Their reference frame was moved (rotated) around the display by telling them to treat different parts of the display as the top. One group of subjects was told to treat the left side as the top, one was told to treat the right side as the top, and one was told to treat the bottom of the display as the top. This procedure unconfounds relations and locations; *above* refers to the left, bottom, and right position in different groups.

The theory underlying the conceptual frame hypothesis assumes that subjects can adjust the orientation of the reference frame voluntarily. If they can do so in practice, there should be an advantage of the conceptual vertical over the conceptual horizontal in all conditions, even when they depart from the environmental vertical and horizontal. The alternative is the *environmental frame* hypothesis, which predicts an advantage for the gravitational or allocentric vertical over the gravitational or allocentric horizontal independent of the conceptual frame subjects impose on the display. The top and bottom positions in the display (defined in terms of gravitational coordinates) should be processed faster than the left and right positions in all conditions.

Method

Subjects. The experiment used three separate groups of eight subjects recruited from the Introductory Psychology subject pool and the general university population. Introductory Psychology subjects received course credit for participating; others received $3.50. All subjects had normal or corrected vision and all passed the Ishihara (1987) color-blindness test. Each subject served in a single 1-h session.

Apparatus and stimuli. These were the same as in Experiment 4.

Procedure. This was the same as in Experiment 4, except that subjects were told to treat either the left side, the right side, or the bottom side of the display as the top.

Results

Mean reaction times, calculated for each position in each condition of the experiment, are presented in Table 1. To assess the importance of environmental versus conceptual reference frames, reaction times were averaged over the conceptual vertical axis (i.e., averaging *above* and *below*) and over the conceptual horizontal (i.e., averaging *left* and *right*). These means are plotted in Fig. 6 along with the corresponding data

TABLE 1

Mean Reaction Time and Percent Correct as a Function of Target Position (Defined Conceptually, Not Environmentally) for Each Orientation Group in Experiment 7

	Above	Below	Left	Right
Top at right	1129	1166	1290	1300
	97	95	93	94
Top at left	1206	1178	1330	1340
	94	94	91	93
Top on bottom	1480	1469	1531	1564
	95	95	93	94

points from Experiment 4. The figure shows a monotonic increase in reaction as the conceptual vertical departed from the environmental. Reaction times were fastest when environmental and conceptual vertical corresponded, slowest when they opposed, and intermediate when they were orthogonal. However, there was an advantage of conceptual vertical over conceptual horizontal in each condition. It appeared to diminish as the conceptual vertical departed from the environmental, but it was still quite large when the (environmental) bottom of the display was defined as the top (mean = 73 ms).

These results were confirmed in a 3 (group: top at left, top at right, and top at bottom) × 2 (target color) × 2 (mapping horizontal or vertical) × 4 (target position) ANOVA on mean reaction times. Target position was defined with respect to the instructed coordinates (e.g., in the top-at-left

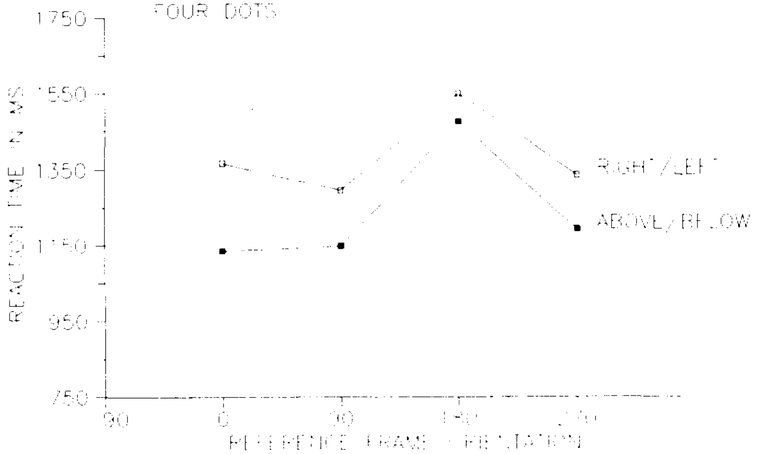

FIG. 6. Mean reaction time for responses to Above and Below cues versus Right and Left cues as a function of the orientation of the conceptual top of the display in Experiment 7 (Above, Below, Right, and Left are defined relative to the conceptual top of the display).

condition, the position on the left side of the display was defined as *above*). The main effect of orientation group approached significance, $F(2,21) = 2.95$, $p < .10$. A planned comparison revealed significant differences between top-at-bottom and the average of top-at-left and top-at-right, $F(1,21) = 5.81$, $p < .05$. The main effect of position was significant, $F(3,63) = 12.44$, $p < .01$. Contrasts comparing conceptual vertical with conceptual horizontal were significant in each group, F's$(1,63) = 18.44$, 13.15, and 4.49 for top-at-left, top-at-right, and top-at-bottom, respectively, all p's $< .05$. The interaction between orientation and target position was not significant, $F(6,63) < 1$.

Mean percent correct scores appear in Table 1. Accuracy was high and correlated negatively with reaction time. Accuracy was higher for *above* and *below* (mean = 95%) than for *left* and *right* (mean = 93%).

Discussion

The results support the conceptual frame hypothesis. They show that the conceptual frame of reference is more important than the environmental frame of reference in cuing attention with deictic relations. There was an advantage of the conceptual vertical over the conceptual horizontal in each condition, even when conceptual and environmental reference frames were orthogonal, when the left and right sides of the display were defined as the top. Thus, the advantage of the vertical over the horizontal in Experiments 4 and 5 was most likely due to the conceptual frame of reference rather than the environmental.

In the present experiment, the axis that showed the advantage was always the axis that was labeled by the experimenter. For example, when the experimenter pointed to the left side of the display and told the subject to treat it as the top of the display, then positions on the left-right (environmental) axis showed an advantage. I conducted a further experiment to rule out the hypothesis that the advantage was due to labeling. Two groups of eight subjects were tested. One was told to treat the top of the display as if it were the right side (thus the left side of the display was the conceptual top) and the other was told to treat the top of the display as if it were the left side (thus the right side of the display was the conceptual top). If labeling produced the differences seen in Experiment 7, then both groups should show an advantage of the environmental vertical over the environmental horizontal. However, if the conceptual frame produced the differences, then both groups should show an advantage of the environmental horizontal over the environmental vertical. The results were consistent with the conceptual frame hypothesis and not with the labeling hypothesis. The conceptual vertical axis (environmental horizontal) showed an advantage over the conceptual horizontal (environmental vertical) in both groups, 93 ms for top-is-left-side, $F(1,21) = 9.81$, $p < .01$, and 39 ms for top-is-right-side, $F(1,21) = 3.56, p < .10$. These differences were slightly smaller than the ones observed in Experiment 7, but they were large enough to suggest that labeling cannot account for all of the effect.

In some respects, Experiment 7 is a conceptual replication of Experiments 1–3. In those experiments, subjects had to compute target position with respect to an axis that ran through the cue and the center of the display. Cue position varied randomly, so sometimes that axis was parallel to the environmental axis and sometimes it was orthogonal to it. Clockwise and Counterclockwise cues, which required subjects to compute left and right, were more difficult than Next-to and Opposite cues, which did not. The difference was observed at each cue position. It was independent of the orientation of the axis with respect to which parity was judged, just as the difference between *above-below* and *left-right* was (largely) independent of the orientation of the conceptual reference frame in Experiment 7.

The results of Experiment 7 contrast with what Franklin and Tversky (1990) found when they pitted the environmental reference frame against the subject's intrinsic reference frame. Whereas Experiment 7 showed the advantage of *above-below* over *left-right* was largely independent of orientation, Franklin and Tversky (1990) found that the advantage of *above-below* over *front-back* was eliminated by having subjects imagine themselves lying on their sides, orthogonal to gravity. However, in their experiments, the advantage of *above-below* over *left-right* was still significant when subjects imagined themselves lying down, so their results may not be so different from the present ones.

The present results also contrast with Levelt's (1984) claims about the importance of the gravitational upright in the computation of *above* and *below*. Levelt attributed some of the advantage of *above-below* over other axes to the usual coincidence of the gravitational and the egocentric or the intrinsic vertical. The present results show an advantage of *above-below* over *left-right* in all orientations, suggesting that coincidence of gravitational and egocentric or intrinsic axes is not necessary to produce the advantage. There was a strong effect of orientation, however: Reaction time was fastest when gravitational and egocentric or intrinsic axes coincided, and became longer as the difference between them increased (i.e., it was longer with 180 degree rotations than with 90 degree rotations). Perhaps there is some difficulty involved when the gravitational upright conflicts with the egocentric or intrinsic upright. However, that difficulty appears not to be related to the advantage of *above-below* over *left-right*. The effect of orientation might reflect processes involved in specifying the orientation of the reference frame with the display; the effect of *above-below* versus *left-right* might reflect subsequent processes that specify the left-right axis once the orientation of the above-below axis is specified (Corballis, 1988).

EXPERIMENT 8: DEICTIC CUING IN COMPLEX DISPLAYS

There are two problems with the simple four-element displays used in Experiments 4–7. First, cues were confounded with positions, which weakens the support the experiments provide for the conceptual frame hypothesis. That hypothesis predicts that the same positions will be hard or easy to access, depending on how they are related to the frame of reference that the subject imposes on the display. Experiment 7 removed the confound by manipulating the orientation of the reference frame between subjects, but within any group of subjects, cues were confounded with positions. The conceptual frame hypothesis should be tested more stringently. Second, the cue was presented simultaneously with the display and it is possible that some of the differences between cues were due to differences in reading the words or accessing their meaning. The fact that the same results occurred when initial letters replaced the word cues in Experiment 5 may allay this concern somewhat, but it is still a potential problem.

Experiment 8 was designed to overcome these problems. Subjects were presented with complex displays of nine colored O's. The nine O's formed the vertices of four diamond shapes that made up a large diamond shape. An example is presented in Fig. 7. The cue was an asterisk presented in the center of one of the smaller diamonds. The target display was preceded by an instruction display that contained a word centered in the screen that specified the relation to be computed between the asterisk and the target: ABOVE, BELOW, LEFT, or RIGHT. If the word

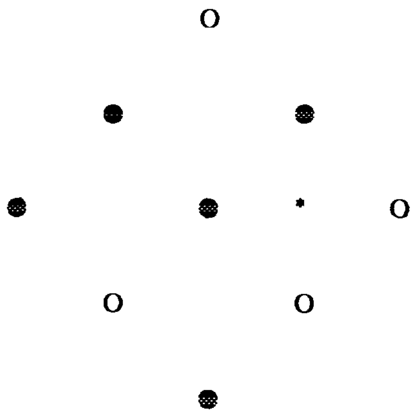

FIG. 7. Sample target display from Experiment 8 (not to scale).

ABOVE appeared, subjects were to report the color of the O that appeared above the asterisk ("dark" in Fig. 7). If the word RIGHT appeared, subjects were to report the color of the O that appeared to the right of the asterisk ("light" in Fig. 7). The asterisk could appear centered in any of the four quadrants, so any of the O's could be cued. Moreover, the same O could be cued in different ways. The central O, for example, could be cued by BELOW if the asterisk was centered in the top diamond, by ABOVE if the asterisk was centered in the bottom diamond, by RIGHT if the asterisk was centered in the left diamond, and by LEFT if the asterisk was centered in the right diamond.

These displays overcome the problems with Experiments 4–7: First, the construction of the target display removes the confound between cues and positions. The same positions can be accessed by different cues, so the conceptual frame hypothesis can be tested more stringently. Second, presenting the instruction display well before the target display separates the processes involved in reading the word and accessing its meaning from the processes involved in applying its meaning to the cue and the target.

A major purpose of Experiment 8 was to determine whether the origin of the reference frame can be moved around the display on a trial by trial basis. The theory assumes it can; the reference frame is a mechanism of attention that can be moved around space voluntarily just like spotlights and spatial indices. If the reference frame can be moved, the positions that are easy and hard to access should move along with it. Access should be easy if the target is on the current vertical axis and hard if it is on the current horizontal axis. The same target can be switched from easy access to hard and back again by moving the origin of the reference frame (e.g., if the target is in the center position and the cue is moved from the top or bottom position to the left or right position; see Fig. 7). More generally, the theory predicts that the conceptual frame hypothesis will be confirmed in each of the four quadrants; what the subject considers vertical should be easier than what the subject considers to be horizontal. If this prediction is not fulfilled, if reference frames cannot be moved voluntarily, it will be hard to think of them as mechanisms of attention.

Another major purpose of the experiment was to test the distinction between basic and deictic relations. The displays required subjects to first locate the asterisk cue and then locate the target with respect to it. By hypothesis, the cue could be located in terms of basic relations (i.e., not in terms of a reference frame or in terms of other objects) whereas the target must be located deictically with respect to a reference frame imposed on the cue. The effects of basic relations were assessed by comparing reaction times when the cue appeared in different quadrants, whereas the effects of deictic relations were assessed by comparing reaction times when the target appeared in different positions relative to the cue. If the distinction between basic and deictic relations is important in attentional cuing, there should be a strong advantage of the vertical axis over the horizontal when target position is assessed but no advantage of vertical over horizontal when the cue position is assessed.

Method

Subjects. Sixteen subjects were recruited from the Introductory Psychology subject pool or the general university population. Introductory Psychology subjects received course credit for participation; others received $3.50. All subjects had normal or corrected vision and all passed the Ishihara (1987) color blindness test. The experiment took 1 h.

Apparatus and stimuli. The stimuli were presented on Amdek Model 720 color monitors controlled by IBM PC/XT or AT computers. Viewing distance was not constrained, varying between 40 and 60 cm. The target displays consisted of nine capital O's. One was the target and the other eight were distractors. The target was red (IBM 12) or green (IBM 10). Half of the distractors were red and half were green. The cue was an asterisk (ASCII 42) presented in white (IBM 15). The targets and distractors were presented at the vertices of four adjacent diamonds that formed one large diamond (see Fig. 7). The cue appeared in the center of one of the diamonds. The positions for targets and distractors, defined in terms of the IBM 24-column × 80-row text screen were 5,40; 9,31; 9,49; 13,22; 13,40; 13,58; 17,31; 17,49; and 21,40. The cue positions were 9,40; 13,31; 13,49; and 17,40.

Instruction displays consisted of the word ABOVE, BELOW, LEFT, or RIGHT presented in white at the center of the screen (beginning at position 13,38). Fixation displays consisted of a period presented in the center of the screen (position 13,40).

Targets were assigned to positions randomly with the constraint that each color appear equally often in each target position (above, below, left of, and right of the asterisk) in each cue position (top, left, bottom, and right quadrant). Distractors were assigned to the nontarget positions randomly with the constraint that half were red and half were green. Four target positions, 4 cue positions, and 2 target colors were combined factorially to produce 32 different trial types, which were replicated 18 times to produce 576 trials. The 576 trials were divided into two sets of 288 trials in which the 32 basic trial types occurred nine times in random order.

Subjects responded on numerical keypad, pressing either the "8" and "2" or the "4" and "6" key to register their responses.

Procedure. Each trial began with a fixation point exposed for 500 ms. Then the instruction display appeared and remained on for 1000 ms. It was extinguished, and the target display was exposed until the subject responded. After the response, the screen went blank for a 1500-ms intertrial interval.

Colors were mapped onto responses using horizontal and vertical rules from Experiments 1–3. Half of the subjects used horizontal rules and half used vertical rules. Half of the subjects with horizontal rules had "red" on the top and "green" on the bottom and half had the opposite; half of the subjects with vertical rules had red on the right and green on the left and half had the opposite.

Subjects were told the sequence in which displays would appear and that their task would be to report the color of one of the forms in the target display. They were shown pictures of a typical instruction display and target display and told how to respond to it appropriately. They were told that the word, the position of the asterisk, and (consequently) the position of the target would vary from trial to trial. Then the mapping rules were described. Subjects were told to rest the index fingers of their right and left hands lightly on the keys at all times.

Once the instructions were understood, the trials began. The computer paused every 96 trials to allow subjects a brief rest.

Results

Mean reaction times averaged across target color and mapping condition are presented in Fig. 8. Reaction times are plotted in positions that correspond to the position of the target they were associated with on the display screen. Reaction times are plotted in normal font for displays in which the cue appeared in the top and bottom quadrants and in bold font for displays in which the cue appeared in the left and right quadrants. Reaction times in different fonts plotted close together reflect responses to the same stimulus addressed by different cues. For example, the four reaction times arranged in a diamond shape in the center of the figure reflect performance on the central item. The top reaction time of the four reflects performance when the cue appeared in the top quadrant and the item was addressed by BELOW. The reaction time on the right reflects performance when the cue appeared in the right quadrant and the item was addressed by LEFT.

Analysis of target position effects reveals an advantage for the vertical axis in all four quadrants. Subjects were faster to respond to targets *above* and *below* the cue than to targets *left of* and *right of* the cue regardless of where the cue appeared. This was true overall. Mean reaction time was 887 ms for the vertical axis and 992 for the horizontal axis, an advantage of 105 ms. It was true in each quadrant. The advantage of vertical over

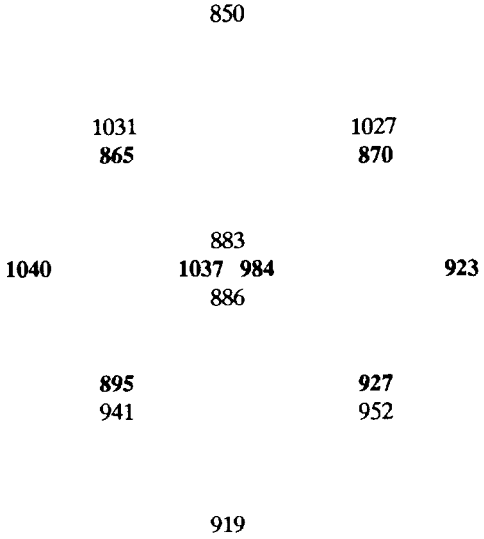

FIG. 8. Mean reaction time for each cue position and target position in Experiment 8.

horizontal was 163, 55, 44, and 159 ms for the top, right, bottom, and left quadrants, respectively. It was true in each array position that could be referred to by two or more relations (see Fig. 8). In the central position, for example, responses with ABOVE and BELOW as cues were 126 ms faster than responses with LEFT and RIGHT as cues even though exactly the same item was cued.

Analysis of cue position effects revealed no advantage for the vertical axis over the horizontal. Mean reaction times averaged across target position were 948, 926, 925, and 959 ms for the top, right, bottom, and left quadrants, respectively. Reaction times in the top and bottom quadrants were only 6 ms faster than reaction times in the left and right quadrants. Cue position did not show the same effects as target position.

These conclusions were confirmed in a 4 (group: mapping) × 4 (target position) × 4 (cue position) ANOVA on the mean reaction times. There was a significant main effect of cue position, $F(3,36) = 3.18$, $p < .05$, a significant main effect of target position, $F(3,36) = 9.04$, $p < .01$, and a significant interaction between them, $F(9,108) = 3.77$, $p < .01$. A contrast comparing the top and bottom cue positions with the left and right cue positions was not significant, $F(1,36) < 1$, whereas a contrast comparing the top and bottom target positions with the left and right target positions was highly significant, $F(1,36) = 25.22$, $p < .01$. These results document the advantage of the vertical axis over the horizontal axis with deictic relations (cue-target relations) and the lack of advantage for basic relations (cue position). The critical interval for the .05 significance level for post-hoc comparisons within the interaction between cue position and target position was 57 ms, according to Fisher's LSD test.

Percent correct scores are presented in Fig. 9 in the same format as the reaction times in Fig. 8. In general, accuracy was high and did not show any trends that compromised the interpretation of the reaction time results.

Discussion

This experiment showed a priority of the vertical axis over the horizontal in complex displays, replicating the results of Experiments 4 and 5, replicating Franklin and Tversky (1990) and Bryant et al. (1992), and confirming the conceptual frame hypothesis. In these displays, cues were not confounded with positions. In many cases, more than one cuing relation could apply to a single position (e.g., the center position). In each of these cases, cues that referred to the vertical axis were better than cues that referred to the horizontal axis.

This experiment separated the processing of the word from the application of the relation specified by the word to the target display. The word appeared in an instruction display 1000 ms before the target display ap-

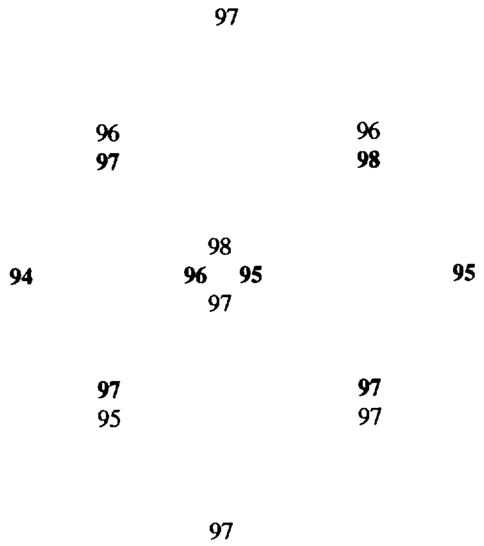

FIG. 9. Mean accuracy scores (percent correct) for each cue position and target position in Experiment 8.

peared. This reduced reaction time, relative to Experiments 4 and 5, but the vertical axis was still better than the horizontal. This suggests that the advantage of the vertical over the horizontal reflects the processing involved in applying the cuing relation to the display (i.e., locating the target relative to the cue) rather than accessing the lexical representation of the word or accessing the semantic representation of the relation. Lexical and semantic access ought to have been complete by the time the display appeared.

This experiment also separated the processes involved in locating the cue from those involved in locating the target relative to the cue. The former were reflected in the effects of cue position (i.e., which quadrant the asterisk appeared in) and the latter were reflected in the effects of target position (i.e., whether the target appeared *above, below, right of,* or *left of* the cue). There was no difference between horizontal and vertical axes in the cue position effect, suggesting that cue location was represented by a basic relation rather than a deictic relation. The strong advantage of vertical over horizontal in the target position effect suggests that target location was represented by a deictic relation.

EXPERIMENT 9: CONCEPTUAL REFERENCE FRAMES WITH COMPLEX DISPLAYS

The conceptual frame hypothesis says that the same display positions may be easy or hard to access, depending on their relation to the spatial reference frame the subject imposes on the display. Experiment 7 tested this hypothesis by rotating the reference frame; Experiment 8 tested it by translating the axes of the reference frame (i.e., to each quadrant of the display). Experiment 9 combined these two manipulations to provide the most stringent test of the conceptual frame hypothesis. Three different groups of subjects were presented with the 9-element displays used in Experiment 8 and were asked to report the color of the target specified by the cue and the instruction display. One group of subjects was told to treat the left side of the display as the top, one group was told to treat the bottom of the display as the top, and one group was told to treat the right side as the top. Otherwise, the design and procedure was the same as in Experiment 8.

If the conceptual frame hypothesis is valid, there should be an advantage of the conceptual vertical over the conceptual horizontal even when they disagree with the environmental vertical and horizontal.

Method

Subjects. Three separate groups of 16 subjects were recruited from the Introductory Psychology subject pool and the general university population. Introduction Psychology subjects received course credit for participation; the others received $3.50. All subjects had normal or corrected vision and all passed the Ishihara (1987) color-blindness test.

Apparatus and stimuli. These were the same as in Experiment 8.

Procedure. The procedure was the same as in Experiment 8, except that one group was told to treat the left side of the display as the top, one group was told to treat the bottom as the top, and one group was told to treat the right side as the top.

Results

Mean reaction times were calculated for each cue position and each target position. The means for each orientation group appear in Table 2. The means from *above* and *below* were averaged together across cue position as were the means from *right* and *left*. These means are plotted in Fig. 10 along with the corresponding means from Experiment 8. Reaction time increased as the conceptual top of the display departed from the environmental top, but the advantage of *above-below* over *right-left* was apparent at each orientation.

These conclusions were confirmed in a 3 (orientation group: top at left, top at right, and top at bottom) × 4 (mapping group) × 4 (cue position) × 4 (target position) ANOVA on the mean reaction times. As in Experiment 7, cue position and target position were defined with respect to the instructed orientation (e.g., in the top-at-left condition, the cue in the leftmost display position was coded as "top" and the leftmost target position was coded as "top"). The main effect of orientation was not significant, nor was a contrast comparing the top-at-bottom condition with the mean of the top-at-left and top-at-right conditions, both F's < 1.0.

FIG. 10. Mean reaction time for responses to Above and Below cues versus Right and Left cues as a function of the orientation of the conceptual top of the display in Experiment 9 (Above, Below, Right, and Left are defined relative to the conceptual top of the display).

There was a significant main effect of cue position, $F(3,108) = 8.20$, $p < .01$, but a contrast comparing the conceptually vertical positions with the conceptually horizontal positions was not significant, $F < 1.0$. The main effect of target position was significant, $F(3,108) = 9.38$, $p < .01$. A contrast comparing conceptually vertical target positions with conceptually horizontal ones was highly significant, $F(1,108) = 27.42$, $p < .01$. The same contrast was significant in each orientation, F's$(1,108) = 8.82$, 15.89, and 5.67, for top-at-left, top-at-bottom, and top-at-right, respectively, p's $< .01$. Similar contrasts comparing conceptual vertical and horizontal cue positions were not significant, all F's < 1.0. The interaction between cue position, target position, and orientation group was significant, $F(18,324) = 1.98$, $p < .05$, but nothing in the interaction compromised the interpretation of the target position main effect (i.e., there was an advantage of the conceptually vertical over the conceptually horizontal in each cell of the interaction).

The accuracy data for each combination of cue position and target position in each orientation group appear in Table 2. There was nothing in the accuracy data to compromise the conclusions drawn from the reaction times.

Discussion

This experiment showed an advantage of the conceptual vertical axis over the conceptual horizontal axis when the conceptual axes were orthogonal to or opposite from the environmental axes, confirming the con-

TABLE 2
Mean Reaction Time and Percent Correct as a Function of Target Position and Cue Position (Defined Conceptually, Not Environmentally) for Each Orientation Group in Experiment 9

	Target position			
	Above	Below	Left	Right
	Top at right			
Top cue	1123	1097	1135	1126
	97	96	94	96
Bottom cue	1045	1049	1085	1081
	96	97	96	95
Left cue	1101	1108	1143	1149
	97	96	96	94
Right cue	985	1072	1088	1091
	97	97	94	95
	Top at left			
Top cue	1116	1141	1197	1147
	96	96	93	95
Bottom cue	1103	1074	1162	1154
	95	96	93	95
Left cue	1135	1133	1147	1191
	95	96	93	94
Right cue	1045	1076	1171	1098
	96	97	93	95
	Top at bottom			
Top cue	1126	1139	1299	1239
	94	95	95	93
Bottom cue	1209	1160	1240	1164
	96	95	92	93
Left cue	1154	1152	1262	1272
	94	95	94	92
Right cue	1172	1138	1155	1211
	97	95	91	95

ceptual frame hypothesis. It replicated Experiment 4 with more complex displays, in which relations were not uniquely associated with positions, and with a procedure that separated reading of the word from application of the relation the word specified to the cue and target. It replicated Experiment 7 with more complex displays, showing that the reference frame could be rotated around its origin. And it replicated Experiment 8, showing that the origin of the reference frame could be moved around space to correspond with the location of the cue.

EXPERIMENT 10: INTRINSIC CUING IN SIMPLE DISPLAYS

Experiment 10 examined cuing with intrinsic relations. As in Experiments 8 and 9, subjects saw an instruction display indicating the relevant spatial relation and then a target display with a cue and some potential targets. In this case, there were four potential targets rather than nine. More important, the cue was a drawing of a human head instead of an asterisk. Examples of the face cues appear in Fig. 11. Intrinsic relations require that the reference object has intrinsic axes, that is, a top and bottom, front and back, and left and right sides. The asterisk has no intrinsic axes and so cannot support intrinsic relations. Human heads

FIG. 11. Examples of face cues from Experiment 10.

have tops and bottoms, fronts and backs, and left and right sides, so they can. Experiment 10 asked whether subjects would be able to direct attention in relation to the intrinsic axes of head cues (cf. Farah, Brunn, Wong, Wallace, & Carpenter, 1990).

The head cue allowed a more complete test of the conceptual frame hypothesis than the previous experiments in this article. It compared the priority of all three reference axes. Subjects saw three different views of the head, each of which specified two reference axes. The *Front* view specified top-bottom and left-right. It supported *above*, *below*, *left of*, and *right of* relations. The *Profile* view specified top-bottom and front-back. It supported *above*, *below*, *in front of*, and *in back of* relations. The *Top* view specified front-back and left-right. It supported *in front of*, *in back of*, *left of*, and *right of* relations. The conceptual frame hypothesis predicts that top-bottom should be easier than front-back and front-back should be easier than left-right. The three different views tested each of the three possible pairwise comparisons. Previous experiments tested only one comparison: top-bottom versus left-right.

The orientation of the head cue was varied; it appeared in one of four different orientations, separated by 90 degrees. The conceptual frame hypothesis predicts that the intrinsic reference frame, not the environmental reference frame or the subject's own (egocentric) reference frame, determines which locations are easier to access. When pitted against environmental and egocentric reference frames, the intrinsic top-bottom should be easier than the intrinsic front-back and left-right. There should be no advantage for environmental or egocentric top-bottom over environmental or egocentric front-back or left-right.

The color of the distractor that was on the same axis as the target was varied to assess the flexibility of reference frame computation. Can subjects set parameters of the reference frame separately? Half of the time, the same-axis distractor was the same color as the target, and half of the time it was the opposite color. In the first case, subjects could respond after aligning the axis, before determining direction. In the second case, both orientation and direction must be set before responding. If subjects can set reference frame parameters separately, they should be faster when the same-axis distractor is the same color than when it is the opposite color. If they have to specify the reference frame completely on every trial, color of the same-axis distractor should have no effect. (The effects of variation in the same-axis distractor could not be assessed in the previous experiments. It was held constant in Experiments 1–3 and varied randomly without being recorded in Experiments 4–9).

The comparisons were all within-subjects. Each subject served for three sessions, seeing a different view of the head cue each session (e.g., front view on Session 1, top view on Session 2, and profile view on

Session 3). The four different instruction words in each session appeared in random order as did the four different orientations of the head cue.

Method

Subjects. Twelve subjects were recruited from the university population. Each subject served in three sessions and was paid $4 per session. All subjects reported normal or corrected vision and all subjects passed the Ishihara (1987) color-blindness test.

Apparatus and stimuli. The stimuli were displayed on IBM 8513 VGA monitors controlled by IBM PS/2 Model 50 computers. In the previous experiments, the stimuli were constructed using text displays. In this experiment, the stimuli were constructed using graphics displays. The fixation display was a white (IBM 15) dot in the center of the screen. The instruction display consisted of the word ABOVE, BELOW, FRONT, BACK, LEFT, or RIGHT presented in white at the center of the screen. The potential targets were red and green dots (IBM 12 and 10, respectively) that were 6.3 mm in diameter appearing 3.5 cm above and below and 3.3 cm to the right and the left of the center of the screen. Each display contained four dots, two red and two green. In half of the displays, the dots opposite each other were the same color. In the other half, opposite dots were different colors. Each color appeared in each position equally often.

The head cues were created by drawing a front view, a profile view, and a top view of a head. The skin was colored yellow, the hair brown, and the eyes blue. Detail lines (e.g., the outline of the nose in front view) were drawn in black. Examples of the three different views of the head are presented in Fig. 11. The front and profile views were 1.9 cm from top to bottom and 1.6 cm from side to side (front to back); the top view was 1.7 cm front to back and 1.7 cm side to side. Eight cues were constructed for each view of the head. Each view was rotated clockwise through four 90-degree steps (i.e., 0, 90, 180, and 270 degrees). Then it was reflected such that left and right were exchanged and rotated through the four 90-degree steps once again. With front and profile views, the 0 degree orientation was defined as the view with the head upright. With the top view, the 0 degree orientation was defined as the view with the nose pointed upward. In this view, the head's left and right corresponded to the subject's left and right.

For each view, there were four instruction words, four different target positions, two target colors, two colors for the same-axis distractor, and four different cue orientations. These factors were combined factorially to produce 256 different trial types. The 768 trials were divided into three sets, in which the 256 basic trial types occurred in random order.

Procedure. Each trial began with a fixation point exposed for 500 ms. Then the instruction display appeared and remained on for 500 ms. It was extinguished and the screen remained blank for 500 ms. Then the target display was exposed until the subject responded. After the response, the screen went blank for a 1500-ms intertrial interval. Subjects saw only one view of the head cue (i.e., front, profile, or top) in a session. The order in which the head cues appeared over sessions was counterbalanced across subjects by assigning two subjects to each of the six possible orders of views.

Subjects responded on the numeric keypad. Half of the subjects pressed "8" if the target was red and "2" if it was green. The other half pressed "4" if it was red and "6" if it was green. Assignment to mapping conditions was orthogonal to assignment to orders of views.

Subjects were told the sequence in which the displays would appear and that their task would be to report the color of one of the circles in the target display. They were shown pictures of a typical instruction display and a target display appropriate to their view condition (i.e., showing a front, profile, or top view of the head, as appropriate) and they were shown how to respond appropriately to the target display. They were told that the instruction word, the orientation of the head, and (consequently) the position of the target would vary randomly from trial to trial. Then the mapping rules were described and subjects were told to rest their index fingers lightly on the keys at all times. The trials began once the instructions were understood. The computer paused every 96 trials to allow subjects a break.

Results

The mean reaction times and accuracy scores in each combination of conditions appear in Table 3. Mean reaction times are plotted as a function of orientation of the head cue in Fig. 12. Overall, the results supported the conceptual frame hypothesis. *Above-below* was faster than *front-back*, and *front-back* was faster than *left-right*. Subjects appear able to direct attention from the intrinsic reference frame of the cue (cf. Farah et al., 1990). The differences were apparent in each of the four orientations in which the head appeared; subjects appear able to rotate intrinsic reference frames into alignment with the objects that possess them. Moreover, the differences were much larger when the distractor across from the target (on the same axis as the target) was opposite in color to the target than when it was the same color as the target. This suggests that subjects first identified the relevant axis and then computed direction: When the distractor on the same axis was the opposite color, subjects had to compute direction, but when it was the same color, they could respond without computing direction.

The data from each viewing condition were analyzed separately because the cues were different—there was no guarantee that the axes were as easy to discern in the different views—and because different relations were tested with different views. Mean reaction times in each viewing condition were analyzed in separate 4 (relation) × 4 (orientation) × 2 (same-axis distractor same or different) ANOVAs.

Front view. In the front view condition, *above* and *below* were contrasted with *left* and *right*. Reaction time, averaged over head orientation, was 828 ms for *above*, 816 ms for *below*, 1211 ms for *left*, and 1182 ms for *right*. The average difference between above-below and left-right was 370 ms. The difference apparent in each orientation (see Fig. 12). The difference was much larger when the distractor across from the target was the opposite color (623 ms) than when it was the same color (116 ms) as the target.

These conclusions were confirmed by ANOVA: The main effect of relation was highly significant, $F(3,33) = 69.64$, $p < .01$, but the main effect of orientation was not, $F(3,33) = 1.99$, $p > .10$, nor was the interaction between relation and orientation, $F(9,99) = 1.40$, $p > .10$. *Above* and *below* were compared with *left* and *right* with a planned comparison, which was highly significant, $F(1,33) = 207.95$, $p < .01$. The main effect of same-axis distractor was highly significant, $F(1,11) = 89.10$, $p < .01$,

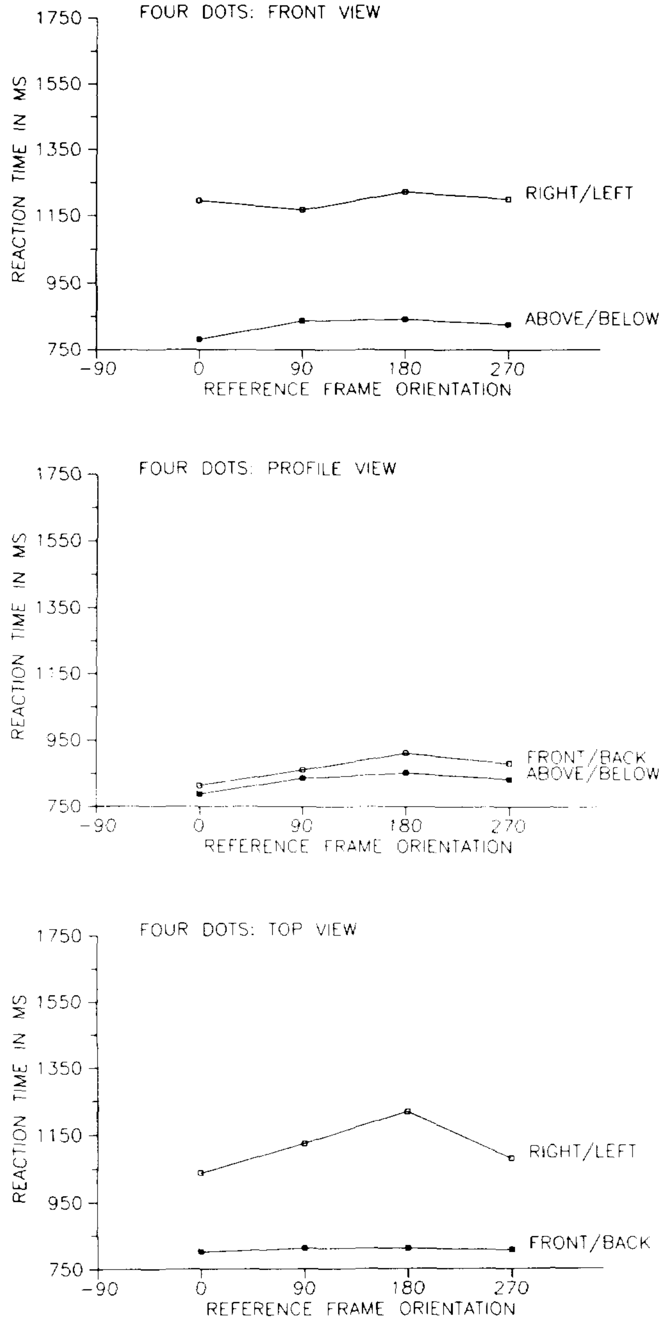

FIG. 12. Mean reaction time from Experiment 10 as a function of the orientation of the head cue for responses to Above and Below cues versus Left and Right cues in the Front View (top panel), to Above and Below cues versus Front and Back cues in the Profile View (middle panel), and to Front and Back cues versus Left and Right cues in the Top View (bottom panel).

TABLE 3

Mean Reaction Time and Percent Correct as a Function of View, Cue Type, Same-Axis Distractor, and Orientation in Experiment 10

Front view

	Same-axis distractor same				Same-axis distractor opposite			
	Above	Below	Left	Right	Above	Below	Left	Right
0 deg	768	803	913	868	778	777	1503	1499
	95	94	92	95	97	96	80	82
90 deg	824	804	985	933	859	859	1375	1378
	95	94	93	90	97	94	85	87
180 deg	832	787	928	940	898	849	1524	1494
	95	95	93	93	94	94	76	82
270 deg	827	807	953	942	832	843	1507	1400
	96	91	92	94	98	93	81	86

Profile view

	Same-axis distractor same				Same-axis distractor opposite			
	Above	Below	Front	Back	Above	Below	Front	Back
0 deg	787	768	784	821	804	796	816	839
	93	92	96	96	94	96	93	93
90 deg	845	836	819	894	844	822	842	893
	89	90	89	93	92	93	93	90
180 deg	838	839	879	898	870	865	915	957
	95	95	95	93	92	90	94	93
270 deg	846	806	856	877	853	830	888	905
	92	91	92	91	94	92	91	92

Top view

	Same-axis distractor same				Same-axis distractor opposite			
	Front	Back	Left	Right	Front	Back	Left	Right
0 deg	785	800	887	846	818	805	1229	1189
	92	95	94	91	95	94	90	91
90 deg	779	829	940	980	818	827	1298	1282
	96	96	93	91	93	94	87	85
180 deg	799	816	910	932	765	867	1541	1485
	96	94	92	94	94	94	86	87
270 deg	795	811	940	910	797	829	1281	1187
	96	95	92	93	95	94	86	94

as was its interaction with relation, $F(3,33) = 51.56, p < .01$. In addition, the interaction between same-axis distractor and orientation, $F(3,33) = 3.06, p < .05$, and the interaction between same-axis distractor, orientation, and target position, $F(9,99) = 2.83, p < .01$, but they did not compromise the effects of primary interest (see Table 3 and Fig. 12).

The accuracy data were generally consistent with the reaction times.

Profile view. In the profile view condition, *above* and *below* were compared with *front* and *back*. Reaction time, averaged over head orientation, was 836 ms for *above*, 820 ms for *below*, 851 ms for *front*, and 886 for *back*. The average difference between *above-below* and *front-back* was 40 ms. The difference apparent in each orientation (see Fig. 12). The nature of the same-axis distractor had little effect, probably because subjects did not have to compute *left* or *right*. The difference between *above-below* and *front-back* was 33 ms when the distractor was the same color as the target and 47 ms when it was the opposite color.

These conclusions were confirmed by ANOVA: The main effect of relation was significant, $F(3,33) = 6.08$, $p < .01$. A planned comparison showed that *above* and *below* were significantly faster than *front* and *back*, $F(1,33) = 12.80$, $p < .01$. The main effect of same-axis distractor was significant, $F(1,11) = 7.02$, $p < .05$, but its interaction with relation was not, $F(3,33) < 1$. The main effect of orientation was significant, $F(3,33) = 16.47$, $p < .01$.

The accuracy data were consistent with the reaction times.

Top view. In the top view condition, *front* and *back* were contrasted with *left* and *right*. Reaction time, averaged over head orientation, was 795 ms for *front*, 823 ms for *back*, 1127 ms for *left*, and 1101 ms for *right*. The average difference between *front-back* and *left-right* was 310 ms. The difference was apparent in each orientation. The difference was much larger when the same-axis distractor was opposite in color to the target (494 ms) than when it was the same (116 ms).

These conclusions were confirmed by ANOVA: The main effect of relation was significant, $F(3,33) = 34.11$, $p < .01$. A planned comparison showed that *front* and *back* were faster significantly faster than *left* and *right*, $F(1,33) = 101.34$, $p < .01$. The main effect of same-axis distractor was significant, $F(1,11) = 33.27$, $p < .01$, as was its interaction with relation, $F(3,33) = 23.96$, $p < .01$. The main effect of orientation was significant, $F(3,33) = 13.60$, $p < .01$, and it interacted significantly with relation, $F(9,99) = 6.32$, $p < .01$, and same-axis distractor, $F(3,33) = 16.14$, $p < .01$. In addition, the three-way interaction between relation, same-axis distractor, and orientation was significant, $F(9,99) = 6.11$, $p < .01$. These results reflect the fact that orientation effects were much larger with *left* and *right* than with *front* and *back*, particularly when the same-axis distractor was the opposite color. None of the interactions with orientation compromised the main results (see (Fig. 12 and Table 3).

The accuracy data were consistent with the reaction times.

Discussion

This experiment showed that subjects could direct attention according to the intrinsic axes of the head cue. Moreover, it showed that the reference frame could be rotated into alignment with the cue, in that the difference between *above-below*, *front-back*, and *left-right* were observed in each orientation of the head. This experiment provides strong support for the conceptual frame hypothesis.

The effects of the same-axis distractor were consistent with the hypothesis that subjects first align the reference frame with the cue and then compute direction. When the same-axis distractor was the same color as the target, subjects could respond as soon as they found the appropriate axis, before they computed direction. When the same-axis distractor was opposite in color, subjects had to find the axis and compute direction before responding.

EXPERIMENT 11: INTRINSIC CUING IN COMPLEX DISPLAYS

The final experiment examined intrinsic cuing in the complex, nine-element displays of Experiments 8 and 9. The head cue appeared in one of the four quadrants, and the subjects were to report the color of the dot that appeared in the instructed position, defined with respect to the intrinsic axes of the head. Front, top, and profile views were used to test all of the contrasts between relations entailed by the conceptual frame hypothesis, as in Experiment 10. This experiment asked whether reference frames could be translated about space as well as rotated around their origin. Experiment 10 addressed only rotation.

Method

Subjects. Three groups of 16 subjects were recruited from the university population. Each subject served in one session for course credit. All subjects reported normal or corrected vision and all subjects passed the Ishihara (1987) color-blindness test.

Apparatus and stimuli. The stimuli were displayed on IBM 8513 VGA monitors controlled by IBM PS/2 Model 50 computers. The stimuli were constructed using graphics displays. The fixation display was a white (IBM 15) dot in the center of the screen. The instruction display consisted of the word ABOVE, BELOW, FRONT, BACK, LEFT, or RIGHT presented in white at the center of the screen. The potential targets were nine red and green dots (IBM 12 and 10, respectively) 6.4 mm in diameter, which formed the vertices of four diamond shapes that made up a large diamond shape (see Fig. 7). The dots were separated from each other by 6.7 cm horizontally and 7 cm vertically. Each display contained four red dots and four green dots. The color of the ninth dot was determined randomly. It was red approximately half of the time. In all displays, the dots across the quadrant from the target (i.e., the same-axis distractor) was opposite in color to the target. Each color appeared in each position equally often.

The head cues were the ones used in Experiment 10 (see Fig. 11). They appeared in the center of the target quadrant in one of four different orientations (0, 90, 180, and 270 degrees from upright).

For each view, there were four instruction words, four different target positions, four different target quadrants, two target colors, and four different cue orientations. These factors were combined factorially to produce 512 different trial types, which were ordered randomly.

Procedure. The procedure was the same as in Experiment 10 except that the displays contained nine dots rather than four, that there were 512 trials rather than 768, and each subject served in only one session, seeing only one view of the head cue.

The mean reaction times from each viewing condition were analyses separately in 2 (group: mapping conditions) × 4 (orientation) × 4 (quadrant) × 4 (relation) ANOVAs.

Results

The mean reaction times and accuracy scores in each combination of conditions appear in Table 4. Mean reaction times are plotted as a function of the orientation of the head cue in Figure 13. The results replicate the previous experiments and generalize the conceptual frame hypothesis to intrinsic cuing with complex displays. *Above-below* was faster than *front-back*, and *front-back* was faster than *left-right*. The differences were apparent in each of the four orientations and in each of the four quadrants; subjects appear able to rotate and translate intrinsic reference frames (cf. Farah *et al.*, 1990).

Front view. The front view allowed *above* and *below* to be compared with *left* and *right*. Averaged over orientation and quadrant, reaction times 961 ms to *above*, 980 ms to *below*, 1380 ms to *left*, and 1340 ms to *right*. The average difference between *above-below* and *left-right* was 389 ms. The difference was apparent in each orientation in each quadrant. Reaction times did not vary much between quadrants, averaging 1185 ms for the top quadrant, 1168 ms for the right quadrant, 1156 for the bottom quadrant, and 1153 ms for the left quadrant. This suggests that the cue location was represented by a basic relation, confirming Experiments 8 and 9. Reaction time varied with the orientation of the cue, averaging 1137 ms for 0 degrees, 1150 ms for 90 degrees, 1258 ms for 180 degrees, and 1118 ms for 270 degrees.

These conclusions were supported by ANOVA: The main effect of relation was highly significant, $F(3,42) = 107.84$, $p < .01$. A contrast comparing *above-below* with *right-left* was highly significant as well, $F(1,42) = 322.24$, $p < .01$. The main effect of quadrant was not significant, $F(3,42) = 2.63$, $p < .07$, nor was a planned contrast comparing top and bottom with right and left, $F(1,42) = 1.22$. The main effect of orientation was significant, $F(3,42) = 21.40$, $p < .01$, and it interacted significantly with relation, $F(9,126) = 5.60$, $p < .01$. In addition, there were significant interactions between mapping condition and relation, $F(3,42) = 8.49$, $p < .01$, and between mapping condition, relation, and orientation, $F(9,126) = 9.59$, $p < .01$. None of these interactions compromised the main effect of interest, the contrast between *above-below* and *right-left*.

The accuracy data were consistent with the reaction times.

Profile view. The profile view allowed a comparison between *above-below* and *front-back*. Averaged over orientation and quadrant, reaction

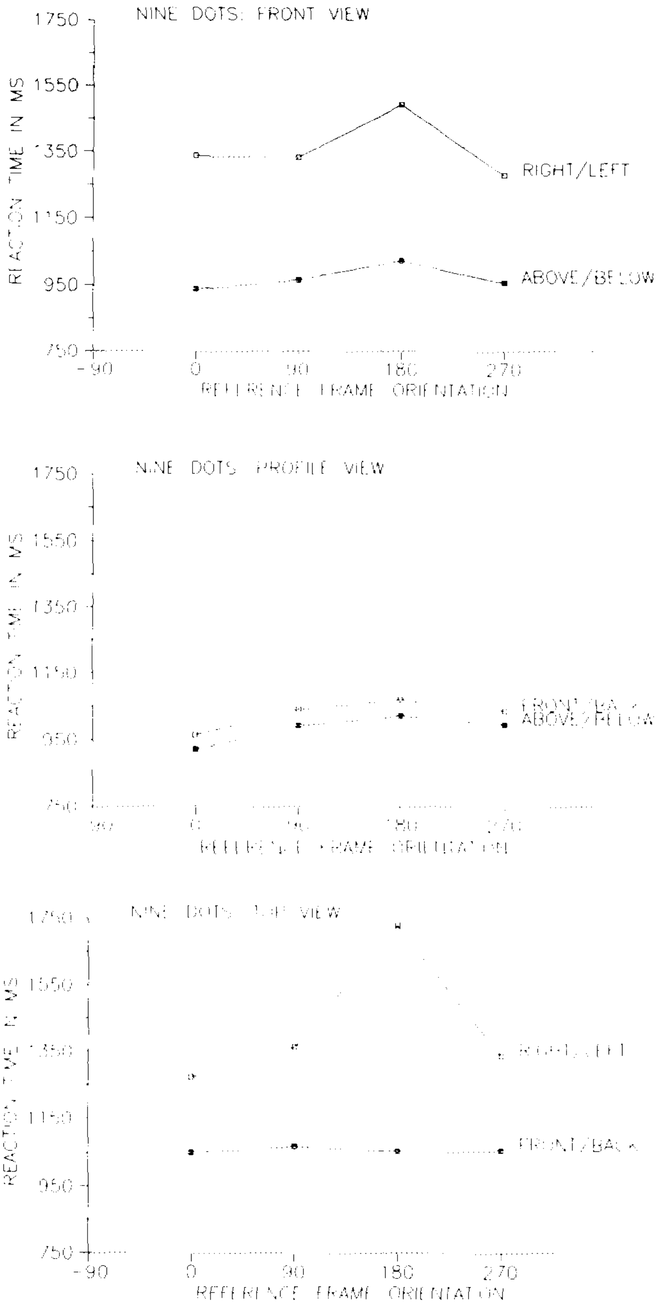

FIG. 13. Mean reaction time from Experiment 11 as a function of the orientation of the head cue for responses to Above and Below cues versus Left and Right cues in the Front View (top panel), to Above and Below cues versus Front and Back cues in the Profile View (middle panel), and to Front and Back cues versus Left and Right cues in the Top View (bottom panel).

time was 969 ms for *above*, 996 ms for *below*, 999 ms for *front*, and 1057 ms for *back*. The average difference between *above-below* and *front-back* was 45 ms, close to the value in Experiment 10. These differences appeared in each orientation in each quadrant. Reaction time did not vary much with quadrant, averaging 1020 ms in the top quadrant, 1002 ms in the right quadrant, 1012 ms in the bottom quadrant, and 985 ms in the left quadrant. This is consistent with previous experiments. Reaction time varied with the orientation of the cue, averaging 943 ms for 0 degrees; 1018 ms for 90 degrees, 1045 ms for 180 degrees, and 1016 ms for 270 degrees.

These conclusions were supported by ANOVA: The main effect of relation was significant, $F(3,42) = 8.87$, $p < .01$. A contrast comparing *above-below* with *front-back* was highly significant as well, $F(1,42) = 13.47$, $p < .01$. The main effect of quadrant was significant, $F(3,42) = 4.95$, $p < .01$. A planned contrast was significant, $F(1,42) = 11.22$, $p < .01$, revealing slower reaction times in top and bottom than in left and right quadrants. The main effect of orientation was significant, $F(3,42) = 27.20$, $p < .01$. The only significant interaction was the one between relation, quadrant, and orientation, $F(27,378) = 1.96$, $p < .01$. It did not compromise the main conclusions.

The accuracy data were consistent with the reaction times.

Top view. The top view allowed a comparison between *front-back* and *left-right*. Averaged over orientation and quadrant, reaction time was 1054 ms for *front*, 1062 ms for *back*, 1457 ms for *left*, and 1393 ms for *right*. The average difference between *front-back* and *left-right* was 367 ms. These differences appeared in each orientation in each quadrant. Reaction time did not vary much with quadrant, averaging 1244 ms in the top quadrant, 1242 ms in the right quadrant, 1258 ms in the bottom quadrant, and 1222 ms in the left quadrant, consistent with previous experiments. Reaction time varied with cue orientation, averaging 1162 ms for 0 degrees, 1216 ms for 90 degrees, 1391 ms for 180 degrees, and 1196 ms for 270 degrees.

These conclusions were supported by ANOVA: The main effect of relation was highly significant, $F(3,42) = 55.41$, $p < .01$. A contrast comparing *front-back* with *left-right* was highly significant as well, $F(1,42) = 163.44$, $p < .01$. The main effect of quadrant was not significant, $F(3,42) = 2.51$, $p < .08$. A planned contrast was significant, $F(1,42) = 4.11$, $p < .05$, revealing slower reaction times in the top and bottom than in the left and right quadrants. The main effect of orientation was significant, $F(3,42) = 46.88$, $p < .01$, suggestive of mental rotation. Orientation interacted significantly with relation, $F(9,126) = 25.75$, $p < .01$, indicating that the effects of orientation were largely confined to *left-right*.

The accuracy data were consistent with the reaction times.

TABLE 4

Mean Reaction Time and Percent Correct as a Function of View, Cue Type, Quadrant, and Orientation in Experiment 11

	Front view							
	Above	Below	Left	Right	Above	Below	Left	Right
	Top quadrant				Right quadrant			
0 deg	908	932	1357	1392	960	956	1401	1317
	98	95	91	93	94	97	94	91
90 deg	995	1013	1354	1350	953	939	1328	1314
	95	98	91	91	98	98	98	97
180 deg	996	1048	1592	1463	972	1034	1604	1456
	99	93	89	92	98	98	88	90
270 deg	962	987	1320	1289	904	1003	1263	1286
	94	95	94	93	94	97	93	96
	Bottom quadrant				Left quadrant			
0 deg	904	959	1326	1267	945	932	1302	1336
	97	95	89	88	99	97	93	91
90 deg	988	929	1326	1338	943	971	1306	1344
	98	96	93	98	98	93	96	93
180 deg	986	1011	1524	1425	1045	1095	1475	1394
	96	94	85	87	95	95	90	83
270 deg	987	926	1354	1246	933	950	1247	1219
	94	95	94	93	94	97	93	96

	Profile view							
	Above	Below	Front	Back	Above	Below	Front	Back
	Top quadrant				Right quadrant			
0 deg	896	909	949	1067	956	946	925	965
	97	94	95	95	98	95	97	94
90 deg	982	1058	944	1093	908	988	1022	1097
	96	97	97	92	98	95	95	93
180 deg	1040	1070	1067	1152	1005	1042	1029	1078
	98	96	93	94	90	92	95	93
270 deg	1033	1027	996	1033	961	969	1034	1114
	92	95	96	92	96	95	95	93
	Bottom quadrant				Left quadrant			
0 deg	864	931	988	973	927	939	932	916
	96	96	90	96	95	95	93	95
90 deg	1050	1039	1066	1043	953	968	1001	1070
	93	93	95	96	93	95	98	91
180 deg	1025	965	1044	1103	962	1059	1008	1068
	96	97	95	96	93	95	96	94
270 deg	998	1042	1003	1061	942	977	969	1071
	96	90	98	94	93	96	95	91

	Top view							
	Front	Back	Left	Right	Front	Back	Left	Right
	Top quadrant				Right quadrant			
0 deg	1087	1042	1295	1220	1093	1043	1286	1264
	92	95	93	89	95	100	90	91
90 deg	1042	1092	1394	1352	1105	1073	1386	1342
	95	95	89	83	91	98	93	93
180 deg	1006	1052	1757	1736	1068	1123	1755	1599
	93	97	83	83	91	95	80	77
270 deg	1074	1085	1348	1323	1008	1026	1369	1324
	94	95	92	90	93	96	91	91

TABLE 4—*Continued*

	Top view							
	Front	Back	Left	Right	Front	Back	Left	Right
	Bottom quadrant				Left quadrant			
0 deg	1070	1069	1318	1273	1003	1000	1266	1256
	95	94	87	88	97	96	97	88
90 deg	1008	1121	1396	1346	1076	1040	1424	1260
	93	95	89	91	95	94	93	90
180 deg	1043	1054	1782	1739	1032	1050	1742	1719
	96	95	78	80	95	94	85	77
270 deg	1099	1097	1419	1287	1042	1018	1371	1245
	94	97	88	94	97	95	88	91

Discussion

This experiment supported the conceptual frame hypothesis with complex displays and intrinsic cuing. It suggests that subjects can use intrinsic relations between cues and targets to direct attention, consistent with Experiment 10 (and contrary to Farah *et al.*, 1990). Experiment 11 also suggests that subjects can translate and rotate the reference frame into alignment with the intrinsic axes of the head cue.

GENERAL DISCUSSION

The theory of conceptual cuing sketched in the Introduction draws a distinction between directing attention to a single object and directing attention from one object to another. The former requires basic relations and the latter deictic or intrinsic relations. The theory assumes that deictic and intrinsic relations require subjects to impose or extract a reference frame before computing the relation, whereas basic relations do not. This prediction was tested and confirmed in the contrast between Experiment 6 and Experiments 4 and 5 in the contrast between cue position and target position effects in Experiments 8, 9, and 11.

Experiment 6 cued target position with a basic relation, naming the alternative positions with arbitrary digits. There was no evidence of a reference frame effect; top and bottom positions were no faster than left and right positions. Experiments 4 and 5 used the same displays but cued target position with the deictic relations *above*, *below*, *left*, and *right*. There were strong reference frame effects; *above* and *below* were substantially faster than *left* and *right*.

Experiments 8, 9, and 11 presented cues in four different positions and targets in four different positions relative to the cue. Cue position could be represented with a basic relation, but the position of the target with respect to the cue required a deictic (Experiments 8 and 9) or intrinsic (Experiment 11) relation. There was no evidence of reference frame computation in the cue position effects. The top and bottom positions were no faster than the left and right positions. There was strong evidence of reference frame computation in the target position effects. *Above* and *below* were much faster than *left* and *right*.

The theory of conceptual cuing assumes that reference frames are mechanisms of spatial attention, like spotlights and spatial indices. It assumes that reference frames are flexible, like other mechanisms of attention. They can be moved around space at will. This prediction was tested and confirmed in Experiments 7–11. Experiments 8, 9, and 11 showed that the origin of the reference frame can be translated across space, and Experiments 7, 9, 10, and 11 showed that the axes can be rotated into alignment with deictic and intrinsic cues. Experiment 10 suggested that subjects may be able to set parameters of the reference frame separately. Subjects were faster when the same-axis distractor was the same color as the target than when it was the opposite color, as if they responded after aligning the reference frame with the axis (setting its orientation), before determining which end was which (setting its direction).

The results support the theory and encourage further investigation. The three steps must be described in more detail and the details must be confirmed experimentally. The theory and results have many implications for future research. The remainder of the article sketches a few of them.

Reference Frames as Mechanisms of Attention

Are reference frames mechanisms of attention? The present experiments showed they have the kind of flexibility associated with attentional mechanisms like spotlights and spatial indices: they can be moved around space and oriented at will. The same sort of flexibility is apparent in studies of reference frame effects in object recognition (Attneave & Olson, 1967; Humphreys, 1983; Marr & Nishihara, 1978; Palmer, 1982, 1983; Rock, 1973), mental rotation (Cooper & Shepard, 1973; Corballis, 1988; Hinton & Parsons, 1981; Koriat & Norman, 1984, 1988; Robertson, Palmer, & Gomez, 1987), and symmetry perception (Corballis & Roldan, 1975; Pashler, 1990): Subjects can adjust the orientation and scale of reference frames at will. Reference frames orient attention to space, whereas spotlights and spatial indices orient attention to objects.

Are reference frames different from spatial indices? Certainly, they support more computations than spatial indices (e.g., direction, orientation, and distance). An interesting possibility is that reference frames are more elaborate versions of spatial indices (for a similar view, see Palmer, 1982, 1983). Reference frames are spatial indices with more parameters specified. Spatial indices are reference frames with only the origin specified. This view predicts that the origin of a reference frame could not be set without spatial indexing and spatial indexing could not occur without setting the origin of a reference frame. Alternatively, reference frames and spatial indices could be separate mechanisms. They may not be coupled closely, so reference frames could be specified without spatial indexing and vice versa. These are empirical issues to be investigated in future research. Experiment 10 provided some support for the elaboration view, suggesting that subjects could align reference frames without assigning direction.

Directing Attention Requires Computing Relations

The novel contribution of the theory is the second step—computing the relation between the cue and the target. The theory assumes that this is a necessary step in directing attention from one object to another. The theory focuses on how the step can occur, sketching the underlying computations. An important direction for future development is to understand the things that determine whether the step occurs: Cues can influence performance without directing attention to targets, and cues can be ignored. Tasks can be performed without computing relations between cues and targets. When will subjects compute relations between cues and targets?

One way to address this question is to delineate conditions under which subjects can perform tasks without computing the relation between the cue and the target. Delineating the conditions will not determine whether subjects compute the relation. Subjects may compute it even if they can do the task without it. However, the conditions will define an arena in which factors that determine subjects' choices can be studied, and that is the first step toward an answer. Subjects can avoid computing the relation if they find some way to do the task without directing attention from the cue to the target: They could attend only to the cue, they could attend only to the target, or they could attend to both cue and target but switch attention between them without directing it from one to the other.

Attending to the cue. If the cue and target are very close together, subjects may not need to switch attention to the cue. Attention to the cue may benefit processing in its neighborhood, so a target close enough (i.e., within the spotlight's beam) may benefit even if subjects do not direct attention to it (see e.g., LaBerge & Brown, 1989). These effects may be limited to targets that can be detected without directing attention to them (without spatial indexing) and to very small cue-target distances. Experiments 1–3 showed no effect of varying distance from 1 to 5 degrees in cuing with *next-to* relations. Subjects must have computed the relations at 5 degrees; the similarity of performance suggests they also did at 1 degree. Benefits from not having to compute the relation must be limited to distances smaller than 1 degree (also see Eriksen & Hoffman, 1972).

Attending to the target. If subjects can find the target without the cue, they may attend to it directly, without first attending to the cue. This could happen in target detection (Posner, 1980) and visual search experiments (Jonides, 1981), where the target is not defined in terms of the cue. Subjects may ignore cues in these experiments, so observed performance is a mixture of performance from trials in which the relation was computed and trials in which it was not. The mixture probability may depend on the difficulty of the cuing relation. Subjects may be more likely to ignore the cue with difficult relations like *left of* and *right of* than easy ones like *above* and *below*. Thus, difficult relations would show smaller benefits from cuing than easy relations.

Switching attention. Subjects may attend to the cue and the target and switch attention from the cue to the target without directing attention. Switching attention involves basic relations whereas directing attention involves deictic or intrinsic relations. In experiments like Müller and Rabbitt's (1989), a push cue directs attention to a target and the "movement" is affected by a pull cue. Subjects compute the (deictic) relation between the push cue and the target and use it to direct attention to the target. The pull cue perturbs the movement, attracting attention to itself momentarily before attention continues on to the position specified by the push cue. Attention may switch from the push cue to the pull cue and from the pull cue to the target, but it is not directed from the pull cue to the target.

Switching attention is clearest in search experiments, in which subjects must direct attention to items one by one but there is no requirement to direct attention from one specific item to another. Attention can be directed to any unexamined item. The next item can be chosen by selecting among locations identified by basic relations without computing a deictic or intrinsic relation between it and the current item.

What does this mean for the attention literature? It means that theorists must address how subjects compute relations between cues and targets. That computation must be explained whenever cue and target are separate objects. Depending on conditions, the computation is either necessary or optional, but theorists must have an explanation in either case. I believe the second step is an important consideration in most cuing experiments, even though the attention literature does not explain it. Attention theorists must find some way to explain how attention is directed from one object to another, or they must restrict their claims to directing attention to single objects (see e.g., van der Heijden, 1992).

Computing Relations Requires Directing Attention

The theory of conceptual cuing is built around a theory of the computation of spatial relations. The theory assumes that attention and intention are necessary to compute spatial relations. Attention is necessary because subjects must choose one out of indefinitely many relations to compute and two out of indefinitely many objects to use as arguments in the computation. These choices correspond to *analyzer selection* and *input selection* in classical analyses of attention (Posner & Boies, 1971; Treisman, 1969). Attention is necessary because each of the arguments must be spatially indexed and a reference frame must be applied to the reference object. Spatial indexing and aligning frames are important mechanisms of attention. Intention is necessary because these acts of attention are deliberate, voluntary choices.

An important corollary of the necessity of attention and intention is that spatial relations cannot be apprehended without them. Subjects should not compute spatial relations between objects they do not attend to, and if they attend to objects, they should not compute spatial relations between them unless they intend to. Greenspan and Segal (1984) tested and confirmed the necessity of intention (although that was not their own intention). They presented subjects with six digits in a column and a sentence describing a spatial relation between two digits (e.g., "6 above 2?"). Subjects' task was to decide whether any of the digits satisfied the relation. Greenspan and Segal presented displays in pairs, repeating the digits but changing the question. There was very little benefit from repeating the digits even though subjects knew they were repeated. Benefit accrued only when the arguments or the relation were repeated (e.g., "6 above 2?" followed by "5 above 2?"). New questions about new arguments produced no benefit (e.g., "6 above 2?" followed by "3 below 4?") even though they referred to displays subjects had seen before. Apparently, having seen the displays does not imply having computed all the spatial relations between all of the digits.

The theory of conceptual cuing appears to take a position in a longstanding debate in the attention literature over what can be done without attention (for reviews, see Hollender, 1986; Johnston & Dark, 1986). The literature contrasts an *early selection* view, which says that only the things that are attended are processed (e.g., Broadbent, 1958, 1982; Kahneman & Treisman, 1984), with a *late selection view,* which says that everything in the display is processed (e.g., Deutsch & Deutsch, 1963; Duncan, 1980).[3] The theory appears to endorse early selection.

---

[3] The contrast between early and late selection often concerns the locus of attentional selection in a hypothetical chain of processing that leads from stimulus to response. Early

Linguistic analyses allow us to dismiss a strong version of late selection immediately: The information that is available in a scene may be defined in terms of the number of ways the scene can be categorized (e.g., Garner, 1962). Linguistic analyses suggest that the number of possible categorizations is infinite. There appear to be no limits on the number of ways an object can be named (Brown, 1958) or the number of properties of an object that can be distinguished (Murphy & Medin, 1985). Actions as well can be described in an indefinitely large number of ways (Vallacher & Wegner, 1987). It would be physically impossible to compute all of these categorizations. Capacity limitations are not the issue. There are too many relations to be computed even if there were no limitations on attentional capacity; infinity is much larger than one (Broadbent, 1958) or 7 ± 2 (Miller, 1956). There may not be enough time and matter in the universe to support computation of all the relations, let alone the time and (gray) matter available in a single person's head.

It is important to be clear about the form of early selection that the theory endorses: The theory assumes there are two representations, a perceptual one that is built by obligatory, bottom-up processes and a conceptual one that is built by attention. The theory assumes that conceptual representations cannot be built without attention. Subjects cannot apprehend conceptual, categorical spatial relations without attending in the sense of choosing relations, spatially indexing arguments, and aligning reference frames. The information necessary to support the categorization may be available in the perceptual representation, but the categorization is not made explicit in the conceptual representation unless subjects pay attention to it (also see Bundesen, 1990). This does not mean that performance cannot be influenced by things outside the focus of attention. It does not mean that behavior cannot be sensitive to spatial relationships that are not attended. It means that subjects will not have an explicit conceptual representation of things and relations they do not attend to.

---

stages deal with "raw physical features," later stages deal with categorical properties, and even later stages deal with meaning. Research on the issue addresses the level of processing that can be reached without attention. I did not frame the contrast this way because I do not accept the idea of chain of processes that deal with progressively more abstract stimulus properties. For one thing, the evidence for such a chain is not compelling (Treisman, 1979). For another, the theory of conceptual cuing distinguishes between two representations—perceptual and conceptual—and assumes that selection occurs at the interface between them (also see Bundesen, 1990). Information corresponding to classical early and late stages can be found in both representations. The perceptual representation contains information that can support propositions about categorical properties and identities as well as raw sensory features. The conceptual representation can contain propositions about raw sensory features as well as categorical properties and identities (Miller & Johnson-Laird, 1976). Thus, the locus of selection is not a relevant issue.

Three Levels of Attention

The theory of conceptual cuing provides new insight into the early selection view that only attended objects are processed. Linguistic analyses suggest that objects are seen in relation to other objects. Both the reference object and the located object are attended, though attention may be "focused" on the located object. Thus, attending to a target *above* a cue involves attending to the cue as well as the target. Similar considerations apply to nonspatial relations (see e.g., Langacker, 1986). In terms of the early selection view, this means that subjects will process more than the object at the focus of attention. The background it is related to will be processed as well. However, that background will be processed with attention, not without it. Only that which is attended is processed. The new insight is that more than the focal object is attended. Thus, there are three levels of attention to consider: The attention paid to focal objects (i.e., located objects), the attention paid to background objects to which the focal object is related (i.e., reference objects), and the attention paid to unrelated objects (i.e., objects whose relation to the target is not computed).

Early selection says no attention is paid to unrelated objects. Most experiments in the literature distinguish only two levels of attention, that paid to the focal object and that paid to the rest of the display. They do not distinguish reference objects from unrelated objects. The failure to make that distinction may be responsible for some of the confusion in the literature, where evidence that nonfocal objects are processed is taken as evidence for late selection (Hollender, 1986; Johnston & Dark, 1986). Nonfocal object processing is consistent with late selection only if the objects are unrelated. Nonfocal objects processed as related objects are consistent with early selection.

This analysis suggests a different interpretation of Eriksen and Eriksen's (1974) flanker task. Subjects are shown three letters and asked to classify the middle one. Interference from the flanking letters is often interpreted as evidence for late selection (e.g., Eriksen & Schultz, 1979). The theory of conceptual cuing interprets it as an effect of attending to reference objects. *Middle* is a spatial relation just like *above, opposite,* and *next-to*. It specifies the position of one letter—the central one—with respect to the others. The central letter is the located object and the flanking letters are the reference objects. Just as subjects must attend to the reference object(s) in computing relations like *above, opposite,* and *next-to,* so must they attend to the reference objects in computing *middle*. Thus, the flanking letters are attended, not unattended. Attending to them as reference objects may activate the responses associated with them and produce the compatibility effect.

This analysis accounts for much of the data in the literature: The effect is ubiquitous, appearing over a wide range of conditions, when the target is defined as the middle character (see e.g., Miller, 1991). The effect is reduced as the distance between the flankers and the target increases (Eriksen & Eriksen, 1974; Kramer & Jacobson, 1991; Miller, 1991) because subjects may be more likely to relate the target to the fixation point than to the flankers as distance increases. The fixation point becomes the reference object and receives attention more often than the flankers. Similarly, the flanker effect is reduced when the target's position is indicated in advance with a bar marker (e.g., Eriksen & Collins, 1969). In this case, the bar marker becomes the reference object and is attended as such. The flankers are not likely to be reference objects and therefore are not attended (also see Yantis and Johnston, 1990).

Linguistic and Conceptual Control of Attention

Speakers can direct listeners' attention to single objects and from one object to another. The important question is how linguistic descriptions of objects and relations are translated into computations that are performed on perceptual representations: How can a sentence turn on a node? Even if we ignore the linguistic processes that derive conceptual representations from print or speech, the question remains: How are conceptual representations translated into computational procedures?

One possibility, explored in recent connectionist models, is that all of the possible representations already exist as nodes and connections in a network and computation involves simply spreading activation through the network. Phaf, van der Heijden, and Hudson (1990) postulated preexisting nodes for each property and each combination of properties a person can attend to. Cohen, Dunbar and McClelland (1990) postulated preexisting nodes for attentional tasks sets (e.g., report color; report identity). These proposals have serious difficulties because of the generativity of language: The Phaf *et al.* model suffers a combinatorial explosion because the number of properties people can attend to is indefinitely large and the number of combinations of properties is even larger (e.g., Murphy & Medin, 1985). The Cohen *et al.* model suffers a similar combinatorial explosion because the number of objects people can attend to—the number of task sets they can adopt—is indefinitely large (e.g., Brown, 1958; Vallacher & Wegner, 1987). There could not be enough nodes in the universe, let alone someone's head, to represent the endless possibilities.

Language gets around the combinatorial explosion by using *compositional* representations and building representations as they are needed instead of precomputing them. Linguistic representations are compositional because they are formed by combining parts that are meaningful in themselves into wholes that have a new meaning that depends on the arrangement of the parts (Fodor & Pylyshyn, 1988). Phonemes form syllables; syllables form words; words form phrases; phrases form sentences; sentences form discourse structures, and so on. Compositional representations generate a large number of wholes by combining a small number of parts.

Compositional representations are created when they are needed, and this imposes strong constraints on computation. Something must keep track of the parts of the representation and the relations between them. Something must put the parts and relations together. A theory of attention must say what it is that does these computations and how it does them. The theory of conceptual cuing takes a step in that direction. Spatial indices keep track of the arguments and spatial reference frames keep track of the relations between them. The first and second steps in the theory say how the arguments and relations are put together. Perhaps future research can generalize the theory to nonspatial compositional representations.

Is This Attention?

The theory proposed in this article differs substantially in topic and content from current theories in the attention literature. It is reasonable to ask whether a theory that different is a theory of attention. The answer depends on the alternatives; if it is not a theory of attention, what is it a theory of? I considered two alternatives: It is a theory of the control of attention, not a theory of attention itself, and it is a theory of comprehension (of spatial relations), not a theory of attention.

Attention or control of attention? Many theorists explain attention by distinguishing a number of mechanisms and saying how those mechanisms interact. There are two interpretations of attention within these theories. One is that attention is one of the basic mechanisms. Spotlight theorists often talk as if the spotlight is attention and the other mechanisms are not. van der Heijden (1992) explicitly identified attention with spatial indexing (selection by location), distinguishing it from mechanisms that identify the target ("expectancy") and prepare responses ("intention"). From this perspective, the theory of conceptual cuing is a theory of the control of attention rather than a theory of attention because it is a theory of the control of spatial indexing.

Another interpretation is that attention is the behavior that emerges from the interaction of the basic mechanisms. Attention is the system in action, not one of the parts of the system. Broadbent's (1958) filter theory explained attention in terms of the joint action of the filter and the limited capacity channel. Kaheneman's (1973) capacity theory explained attention in terms of the availability of capacity and the subject's policy for allocating capacity. Posner explained attention first in terms of interacting information processing components (e.g., Posner & Boies, 1971) and then in terms of interacting brain systems (e.g., Posner & Petersen, 1990).

I prefer the second interpretation. Attention is the application of basic mechanisms, like spatial indices and reference frames, to perceptual and conceptual representations. The idea that attention is one of the basic mechanisms involved in the application seems to confuse the thing to be explained with the explanation.[4] But however attention is interpreted, theorists must explain how the basic mechanisms are controlled. Theories that assume attention is one of the mechanisms have to explain how that mechanisms can be controlled.

Attention or comprehension? The theory of conceptual cuing is a theory of the comprehension of spatial relations. Is it also a theory of attention? That depends on the relation between attention and comprehension. The literature suggests that attention and comprehension are different processes. They are addressed in different experiments and accounted for by different theories. My theory suggests that attention and comprehension cannot be different processes because they involve the same basic mechanisms. Directing attention and comprehending spatial relations both involve spatial indexing and aligning reference frames. The representations and processes necessary to direct attention are the same ones necessary to comprehend spatial relations.

Attention and comprehension could be different levels of analysis of the same phenomenon. Comprehension could be a "higher level" description and attention a "lower level" description; people attend in order to comprehend. My theory suggests attention is not subordinate to comprehension. People attend to comprehend in some cases, as in computing spatial relations, but they comprehend in order to attend in other cases, as in linguistic and conceptual cuing.

The theory of conceptual cuing suggests that attention and comprehension are different perspectives on the same phenomenon, emphasizing different aspects of the same thing. The attentional perspective focuses on process more than content and addresses events that unfold moment by moment in real time. The comprehension perspective focuses on content more than process, and addresses the representations that result from processing rather than the processing itself. Comprehension is goal di-

---

[4] van der Heijden (1992) did not confuse the explanation with the thing to be explained; he wanted to explain something different. I assume attention is a natural phenomenon and the goal of theory (ultimately) is to explain the natural phenomenon. van der Heijden (1992) assumes (explicitly; see his Chapter 1) that attention is an internal mechanism and the goal of his theory is to explain experimental performance in terms of internal mechanisms. Researchers who adopt my goal (explaining attention) and van der Heijden's assumption (attention is an explanatory mechanism) confuse the explanation with the thing to be explained.

rected; comprehending implies succeeding in attaining a goal. Attention is what one does to direct oneself toward a goal. Attending implies trying, not succeeding. One can attend and fail as well as attend and succeed.

From this perspective, the theory of conceptual cuing is a theory of attention and it is a theory of comprehension. From this perspective, all theories of attention are theories of comprehension, and vice versa. The theorists themselves may not think that way, but this article has shown that it can be profitable to think that way nevertheless. The attempt to explain comprehension and attention at the same time emphasizes cognitive constraints on attention. After a decade of exploring neurological constraints and four decades of exploring psychophysical constraints, it is time to examine constraints imposed by language and cognition.

Conclusions

How do we direct attention from one object to another? The theory and data presented in this article suggest that directing attention involves computing a spatial relation between one object and another—between a cue and a target—and they demonstrate the importance of spatial reference frames in computing the relation. The experiments suggest that reference frames are important mechanisms of attentional selection. They can be rotated and translated across space according to the intentions of the observer and they can be aligned with the intrinsic axes of attended objects. The experiments also suggest that the semantics of the relations between cues and targets have powerful effects on performance. The semantics specify the computational goals that the attention system must satisfy. The goals shape the algorithms that the attention system implements. And the algorithms determine how performance unfolds in real time.

REFERENCES

Anstis, S. M. (1974). A chart demonstrating variations in acuity with retinal position. *Vision Research*, 14, 589–592.

Attneave, F., & Olson, R. K. (1967). Discriminability of stimuli varying in physical and retinal orientation. *Journal of Experimental Psychology*, 74, 149–157.

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.

Broadbent, D. E. (1985). *Perception and communication*. London: Pergammon.

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

Brown, R. (1958). How shall a thing be called? *Psychological Review*, 65, 14–21.

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.

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

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

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.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, **97**, 332–361.

Colegate, R. L., Hoffman, J. E., & Eriksen, C. W. (1973). Selective encoding from multielement visual displays. *Perception and Psychophysics*, **14**, 217–224.

Cooper, L. A., & Shepard, R. (1973). The time required to prepare for a rotated stimulus. *Memory and Cognition*, **1**, 246–250.

Corballis, M. C. (1988). Recognition of disoriented shapes. *Psychological Review*, **95**, 115–123.

Corballis, M. C., & Roldan, C. E. (1975). Detection of symmetry as a function of angular orientation. *Journal of Experimental Psychology: Human Perception and Performance*, **1**, 221–230.

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

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

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

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

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

Egly, R., & Homa, D. (1991). Reallocation of visual attention. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 142–159.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics*, **16**, 143–149.

Eriksen, C. W., & Collins, J. F. (1969). Temporal course of selective attention. *Journal of Experimental Psychology*, **80**, 254–261.

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

Eriksen, C. W., & Murphy, T. (1987). Movement of the attentional focus across the visual field: A critical look at the evidence. *Perception and Psychophysics*, **42**, 229–305.

Eriksen, C. W., & Schultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception and Psychophysics*, **25**, 249–263.

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.

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.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.

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

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

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

Hollender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, **9**, 1–66.

Humphreys, G. W. (1983). Reference frames and shape perception. *Cognitive Psychology*, **15**, 151–196.

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

Jackendoff, R. (1983). *Semantics and cognition*. Cambridge, MA: M.I.T. Press.

Jackendoff, R., & Landau, B. (1991). Spatial language and spatial cognition. In D. J. Napoli & J. A. Kegl (Eds.), *Bridges between psychology and linguistics: A Swarthmore festschrift for Lila Gleitman*. Hillsdale, NJ: Erlbaum.

Johnston, W. A., & Dark, V. (1986). Selective attention. *Annual Review of Psychology*, **37**, 43–75.

Jolicoeur, P., Ullman, S., & MacKay, L. (1986). Curve tracing: A possible basic operation in the perception of spatial relations. *Memory and Cognition*, **14**, 129–140.

Jolicoeur, P., Ullman, S., & MacKay, L. (1991). Visual curve tracing properties. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 997–1002.

Jonides, J. (1981). Voluntary vs. automatic control over the mind's eye movement. In J. Long & A. D. Baddeley (Eds.), *Attention and Performance IX*. Hillsdale, NJ: Erlbaum.

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

Juola, J. F., Bouwhuis, D. G., Cooper, E. E., & Warner, C. B. (1991). Control of attention around the fovea. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 125–141.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

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

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

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

Koriat, A., & Norman, J. (1984). What is rotated in mental rotation? *Journal of Experimental Psychology: Learning, Memory and Cognition*, **10**, 421–434.

Koriat, A., & Norman, J. (1988). Frames and images: Sequential effects in mental rotation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **14**, 93–111.

Kosslyn, S. (1980). *Image and mind*. Cambridge, MA: Harvard Univ. Press.

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

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

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

Langacker, (1986). An introduction to cognitive grammar. *Cognitive Science*, **10**, 1–40.

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

Logan, G. D., & Sadler, D. (1995). A computational analysis of the apprehension of spatial relations. In P. Bloom, M. A. Peterson, L. Nadel, & M. Garritt (Eds.) *Language and space*, Cambridge, MA: M.I.T. 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. (1956). The magical number seven plus or minus two: Some limits on our capacity to process information. *Psychological Review*, **63**, 81–97.

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

Miller, J. (1989). The control of attention by abrupt visual onsets and offsets. *Perception and Psychophysics*, **45**, 275–291.

Miller, J. (1991). The flanker compatibility effect as a function of visual angle, attentional focus, visual transients, and perceptual load: A search for boundary conditions. *Perception and Psychophysics*, **49**, 270–288.

Morrow, D. G., & Clark, H. H. (1988). Interpreting words in spatial descriptions. *Language and Cognitive Processes*, **3**, 275–291.

Müller, H. J., & Rabbitt, P. M. A. (1989). Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. *Journal of Experimental Psychology: Human Perception and Performance*, **15**, 315–330.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, **92**, 289–316.

Nissen, M. J. (1985). Accessing features and objects: Is location special? In M. I. Posner & O. S. M. Marin (Eds.), *Attention and Performance XI*. Hillsdale, NJ: Erlbaum.

Palmer, S. E. (1982). Symmetry, transformation, and the structure of perceptual systems. In J. Beck (Ed.), *Organization and representation in perception* (pp. 94–144). Hillsdale, NJ: Erlbaum.

Palmer, S. E. (1983). The psychology of perceptual organization: A transformational approach. In J. Beck, B. Hope & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 269–339. New York: Academic Press.

Pashler, H. (1990). Coordinate frame for symmetry detection and object recognition. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 150–163.

Pashler, H. (1991). Visual selective attention and the two-component theory of divided attention. *Journal of Experimental Psychology: Human Perception and Performance*, **17**, 1023–1040.

Phaf, R. H., van der Heijden, A. H. C., & Hudson, P. T. W. (1990). SLAM: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, **22**, 273–341.

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

Posner, M. I. (1988). Structures and functions of selective attention. In T. Boll & B. Bryant

(Eds.), *Master lectures in clinical neurology and brain function*. American Psychological Association.

Posner, M. I., & Boies, S. J. (1971). Components of attention. *Psychological Review*, **78**, 391–408.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. in H. Bouma & D. Bouwhuis (Eds.), *Attention and Performance X*. Hillsdale, NJ: Erlbaum.

Posner, M. I., & Petersen, S. (1990). The attention system of the human brain. *Annual Reviews of Neuroscience*, **13**, 25–42.

Posner, M. I., & Presti, D. E. (1987). Selective attention and cognitive control. *Trends in Neuroscience*, **10**, 12–17.

Pylyshyn, Z. (1984). *Computation and cognition*. Cambridge, MA: M.I.T. 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.

Remington, R., & Pierce, L. (1984). Moving attention: Evidence for time-invariant shifts of visual selective attention. *Perception and Psychophysics*, **35**, 393–399.

Robertson, L. C., Palmer, S. E., & Gomez, L. M. (1987). Reference frames in mental rotation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **13**, 368–379.

Rock, I. (1973). *Orientation and form*. New York: Academic Press.

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

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.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception and Psychophysics*, **50**. 184–193.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception and Psychophysics*, **51**, 599–606.

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

Treisman, A. (1979). The psychological reality of levels of processing. In L. S. Cermak & F. I. M. Craik (Eds.), *Levels of processing in human memory* (pp. 301–330). Hillsdale, NJ: Erlbaum.

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

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

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

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

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

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

Vallacher, R. R., & Wegner, D. M. (1987). What do people think they're doing? Action identification and human behavior. *Psychological Review*, **94**, 3–15.

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

van der Heijden, A. H. C. (1992). *Selective attention in vision.* London: Routledge.

Warner, C. B., Juola, J. F., & Koshino, H. (1990). Voluntary allocation versus automatic capture of visual attention. *Perception and Psychophysics*, **48**, 243–251.

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.

Yantis, S. (1988). On analog movements of visual attention. *Perception and Psychophysics*, **43**, 203–206.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective visual attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, **10**, 601–621.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 121–134.

Yantis, S., & Johnston, J. C. (1990). On the locus of visual selection: Evidence from focused attention tasks. *Journal of Experimental Psychology: Human Perception and Performance*, **16**, 135–149.

(Accepted June 28, 1993)