

# Psychological Review

VOLUME 92 NUMBER 2 APRIL 1985

## Cognitive Coordinate Systems: Accounts of Mental Rotation and Individual Differences in Spatial Ability

Marcel Adam Just and Patricia A. Carpenter  
Carnegie-Mellon University

Strategic differences in spatial tasks can be explained in terms of different cognitive coordinate systems that subjects adopt. The strategy of mental rotation that occurs in many recent experiments uses a coordinate system defined by the standard axes of our visual world (i. e., horizontal, vertical, and depth axes). Several other possible coordinate systems (and hence other strategies) for solving the problems that occur in psychometric tests of spatial ability are examined in this article. One alternative strategy uses a coordinate system defined by the demands of each test item, resulting in mental rotation around arbitrary, task-defined axes. Another strategy uses a coordinate system defined exclusively by the objects, producing representations that are invariant with the objects' orientation. A detailed theoretical account of the mental rotation of individuals of low and high spatial ability, solving problems taken from psychometric tests, is instantiated as two related computer simulation models whose performance corresponds to the response latencies, eye-fixation patterns, and retrospective strategy reports of the two ability groups.

The main purpose of this article is to provide a theory of how people solve problems on psychometric tests of spatial ability, focusing on the mental operations, representations, and strategies that are used for different types of problems. The theory is instantiated in terms of computer simulation models whose performance characteristics resemble human characteristics. A second purpose of the article is to analyze the processing differences between people of high and low spatial ability. One computer model simulates the processes

of the low-spatial subjects, and the other simulates the processes of the high-spatial subjects. The differences between the two models are small and localized, but they produce performance differences that are large and general. This approach to explaining processing commonalities and differences among individuals progresses beyond the classification of abilities, and specifies exactly what high- and low-spatial subjects do differently while solving problems (see also Carpenter & Just, in press; Carroll, 1976; Egan, 1978; Pellegrino & Kail, 1982; Snow, 1980; Snow & Lohman, 1984; Sternberg, 1981).

This research was supported in part by Grant MH-29617 from the National Institute of Mental Health and Contract N-00014-82-C-0027 from the Office of Naval Research. The order of authorship is arbitrary and was decided by the toss of a coin.

We thank Randy Mumaw and Bill Chase for their help in obtaining the subjects for Experiments 1 and 3, and for providing their percentile ranking in the psychometric battery.

Requests for reprints should be sent to either Marcel Adam Just or Patricia A. Carpenter, Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213. Copies of the CAPS system (for VAX/VMS or VAX/UNIX) on magnetic tape will be provided on arrangement.

### Cognitive Coordinate Systems

We begin our analysis by considering some of the properties of coordinate systems, formalisms that can be used to describe spatial objects and their transformations. Although coordinate systems are mathematical rather than psychological formalisms, they provide a possible starting point for characterizing human spatial representations. The most psychologically relevant attribute of a coordinate

system is its usefulness for describing quantitative relations among geometric objects. The value of this property becomes clear by considering the classical geometry developed by the ancient Greeks, which lacked a coordinate system. Classical Euclidean geometry provided an axiomatic system for describing properties of physical objects such as points, lines, angles, and polygons, and certain relations among the objects, such as equality, congruence, and parallelism. Because it lacked any inherent numerical system, Euclidean geometry could not deal with many kinds of metric relations and transformations, such as generalized rotation, translation, and size scaling of a geometric object. For example, it would be difficult within Euclidean geometry to express the fact that two polygons with the same structure differed by a translation of 1 inch, a rotation of  $45^\circ$ , and a scaling factor of 2. It was not until about 2,000 years after the Greeks that Descartes combined algebra with geometry, to create analytic geometry. This innovation provided a coordinate system that allowed physical objects to be not only represented, but also mathematically transformed. A Cartesian coordinate system, consisting of an origin and a set of mutually perpendicular axes, established a one-to-one mapping among three domains: real numbers, points in physical space, and points (ordered triples) in a mathematical coordinate system. These mappings allowed properties of one domain to be imported into another. In particular, the mapping between real numbers and points in the coordinate system allowed algebraic operations that correspond to spatial transformations to be applied to geometric objects.

Because a Cartesian coordinate system allows geometric objects to be represented and transformed (say, by rotation), mathematical terms can be used to precisely describe human spatial processes, including mental rotation. However, there are many ways to mathematically describe a given rotation, and it is not easy to tell which of the variations are psychologically interesting. Some mathematical descriptions may be notational variants of each other, whereas other variations may correspond to important psychological differences. One variation that appears to reflect important psychological differences is the

variation in possible coordinate systems within which an object can be embedded. Specifically, we can consider how people select the axes for a cognitive coordinate system, and how they mentally rotate within that system.

### *Selecting a Cognitive Coordinate System*

Physical objects are perceived with respect to a cognitive coordinate system, which consists of at least an implicit origin and some directional axes. The existence of an implicit coordinate system has been demonstrated by research on the recognition of objects that have previously been seen from a different perspective (e.g., Marr, 1982; Rock, 1973). Certain familiar shapes (such as the outlines of countries) are often unrecognized and misidentified if presented in an unusual orientation (Rock, 1973). Rock argued that part of the recognition process includes assigning an implicit *up* and *down* direction to the perceived object. In other words, the mental description of some objects contains an implicit reference to a coordinate system that is extrinsic to the object (such as the object being upright with respect to the environment). The consequence is that it is harder to recognize an object if its orientation does not match the previously stored one.

Adopting a new coordinate system, different from the system within which the object was originally encoded, can interfere with the ability to extract information from the representation. For example, the most common cognitive coordinate system for representing a cube contains axes orthogonal to the faces, and within this system it is very easy to mentally specify the location of the eight cube vertices in the representation. But if subjects are first asked to perform a task that induces a different coordinate system, then finding the vertices becomes very difficult (Hinton, 1979; see also Humphreys, 1983). The first task requires the subjects to mentally tilt a cube so that the diagonal that passes through center of the cube is vertical. That diagonal then becomes one of the axes of the induced cognitive coordinate system. Subsequently, the subjects make many errors in locating the vertices of the cube in their mental representation. Thus, even rudimentary infor-

mation that would be readily visible in a physical object is relatively inaccessible in a mental representation if the cognitive coordinate system is uncongenial to the retrieval of that type of information.

The existence of a cognitive coordinate system can also be demonstrated in mental rotation tasks. One series of studies attempted to discover the determinants of the vertical axis of the cognitive coordinate system in a mental rotation task, disassociating the retinal upright from the gravitational/room upright by having subjects tilt their heads in some conditions (Corballis, Zbrodoff, & Roldan, 1976). The reaction time is generally shorter if the major axis of one of the figures to be compared coincides with a major axis of the cognitive coordinate system, so one can empirically determine which axis is being used in the cognitive coordinate system. The results of one such study showed that the choice of axes was partially determined by the nature of the stimulus figure. For figures that had no intrinsic upright, like an array of random dots, the retinal upright was used as the vertical axis in the cognitive coordinate system. However, for familiar figures with a clear structural dimensionality of their own, namely alphabetic characters, the gravitational/room upright was used as the vertical axis. For familiar figures that are haptically presented to blindfolded subjects, the subjects' hand position (parallel to or at a 45° angle to the table edge) determined the vertical axis of the cognitive coordinate system (Carpenter & Eisenberg, 1978). It is interesting that blind subjects in the same task used a physical context (e.g., the tabletop) to define the vertical axis.

If an object has more than one main structural component (i.e., several major axes, like a giraffe's neck, trunk, and legs), then each component can be represented within its own local frame of reference. Such a representation produces a separate cognitive coordinate system for each part of a complex object, with labeled pointers from each part to every other contiguous part, indicating the point and angle of attachment (Marr & Nishihara, 1978). The advantage of this type of representation is that each part of a figure can be dealt with separately, and each separate part is eminently manipulable. The way this

type of representation allows a person to deal with a complex object is to divide and conquer.

These studies demonstrate that spatial information is coded with respect to a coordinate system and that there often exist alternative coordinate systems. They also demonstrate that the cognitive coordinate system has effects on recognition, information retrieval, and on spatial transformations, such as mental rotation. The article specifies in detail the coordinate system that is used in a mental rotation task. We suggest that alternative coordinate systems can explain some (although not all) individual differences in spatial ability, as well as strategic differences in spatial tasks.

*Human and mathematical coordinate systems.* There are some known ways in which mathematical coordinate systems and cognitive coordinate systems differ. Unlike the mathematical system, the human representation of an object also has a viewing point, a location from which the mind's eye views the object. The linguistic terms we use to name parts of objects often reflect the existence of the viewing point, such that we talk about the *front* or *back* of a child's toy block, even though those two surfaces may be identical in all other respects besides their relation to the viewing point. The viewing point may be different from the origin of the cognitive coordinate system or it may coincide with it, depending on the nature of the object and the task. The origin of the cognitive coordinate system is usually at the object's center of gravity. If the object is larger than a person, then the viewing point can coincide with that origin or it can be outside the represented boundaries of the object. For example, when viewers are asked to describe a room or apartment, some people mentally place themselves in the room, whereas others describe it as though from a distance (Levelt, 1982; Linde & Labov, 1975).

The existence of a viewing point suggests that certain portions of an object may be "hidden" when viewed from that point. The surfaces of real objects made of opaque material occlude other surfaces, so that an observer cannot see the back of a solid cube, for example. It seems that the representations of occluding surfaces are also occluding, al-

though the representations are only symbolic. This property of representational occlusion has implications for information retrieval from a cognitive coordinate system. When subjects are asked to imagine one object hidden behind another object, they are less likely to recall the hidden object than the visible object (Keenan & Moore, 1979). We show that the information on the hidden faces of a cube is also susceptible to loss.

When the viewing point is outside the object, it can be at varying distances from the object, but there seems to be a normative distance, one at which the object subtends about 50° of visual angle (Kosslyn, 1980). In other words, when the viewing point is outside the object, then the distance between the viewing point and the object is largely determined by the size of the object. The distance from the viewer influences the amount of detail that is easily accessible in the presentation, something analogous holding a photograph at a nearer or farther viewing distance, depending on whether one is interested in fine-grain detail or the broad strokes (Kosslyn, 1980).

There appears to be an upper bound on the amount of detail that can be represented within a cognitive coordinate system. We can imagine a tree and some leaves on the tree, but it is difficult to imagine the veins in the leaves at the same time as one imagines the entire tree. We typically deal with this problem by creating a "window" on the component we are interested in. The window is an embedded cognitive coordinate system usually centered on the component of interest, like an insert of a map that shows a smaller region in greater detail than the scale of the main map would allow. Unlike maps, our working memories appear too limited in capacity to keep both the main cognitive coordinate system and the embedded cognitive coordinate system in an activated state simultaneously. We can shift our attention from one embedded cognitive coordinate system to another (effectively, a translation) and the amount of time taken for the shift may vary with the distance (Kosslyn, 1980). Another manifestation of the capacity limitation is that the parts of the object at the center of a representation seem to contain more detail than do parts distal from the center, with a

decreasing gradient of resolution. By contrast, mathematical systems generally have sharply defined boundaries.

In sum, cognitive coordinate systems have several properties that distinguish them from mathematical systems. We specify in detail some of the alternative coordinate systems that can be used in rotation problems that appear in tests of spatial ability, and show how some of the psychological properties of these systems affect the qualitative and quantitative aspects of performance.

*Outline of this article.* The largest part of this article explains how people solve different problems from the Cube Comparisons test of spatial ability, contrasting the performance of people who are low or high in spatial ability (as measured by psychometric tests). The theoretical explanation takes the form of two related computer simulation models (one for the low-spatial and the other for the high-spatial subjects) expressed as production systems. Many of the observed individual differences can be ascribed to differences in the choice of cognitive coordinate systems. Two additional experiments briefly demonstrate that the theoretical explanation generalizes to a larger group of subjects taking a psychometric test and also generalizes to a second spatial test. The final discussion considers the interdependence between the choice of a cognitive coordinate system and the choice of a strategy for performing a spatial task. The discussion ends by suggesting that some of the difficulties encountered by psychometric classifications of spatial factors may have been due to the concomitant variation in cognitive coordinate systems and strategies.

#### *Structure of the Cube Comparisons Test*

We took psychometric tests as a starting point for an analysis of spatial ability because the problems are moderately interesting and because the tests have some predictive validity. Performance in paper and pencil tests of spatial ability is modestly correlated with performance in real world situations that require spatial ability (Ghiselli, 1966, 1973; Smith, 1964). The research focuses on two psychometric tests that appear to tap a component of spatial ability involving the manipulation of spatial representations. Items from

such tests typically consist of two drawings of an object that differ in orientation, and the subject's task is to decide whether the drawings could depict the same object. The scores across different instantiations of such tests are correlated, and the correlation is often attributed to a factor labeled *visualization* (Guilford, Fruchter, & Zimmerman, 1952; Lohman, 1979; McGee, 1979; Michael, Guilford, Fruchter, & Zimmerman, 1957; Smith, 1964).

The problems that our research has examined most closely were developed from the Cube Comparisons test (French, Ekstrom, & Price, 1963), an old psychometric tool, a version of which appeared in Thurstone's (1938) original Primary Mental Abilities battery. Figure 1b presents a typical problem—a pair of cubes that are described as drawings of children's blocks. The subject is told to assume that each block has a letter or number on each of its six faces, with the constraint that the same figure cannot appear more than once on a block. The task is to determine whether the two drawings could possibly depict the same block. One commonly reported method of solving the problem in Figure 1b is to mentally rotate the *A* on the front face of the right cube to make it upright, like its mate on the left cube. The *E* on the right cube would then be rotated to the top face, where it would match its left-hand mate in location and orientation. The *J* would be rotated out of view, where it would match a hidden face of the left cube, whereas the *P* on the left cube would match a hidden face of the right cube. These two drawings could depict the same block, so the correct response is *same*.

An analysis of the problem space revealed two main variables that could determine the difficulty of a *same* problem. The first variable is the length and complexity of the trajectory through which one cube has to be manipulated to bring it into alignment with the other cube. The second variable is the presence of letters whose orientation is ambiguous.

*Standard trajectories.* The five possible non-null trajectories for same trials can be described in terms of rotations around axes that are perpendicular to the faces of the cube. These are called *standard trajectories*. The trajectories that we present in this para-

a.		0 DEGREES 3 MATCHES	
b.		90 DEGREES 2 MATCHES	
c.		180 DEGREES (Same Axis) 1 MATCH	
d.		180 DEGREES 3 MATCHES	
e.		180 DEGREES (Different Axes) 1 MATCH	
f.		270 DEGREES 2 MATCHES	

Figure 1. An example of each type of *same* problem in the Cube Comparisons task.

graph are intended as descriptions of the stimulus, whereas the psychological processes are discussed below. The *same* problems require zero, one, two, or three 90° rotations to equalize the location and orientation of one pair of visible letters of the same identity, hereafter called *matching letters* or *matches*. There are either one, two, or three pairs of matching letters in each problem type. Thus, the six problem types shown in Figure 1 can be labeled as 0°–3 Matches (the identity condition); 90°–2 Matches; 180° (same)–1 Match, where there are two 90° rotations around the same axis; 180°–3 Matches; 180° (different)–1 Match, where the rotations are around two different axes; and 270°–2 Matches. (The order in which the three problem types involving 180 Degrees are presented in Figure 1 and in subsequent figures is motivated by expository rather than theoretical considerations.)

*Alternate trajectories.* Whereas the standard trajectories can be used in the solution process for all six problem types, alternative trajectories can be used to solve three problems—the 180°–3 Matches, 180° (different)–1 Match, and 270°–2 Matches conditions.

The alternative trajectories, illustrated in the right-most column of Figure 1d, 1e, and 1f, are around axes that are not perpendicular to the faces of the cubes and are shorter than the standard trajectories. The alternative trajectory for the  $180^\circ$ -3 Matches problem in Figure 1d is a  $120^\circ$  twist around an oblique axis that passes through the entirely visible corner and through the center of the cube. The alternative trajectory for the  $180^\circ$  (different)-1 Match problem in Figure 1e is a  $120^\circ$  twist around an oblique axis that passes through the top-left corner of the front face and through the center of the cube. The alternative trajectory for the  $270^\circ$ -2 Matches problem in Figure 1f is a  $180^\circ$  flip around an axis that passes through the middle of the right edge of the front face and through the center of the cube. The choice of trajectory has implications for the computations that must subsequently be performed. Moreover, we show that the low-spatial subjects never used these shorter trajectories, whereas high-spatial subjects usually did use them.

### *Alternative Strategies*

In spatial tasks that at least superficially seem to involve a spatial transformation, there are four main strategies that subjects reported using. We describe the strategies and the cognitive coordinate system upon which each is based:

1. *Mental rotation around standard axes.* This is the form of mental rotation that is most frequently discussed in the psychological literature (e.g., Cooper & Shepard, 1973; Shepard & Metzler, 1971). Often the object is mentally rotated in the plane of the picture so that the axis of rotation is the depth ( $z$ ) axis, or the object is mentally rotated in depth so that the rotation axis is the vertical ( $y$ ) axis. In all instances of this strategy, the axis of rotation is one of the usual three, the  $x$ ,  $y$ , or  $z$  axis, as defined by the visual environment, gravity, or the retina, although these frames of reference usually coincide. These frames of reference are external to the object that is being mentally rotated.

2. *Mental rotation around task-defined axes.* Some subjects can mentally rotate around any arbitrary axis that is useful or necessary for a particular task. The alternative

trajectories in Figure 1d, 1e, and 1f illustrate three arbitrary, task-defined axes. The process by which subjects compute the axis of rotation becomes interesting and important when the axis is determined by the properties of each individual problem. By contrast, the axis-finding process is trivial if the same rotation axis is used repeatedly from trial to trial. The ability to find and mentally rotate around a task-defined axis implies that at least in a limited way, the axis of rotation is being used as an axis of a cognitive coordinate system.

3. *Comparison of orientation-free descriptions.* A representation generated within an object-defined cognitive coordinate system is invariant with the object's orientation in space. Two such representations of the cubes in the Cube Comparisons task can be directly compared without regard to the orientation of the two depictions. A subject using this strategy codes the relation between one pair of letters on the left cube (e.g., the top of the *A* points to the bottom of the *E*) and then codes the corresponding relation on the right cube to determine if the two codes are consistent with each other. The two codes are consistent if they are identical or if one member of the letter pair on the left cube corresponds to a hidden letter on the right cube. The use of an orientation-free code requires that each major part of the object (each face of a cube, in this instance) be coded within its local coordinate system, such that each part has a top and bottom direction to represent the local orientation of the components. In addition, the relative orientations of adjacent parts (or their respective coordinate systems) are also represented.

4. *Perspective change.* The problems in the Cube Comparisons test, in the Vandenberg (1971) Mental Rotation test, and other similar tasks can be solved by mental perspective change. In this strategy, the object's position and the observer's position are coded within a cognitive coordinate system that includes both the observer and the object, with the object's represented position used as the origin. The use of this strategy entails mentally changing the representation of the observer's position relative to the object and hence his or her view of the object, but keeping the representation of the object's orientation in

space constant. In the Cube Comparisons task, one can imagine how the right-hand cube in Figure 1f would look when viewed from directly below. That view is consistent with the view depicted on the left, and so the correct response is *same*. The axis-finding process becomes a decision of which view to take of the object.

### *Representations Used in Mental Rotation*

The standard rotation strategy has revealed a close correspondence between physical objects and processes on one hand, and mental representations and processes on the other hand. The main empirical observations in mental rotation research are that the response time increases monotonically with the angle of rotation (Cooper & Shepard, 1973; Shepard & Metzler, 1971) and that an object that is being mentally rotated from one orientation to another mentally passes through intermediate orientations (Cooper & Shepard, 1973).

One unresolved issue in the standard strategy is the content of the rotated representation, particularly in the case of a fairly complex stimulus object like the figures used by Shepard and Metzler (1971). On one hand, it is possible that the representation that is being rotated is the representation of the entire object, including all the represented information about the object's shape and possible ornamentation of surfaces. On the other hand, the representation that is mentally rotated could be a subset of the representation of the entire object, such as a skeletal outline of the object, or even just a part of the object. We have previously proposed that in the Shepard-Metzler task, subjects rotate a skeletal representation, consisting of vectors that correspond to the major axes of each segment of the figure. Representing a Shepard-Metzler figure with this type of skeletal representation is similar to representing the shape of an animal (like a giraffe, ostrich, or rabbit) with a figure made of pipe cleaners (Just & Carpenter, 1976). The pipe cleaners (or vectors) capture the essence of certain shapes without representing the surface of the object (cf. Marr & Nishihara, 1978). One advantage of such a representation is that it is easy to manipulate mathematically, and perhaps mentally as well.

The question of what is rotated has been studied by investigating the effects of object complexity on task performance. If only a skeletal representation of a figure were being rotated, and if that skeletal representation were rotated one piece at a time, then the rotation transformation itself should be unaffected by the complexity of the figure from which it was extracted. One study compared the rotation of Shepard-Metzler figures and simple two-dimension rectilinear nonsense figures (Carpenter & Just, 1978). Even though the total reaction time to do a large rotation of a complex figure was approximately twice as long as for the rotation of a simple figure, the actual time spent in applying the rotation transformation (estimated from eye-fixation behavior) was only marginally longer for the complex figure. Most of the extra time on the complex figure was spent in the encoding stage prior to rotation, presumably extracting the skeletal features to be included in the representation, and in the confirmation stage, relating the rotation of those features to the remaining parts of the figure. According to this interpretation, the complexity of a figure affects the difficulty of extracting the representation to be rotated, but not the rotation. In addition, it suggests that a representation of only one part of a complex object may be mentally rotated at a time.

Further support for this position comes from a series of studies that showed that the increased complexity (additional structural features) of an object did not affect response time in a rotation task if the complexity was irrelevant to the discrimination, but did affect response time if the complexity was critical to the discrimination. This result suggests that in the former case, not all of the properties of the object were contained in the representation that was being rotated (Yuille & Steiger, 1982). These results also question the suggestion that a complex object can be rotated as a whole (Cooper & Podgorny, 1976). Our results and theory speak to this issue, indicating that mental rotation of a complex figure is performed by rotating different parts of the figure in separate rotation episodes.

Of course, it is difficult to specify the content of a representation without saying something about its format, and very much

has already been said about the possible formats of spatial representations, whether analogue or propositional (Anderson, 1978; Hayes-Roth, 1979; Hinton, 1979; Kosslyn, 1981; Pylyshyn, 1973, 1979). The format of representation that we have used in our previous and current models is a propositional representation in which the values of some attributes can be specified numerically. Thus structural relations can be represented in terms of conventional propositional relations, and metric information can be represented with the numerical values of attributes like length. Other formats could accommodate the same content, but the format we have used is particularly congenial to the processes we propose and it is compatible with representations we have proposed for nonspatial tasks (Thibadeau, Just, & Carpenter, 1982).

#### *Processes Used in Mental Rotation*

Closely related to the issues of representational content and format is the nature of the processes that operate on the representation. The suggestion from our previous work is that the rotation process is discrete, with fairly large step sizes in the tasks we examined. In addition, the rotation process is not ballistic; that is, it is not unchangeable once set in motion toward some target orientation. Rather, it is monitored after every rotation step to determine if the new orientation is sufficiently close to the target orientation (Carpenter & Just, 1978; Just & Carpenter, 1976). The experiments we report support this general characterization of rotation.

The model we have proposed (Just & Carpenter, 1976) has three major processes. Stated in terms of the stimulus properties involved in the Cube Comparisons test, these are

1. *Search*—finding a pair of matching letters on the two cubes.

2. *Transformation and comparison*—mentally rotating a letter through a trajectory that will eventually bring its location and/or orientation into congruence with its mate's. The orientation is transformed by some increment, and after each step the two locations/orientations are compared to determine whether they are sufficiently similar. If they are not, another transform-compare iteration is executed.

3. *Confirmation*—determining that each of the remaining letters, after being subjected to the same transformations, match the location and orientation of their counterparts.

The second and third processes should differentiate the six problem types in the Cube Comparisons test. Because rotation time increases with rotation angle, we can predict the relative difficulty of the problems for subjects who use standard trajectories. The time needed to transform the initial pair of matching letters should increase from 0° to 90° to 180° to 270°. In addition, confirmation time should increase with the longer trajectories because the same transformations are applied to the other letters.

#### Cube Comparisons: A Model of Human Performance

The purpose of this experiment was to analyze how people perform the Cube Comparisons task and to determine which processes distinguish subjects of high-spatial ability from subjects of low-spatial ability. Subjects who had been psychometrically classified as being high or low in spatial ability solved Cube Comparisons problems while their eye fixations were recorded to trace the sequence and duration of the component processes.

#### *Experiment 1*

Experiment 1 included six exemplars of each of the six Cube Comparisons problem types, with each axis and direction of rotation represented equally often within each problem type. The 36 *different* trials were formed by first constructing the *same* pair and then altering the right cube by either changing the location or orientation of a matching letter, or exchanging the locations of two letters.

The subject initiated a trial by pressing a button while looking at a fixation point located where the center of the front face of the left cube would appear. The subject indicated a judgment of *same* or *different* by pressing one of two response buttons, which terminated the display. Immediately afterward, the stimulus cubes were displayed a second time and the experimenter recorded the subject's verbal account of the solution process. Each subject went through 6 practice

trials followed by the 72 test trials in random order. The graphics and eye-fixation instrumentation and some of the data acquisition procedures are described in more detail in Appendix A.

The subjects were 4 students who had scored well on a battery of nine psychometric spatial tests (mean percentile of 80 in a population of 144 university students), and 4 who had scored poorly (mean percentile of 21). (In Experiment 2 we examine a larger group of subjects performing a similar task.) The low-spatial subjects were academically successful but were low in spatial ability. Two were undergraduates in the humanities, 1 was a graduate student, and 1 was in a professional school. The high-spatial subjects were undergraduates in science and engineering. The psychometric test battery consisted of 9 tests, including several rotation tests, a number comparison test, an identical pictures test, a surface development test, and a paper form board test.

*Strategy reports.* Three of the 4 high-spatial subjects and all 4 low-spatial subjects described a rotation strategy on all nonidentity trials. The 4th high-spatial subject reported a strategy of comparing orientation-free descriptions. His pattern of response times differed from the others, and his data were analyzed separately and are reported separately from all of the other subjects.

High-spatial subjects usually reported using a nonstandard trajectory for those problems in which it was applicable, namely the 180°–3 Matches, 180° (different)–1 Match, and 270°–2 Matches conditions. The retrospective reports usually described the trajectories in sufficient detail for us to categorize them (59% of the reports could be categorized for the high-spatial subjects, 49% for the low-spatial subjects). On those trials in which the trajectory could be categorized, the 3 high-spatial subjects reported a nonstandard trajectory 81% of the time, compared to just one single report of a nonstandard trajectory among the low-spatial subjects,  $F(1, 5) = 20.08$ ,  $p < .01$ . A similar effect was found when protocols were scored for the corresponding *different* trials,  $F(1, 5) = 11.57$ ,  $p < .02$ . The statistical analyses above were performed on arcsin-transformed proportions of reported trajectories that were nonstandard.<sup>1</sup>

The reports were classified as indicating a standard trajectory if the subject clearly described two or three distinct movements. For example, a typical description for a 180°–3 Matches condition was “if you first rotate the *B* on the top to the front and then turn the cube so that the *B* will match (in orientation).” A description was classified as indicating a nonstandard trajectory if subjects made it clear that they had executed the trajectory in a single movement or had described a nonstandard axis of rotation. For example, a typical protocol of a high-spatial subject for a 180°–3 Matches trial was “I spun it around the corner of the three sides until the letters lined up.” To summarize, the low-spatial subjects characteristically described using standard axes, whereas the high-spatial subjects most often described trajectories that are the shortest for solving that particular problem.

*Response times.* The problems with more complex trajectories generally took more time to be solved than did simpler problems,  $F(5, 75) = 11.38$ ,  $p < .01$ , and the low-spatial subjects took much more time to respond than did the high-spatial subjects,  $F(1, 15) = 27.65$ ,  $p < .01$ . As shown in Figure 2, the low-spatial subjects took particularly long on problems with longer trajectories, resulting in an interaction of problem type and subjects,  $F(5, 75) = 3.38$ ,  $p < .01$ . There was almost no difference between the two groups in the identity condition, which involved no rotation, whereas the low-spatial subjects took more than twice as long as the high-spatial subjects on the most difficult trial type (13,864 ms vs. 6,349 msec in the 270°–2 Matches condition).<sup>2</sup>

<sup>1</sup> To assess the reliability of the classification procedure, an independent judge classified the trajectories on the basis of retrospective reports from 72 trials that allowed for alternative trajectories, selected from 4 randomly chosen high- and low-spatial subjects. There was complete agreement between the two judges as to whether the trajectory was classifiable and whether it was standard or nonstandard in 94% of the cases.

<sup>2</sup> Two statistical analyses were performed on the response times and gaze durations using data from only those trials that had correct responses and scorable eye-fixation protocols. In the first analysis, the problem of missing data was dealt with by including three observations per cell, out of a possible total of six. If there were more than three usable observations, then three were randomly

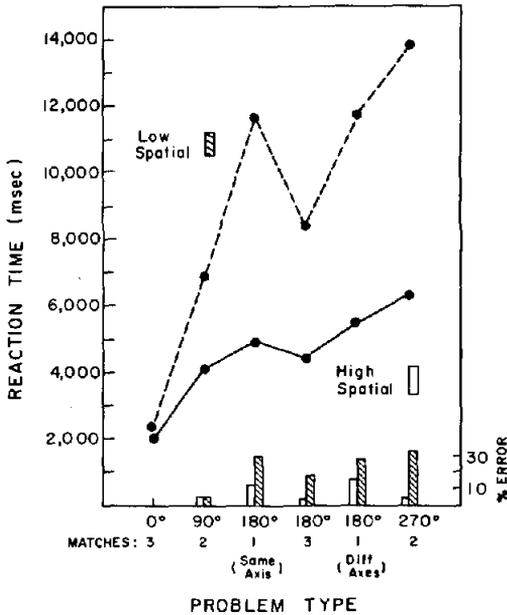


Figure 2. Reaction times and error rates for *same* problems for low-spatial and high-spatial subjects in the Cube Comparisons task. (The rotation angle referred to in the problem labels applies only to standard axis transformations.)

**Eye fixations.** The eye-fixation behavior was analyzed to determine the order in which the component processes were executed, to measure the time spent executing each process, and to identify the sources of individual differences between the high- and low-spatial subjects. The eye-fixation protocols were decomposed into segments corresponding to the three stages that were described earlier: searching for matching letters, determining and executing the trajectory that rotates one letter of a matching pair into congruence, and confirming the correspondence of the locations and orientations of other letters.

The segment that was easiest to identify was *initial rotation*, consisting of the first

sequence of eye fixations that alternated between matching letters on the two cubes. Fixations that occurred before initial rotation were classified as *search*. In those problems that contained more than one pair of matching letters, subjects generally processed the pairs one after another, fixating between the members of one pair at a time. The alternation between the members of the first pair was classified as initial rotation, whereas the alternation between members of the remaining pairs was classified as part of *confirmation*. The subsequent time spent looking between letters that had no mates was also classified as confirmation. In some cases, subjects made a second sequence of alternating fixations between the initially rotated letters; this was called *subsequent rotation*.<sup>3</sup>

The monitored rotation episodes consist of rotations plus comparisons to determine whether the orientations of the two representations are sufficiently similar. Therefore, the measured durations of such episodes are greater than zero even in the identity condition in which there is no rotation to be done because of the time taken by the comparisons. Our methodology cannot separate rotation time from comparison time, but we refer to the episodes as *rotation*, instead of the more accurate but cumbersome *rotation and comparison*.

Figures 3A and 3B and adjoining Tables 1 and 2 present two protocols that illustrate

<sup>3</sup> The eye-fixation data were too noisy to score on 10% of the trials for both the high- and low-spatial subjects. All of the remaining same trials on which a correct response was given were analyzed. The instrumentation for displaying the stimuli and monitoring eye fixations is described in Appendix A. Fixations that were not classifiable included occasional fixations to other letters during a rotation episode. If a single such fixation occurred during a rotation episode, it was ignored; if there was more than one fixation, then the rotation was considered to be terminated and the two fixations were attributed to confirmation. Subjects sometimes claimed to have rotated two letters at once and, indeed, their protocol would indicate fixations between the two pairs of matching letters. This was infrequent, and we found that the total gaze duration for rotating two letters was double the time for rotating a single letter. Hence, in the few times that this occurred, the rotation time was divided in half, with one half attributed to initial rotation and the other half to confirmation.

selected. This produced an analysis of variance (ANOVA) with subjects of two ability levels, six trial types, and three observations. In the second analysis, an ANOVA was done on the means of the six or fewer usable observations per cell. The results were generally similar and we report only the first analysis, which has somewhat more power.

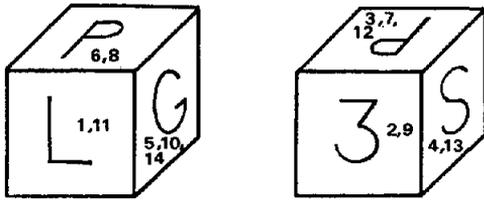


Figure 3A. Sequence of gazes on a 180 Degrees (same)-1 Match problem. (The location of the numbers from 1 to 14 indicates which face was fixated, but not the precise location of the fixations on the face.)

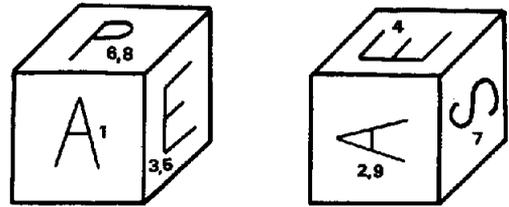


Figure 3B. Sequence of gazes in a 90 Degrees-2 Matches trial.

Table 1  
Location and Duration of Gazes on a 180° (Same)-1 Match Problem

Gaze no.	Cube	Letter	Duration (ms)	Process
1	Left	L	466	Search (1,384 ms)
2	Right	3	201	
3	Right	P	284	
4	Right	S	200	
5	Left	G	233	
6	Left	P	300	Rotation and comparison (1,133 ms)
7	Right	P	450	
8	Left	P	383	
9	Right	3	450	Confirmation (2,632 ms)
10	Left	G	349	
11	Left	L	216	
12	Right	P	617	
13	Right	S	267	
14	Left	G	733	

Table 2  
Location and Duration of Gazes on a 90°-2 Matches Problem

Gaze No.	Cube	Letter	Duration (ms)	Process
1	Left	A	369	Rotation and comparison (719 ms)
2	Right	A	350	
3	Left	E	351	Confirmation (868 ms)
4	Right	E	184	
5	Left	E	333	
6	Left	P	133	Confirmation (1,183 ms)
7	Right	S	400	
8	Left	P	233	
9	Right	A	417	

the nature and classification of the eye fixations. Consecutive eye fixations on the same face of a cube were aggregated into units called gazes. The protocol shown in Figure 3A shows the sequence of gazes produced by a subject solving a 180° (same)-1 Match problem. The numbers on the cube indicate the sequence of gazes on the faces of the two cubes, but not the exact location of the fixations; the braces show how the gazes were assigned to processes. The first set of gazes (Gazes 1-5) is attributed to the search for matching letters; this search is much longer than average, partly because there is no mate for the L that appears on the front face of the left cube. The rotation and comparison process is identified by the three consecutive gazes between the Ps (Gazes 6, 7, and 8).

This is followed by a confirmation process (Gazes 9 to 14) in which the subject computed the final location of the S and the 3 on the right cube.

The protocol shown in Figure 3B shows the sequence of gazes of another subject solving a 90°-2 Matches problem. In this protocol there is no initial search because the letter on the front face of the left cube (the A) matches the letter on the front face of the right cube. The first two consecutive gazes on matching letters (Gazes 1 and 2) are attributed to the rotation and comparison process. The rest of the gazes (3-9) are attributed to confirmation. The first part of the confirmation (Gazes 3, 4, and 5) operated on the Es; the duration of this process is similar to the time for the initial rotation and comparison of the As, a point we will return to in the next section. Then the subject computed the final location of the S on the right cube.

*How much of the cube is rotated at one time?* Can the representation of the entire cube be transformed with a single mental process? Does the rotation of one cube face (and the letter it contains) automatically bring along some or all of the other parts of the cube for a free ride? For example, if we attached a magnet to one face of a metallic cube and rotated that face by physically moving the magnet, then the rest of the cube, including the other five faces, would automatically have their positions altered by the same transformation. Does mental rotation of one face of a cube and the letter it contains similarly move the other faces into the transformed position? Or must each face be transformed separately? We can look to the eye-fixation data for an empirical answer to the question, by determining how much time subjects spend on the rotation of the first pair of matching letters they consider, as compared to the rotation of other matching letters. The eye-fixation protocols generally contain an episode of looking back and forth between a first pair of matching letters that are being rotated into congruence, and then, in the problems containing more than one pair of matching letters, another episode involving a second pair of matching letters. If the rotation of the first letter pair entailed the rotation of the second, then the second episode should be considerably shorter in duration because the rotation would not have to be executed a second time.

The data indicate that the duration of the second episode is generally similar to the first, implying that each face of the cube must be rotated separately. To examine this point quantitatively, we compared the durations of the two episodes of looking between matching letters in the 90°-2 Matches condition, where there is only one possible trajectory but two matching pairs of letters. For the low-spatial subjects, the initial rotation episode averaged 1,625 ms and the second episode, 1,712 ms. In other words, there were no savings. For the high-spatial subjects, the initial episode averaged 1,035 ms and the second episode consumed very little less, 974 ms,  $t(18) < 1$ . Thus, the time required to determine the relation between the second pair of matching letters is very similar to the

time for the first, suggesting that both episodes involve a similar mental rotation, and that only one face of the cube (including the letter it contains) is rotated at a time. The current fine-grain analysis of the performance indicates that even the representation of an object as structurally simple as a cube with letters on it is not rotated holistically.

Similar analyses of the time to rotate a second pair of letters were also obtained in a pilot study using 6 unselected subjects. These subjects had response times intermediate between those of the high- and low-ability groups. Their response times were 2,367; 5,775; 8,166; 6,812; 8,282; and 12,548 ms for the six conditions shown in Figure 2, respectively. For these subjects also, the mean duration of the first rotation episode on a 90°-2 Matches trial, 1,217 ms, was close to the mean duration of the second rotation episode, 1,335 ms. For 4 of the 6 subjects, the mean duration of the second episode was longer (differences of 17, 89, 312, and 315 ms), and for the other 2 subjects, the mean duration of the first episode was slightly longer (differences of 10 and 19 ms). Thus, the data from these subjects also supported the conclusion that subjects rotate the cube one face at a time and that there are no savings accrued toward the rotation of subsequent faces. In general, the eye-fixation data are consistent with a model of piecemeal rotation rather than with a holistic model (e.g., Funt, 1983).

#### *Processing Model for Low-Spatial Subjects*

Based on the subjects' strategy reports, response times, and eye fixations, we developed simulation models of the processes that high-spatial and low-spatial subjects use to perform the Cube Comparisons task. We first describe the model for the low-spatial subjects because it provides a baseline description of the mental rotation process for the six different problem types. The models are expressed as production systems operating within CAPS (Collaborative activation-based production system), a theory of the human information-processing architecture that is described in more detail elsewhere (Thibadeau, Just, & Carpenter, 1982). Production systems consist

of rules, or productions, that describe actions to be taken when certain information resides in working memory (Newell, 1973). One feature that distinguishes CAPS from most other production systems is that all of the productions whose conditions are satisfied at a given time, fire concurrently. The scanning of working memory and the subsequent firing of the enabled productions constitute one CAPS cycle. This feature provides CAPS models with an inherent metric of processing time because the CAPS cycles can then be related to human processing time.<sup>4</sup>

*Representation.* The visible letters on each face of each cube are represented by a propositional structure that has some correspondence to the subjects' verbal descriptions. The representation contains the name of the letter; the cube it is on (the left or right one); the location of the letter on the cube, indicating which one of the six faces it occupies (front, top, right, back, bottom, or left); and the orientation of the letter on the face. For example, the *B* on the left cube of Figure 1d is represented as

(LETTER: B, CUBE: left, FACE: "front",  
ORIENTATION: "upright").

The two arguments in quotation marks, which specify the face the letter is on and the orientation it has, are in numerical rather than symbolic form. Even though a subject may describe a surface as the *front* one, that description contains a considerable amount of information concerning its location in space relative to other surfaces. A similar point can be made for the orientation information; presumably people represent orientation information in terms of reference lines (such as vertical or oblique) and rough indications of deviations from a reference line. The formation of the representation is not simulated, but is automatically available whenever the stimulus is scanned. In a later section we discuss the cognitive coordinate system that is implicit in this description.

*Search.* The subjects usually reported that they first examined the left cube, encoded the letter on its front face, and then searched the right cube to find a matching letter. The model treats the left cube as a standard, just

as the human subjects do, because its letters are in an upright orientation. Consequently, the model's first step is also to find a letter on the right cube that has the same name as the letter on the front face of the left cube. If there is no mate, then there is an attempt to match the letter on the top or right of the left cube. Once a pair of matching letters is found, each member of the pair is tagged as a *focused* letter.

*Rotation.* After matching letters have been found, there are two successive rotation processes, one that equalizes their location and one that equalizes their orientation. The low-spatial subjects report performing these two operations separately, and the model also performs them this way. If the two matching letters are in corresponding locations to start with, as are the *As* in Figure 1b, then nothing more need be done to equalize locations. When the production system solves this problem, the location-equalizing phase is skipped simply because the relevant production, one of whose conditions is that the location of the focused letters be different, would not be enabled. If the two matching letters are on different faces, then the low-spatial model equalizes their locations by rotating the matching letter on the right cube to the same location as its mate. For example, in the 180°-3 Matches problem shown in Figure 1d, the model of the low-spatial subjects would rotate the *B* from the top of the right cube to the front face, to match the location of the *B* on the left cube. This transformation requires a 90° rotation around the *x* axis (i.e., the axis perpendicular to the right side of the cube and passing through its center).

<sup>4</sup> Two other distinguishing features of CAPS are that all propositions have an activation level that is manipulated by the productions and that all cycles take the same amount of time, regardless of which or how many productions fire in parallel on a given cycle. Other CAPS models that share these assumptions simulate the word-by-word time course of human reading and the solving of Raven (1962) Progressive Matrices items. The CAPS assumptions enable these other models to fit the human performance characteristics in several interesting ways. Although some of the assumptions are not essential to the models for the Cube Comparisons task, the assumptions are retained to maintain theoretical consistency across these very different task domains.

Once the *B* is mentally represented on the front face, the model of the low-spatial subject rotates the *B* counterclockwise around the *z* axis to equalize the orientations of the two *B*s.

*Computing the axis and direction of rotation.* Mental rotation entails a rotary motion around a given axis in a given direction, but the process by which a subject determines the axis and direction of rotation in a given problem has not been investigated to date. The process is particularly important in this task because the problems involve a variety of axes and directions, and a new determination must be made on each trial. The retrospective reports of the subjects did not describe how the determination of axis and direction was made. The process seems too rapid and automatic to introspect about in this task. The following description of how the simulation model computed the axis and direction of rotation seems a plausible first approximation of how human subjects might compute the information.

The axis of rotation can be computed if the starting and ending locations of a few points of the object to be rotated are known. The points on the surface of a rotated object move through a circular trajectory, from some original location to a new location. That circle of locations defines a plane that is perpendicular to the axis of rotation. The simulation model determines this plane from some locations along the circle, and then computes the normal of this plane that passes through the center of the circle, and that normal is the axis of rotation. All of the information necessary to compute the axis of rotation this way is available to the subjects before they do any mental rotation. If the locations of two matching letters are different, the starting location of the letter to be rotated is given by the location of the letter on the right cube and the destination location is given by the location of its mate on the left cube. For example, to equalize the locations of the *B*s in Figure 1d, the *B* on the right cube must be rotated from the top face to the front face. More specifically, the center of the *B* will move from the center of the top face to the center of the front face. For the model of the low-spatial subjects, the rotation

axis will always turn out to be perpendicular to one of the cube faces, and passing through the center of the cube, simulating the low-spatial subjects' choice of standard axes. The same general algorithm can be used to compute nonstandard axes. A more detailed description of the algorithm and an example appear in Appendix B.

The direction of rotation that the subjects choose in this task is the one that minimizes the rotation angle, choosing the direction that results in a rotation of no more than 180°. The simulation also chooses the direction of rotation by determining which of the two possible rotation angles is smaller. It might be noted though that if subjects are set to rotate in a particular direction on trial after trial, they sometimes continue to choose that direction even though it produces a mental rotation greater than 180° (Metzler & Shepard, 1974).

*Initial rotation.* Once the axis and direction of rotation have been determined, a production rotates the relevant letter of the right cube, by a given number of degrees, called the *step size*. We will discuss how we selected the step size in more detail, but for now we assume a 15° step size. The rotation production has as one of its conditions that the focused letter on the left cube is located on a different face than is its mate on the right cube. Thus, this production continues to fire iteratively over successive cycles, rotating the letter by 15° on each cycle, until the letter on the right has the same location as its mate on the left. The representation of one of the two focused letters (i.e., the one on the right cube) is changed by the rotation production on each iteration, by changing the numerical value of the attribute FACE. For example, during this phase of the processing of the problem shown in Figure 4, the representation of the *B* on the right cube will change from (1) to (2):

1. (LETTER: B, CUBE: right, FACE: "top", ORIENTATION: "90 Degrees")
2. (LETTER: B, CUBE: right, FACE: "front", ORIENTATION: "90 Degrees")

where "top" and "front" have numerical values that contain the coordinates of the

center of the face. A record is made of each rotation step as it is executed, including the axis and direction of rotation. This record, called the *transformation list*, is later used to apply the same transformations to other letters.

Table 3 describes some of the main productions in conversational form. Each one is labeled by its major function, such as *Get-front-face-match* for the first production. Next, the table lists the information that must be available (either present in working memory or available by scanning the stimulus) before the production is enabled. This information constitutes the conditions that must be fulfilled to make the production "fire." For example, *Get-front-face-match* is enabled if there is a letter on the right cube that matches the letter on the front face of the left. Finally, the table lists the actions that are taken when the production fires. The action of this particular production is to mark the pair of matching letters as the current focus of attention. Table 4 (p. 154) provides a trace of the model's processes as it solves a typical problem, the 180°-3 Matches problem shown in Figure 4. The trace is abbreviated, so much of the control information is not shown.

The letter or pair of matching letters that are being transformed are marked as the currently focused ones, with no more than two letters focused at any one time. Prior to being focused, the representations of the letters are assumed to not be in working memory, and after the representations have been transformed, the focus marker is removed and the representations are marked as *done*. When the two focused letters match completely in both location and orientation, the status marker on their representations is changed from focused to done, and a counter of the number of matching faces is incremented from zero to one. The *done* marker makes these letters ineligible to fulfill the conditions of the productions that try to find matches and perform transformations, so these letters are not subjected to any further processing. The search process and the two phases of the initial rotation process are executed identically on same and different trials.

*Confirmation.* After the first pair of matching letters has been brought into congruence, there is an attempt to bring other pairs of matching letters into congruence, using the same transformations. There may be zero, one, or two pairs of secondary matches remaining after the first pair of matching letters has been brought into congruence. If there are two secondary matches, only one pair is processed at a time, and the choice of which one to process first is made randomly. The processing is straightforward, involving the stepwise application of the stored rotations on the transformation list to the selected letters on the right cube. For example, in Figure 4, the same sequence of rotations that brought the *Bs* into congruence must be applied to the right-hand *A*. After all the transformations have been applied, the location and orientation of the letter from the right cube are compared to the those of its left-hand mate. If all is not identical, then the two cubes are different and the process terminates. If all three visible letters of the right cube can be rotated into congruence, then the two cubes are the same and the process terminates.

If all the pairs of matching letters can be rotated into congruence with the same transformations, but there are fewer than three such pairs, then there is a check of the remaining letters that have no mates, hereafter called the *singleton letters*. Specifically, the singleton letters on the right cube are subjected to the same transformations as the matching letters, and after transformation they must be located on one of the hidden faces if the two cubes are the same. If not, the cubes are different. For example, in Figure 1b, the *J* from the right cube must be rotated onto the left, hidden face. If there is more than one singleton (and there can be 0, 1, or 2), then they are processed one at a time, and the choice of which one to process first is made at random.

The confirmation process was the major source of error for both the low- and high-spatial subjects. For the low-spatial subjects, the error rates were 21.5% and 16.7% for the same and different trials, respectively, with most of the errors (71% of them) clustered in the 180°-3 Matches and 270°-2 Matches

Table 3  
*Some of the Main Productions in Conversational Form*

Condition	Action
<b>Get-front-face-match</b>	
If there is a letter somewhere on the right cube that has the same name as the letter on the front face of the left cube,	Mark that letter (on both cubes) as the currently focused letter.
(There are two similar productions that will detect matches on other faces if there is no match to the letter on the front face of the left cube).	
<b>Get-axis + direction-for-location-match</b>	
If the 2 focused letters mismatch in location (i.e., which face they are on),	Compute the axis and direction of rotation and indicate there is some rotating to be done.
<b>Rotate-until-locations-match</b>	
If there is rotating to be done and the focused letters do not match in location,	Rotate (amount = step size) the focused letter on the right cube using the axis and direction specified. Enter a record of this rotation on the transformation list.
<b>End-location-matching</b>	
If there is an indication to rotate but the two focused letters have the same location,	Remove the indication to rotate.
<b>Get-axis + direction-for-orientation-match</b>	
If the two focused letters have the same location but different orientations on that face,	Compute the axis and direction of rotation and indicate a need to turn <sup>a</sup> that letter.
<b>Rotate-until-orientations-match</b>	
If there is turning to be done and the focused letters do not match in orientation,	Rotate (amount = step size) the focused letter on the right cube using the axis and direction specified. Enter a record of this rotation on the transformation list.
<b>End-main-match</b>	
If the two focused letters match completely in both location and orientation,	Remove the focus from the pair of letters. Mark the pair as <i>done</i> . Store the fact that one pair of letters has been completely matched. Remove any indications of a need to turn or rotate anything. Indicate that the main match has been done.
<b>Get-secondary-matches</b>	
If there is a pair of matching letters on the two cubes (other than the letters marked done),	Mark them as the currently focused letters.
<b>Reduce-multiple-foci-to-one</b>	
If there are two pairs of focused letters,	Arbitrarily choose one of those pairs and delete the focus mark from it.
<b>Move-secondary-letter</b>	
If there is a pair of focused letters and a transformation list,	Apply the next unmarked transformation and mark it as done.

Table 3 (continued)

Condition	Action
Finish-secondary-letter	
If a pair of focused secondary letters match in both location and orientation and there are no more unmarked rotations on the transformation list,	Mark the pair as done. Increment the count of letters that have been completely matched.
(There are productions that attempt to bring singleton letters into congruence with hidden faces on the left cube. They are similar to the productions that manipulate secondary matching letters.)	
Detect-SAME	
If the count of completely matched letters is 3,	Respond "SAME" and stop.

*Note.* For the high-spatial model, the second to the fifth productions above are condensed into two productions: One detects location or orientation disparity or both, and computes the appropriate axis and direction of rotation; the second production rotates until all disparity is eliminated.

\* The word *turn* is used in this table to indicate a rotation of a letter's orientation without changing its location.

problems. For the high-spatial subjects, the error rates were lower, 7.4% and 9.3%, for the same and different trials, respectively.

The low-spatial subjects' retrospective reports indicated that they lost track of a letter during the confirmation process by neglecting to trace its complete trajectory. Errors occurred fairly often if the trajectory required that a letter was to be temporarily moved to a hidden face on its way to a visible face. This type of trajectory occurred in the 180°-3 Matches and the 270°-2 Matches conditions, precisely where the low-spatial subjects made most of their errors. These subjects' retrospective reports on the error trials consistently indicated that, after having mentally rotated a letter to a hidden face, they neglected to make it re-emerge to a visible face, as the appropriate transformation would require. Less often, they did make the letter emerge after having moved it to a hidden face, but with an incorrect orientation. These failures caused errors on both same and different trials.

The interesting aspect of the distinction between visible and hidden faces is that they are hidden or visible only in the cognitive coordinate system; the letters are all equally visible in the physical display, and visibility has no definition in the mathematical coordinate system. It is only in a cognitive coordinate system that it makes sense to speak of a face being visible if there is no opaque

surface represented between it and the viewing point. The errors caused by letters traveling to hidden faces can be simulated in the model by making the activation level of each letter depend partially on its visibility from the viewing point. If the activation level of a letter rotated to a hidden face falls below the threshold required by the rotation productions, the rotation of that letter will stop, even if not all the transformations on the transformation list have been applied. Stopping before all the transformations are applied will produce an error on some proportion of the trials that involve a confirmation trajectory passing through a hidden face.

A second source of error was a failure to confirm all of the letters. This kind of book-keeping error would occur if the subject did not keep track of which or how many letters had been checked.

*Different trials.* The production system model stops as soon as it detects an inconsistency between the two cubes, and states that the cubes are different. However, in this experiment and in others (Just & Carpenter, 1976) we have found that subjects differ from trial to trial and from each other in how they react to detecting an inconsistency in a same-different task involving complex figures. Sometimes subjects terminate immediately on detecting the inconsistency, as the model does. Sometimes they reprocess the entire trial the same way, or they try a different

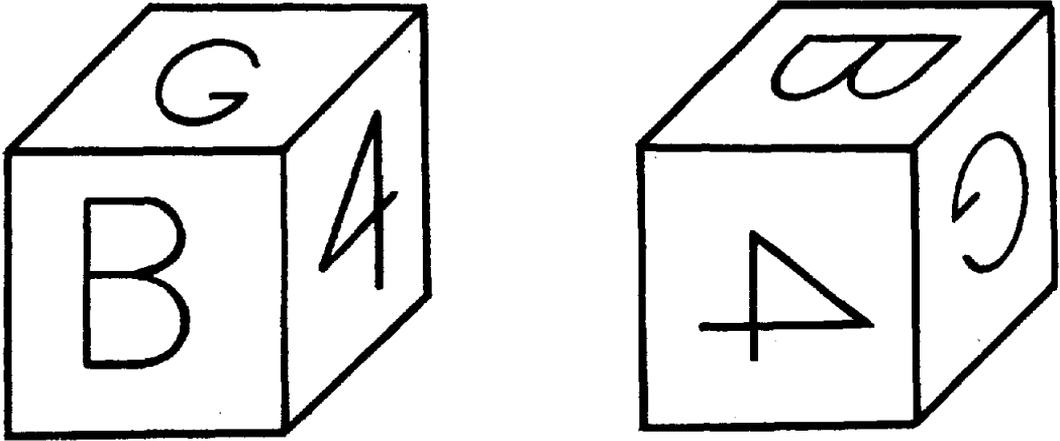


Figure 4. Example of a 180°-3 Matches problem.

Table 4

*Abbreviated Trace of the Low Spatial Model Solving the 180 Degrees-3 Matches Problem*

Cycle	Process
1	Initialization occurs, and the stimulus is made available for encoding.
2	The two matching <i>B</i> s are noticed and marked as focused.
3	The disparity in location of the two <i>B</i> s is noticed, and the computation is made to rotate counterclockwise around the <i>x</i> axis.
4-9	The <i>B</i> on the right cube is rotated as specified, in 15° steps, until its location matches the location of the <i>B</i> on the left cube. A record of each rotation is made on the transformation list.
10	The match in location is noticed (so the indication to rotate is removed). The mismatch in orientation is noticed, and the computation is made to turn counterclockwise around the <i>z</i> axis.
11-16	The <i>B</i> on the right cube is rotated as specified, in 15° steps, until it matches the orientation of the <i>B</i> on the left cube. A record of each rotation is made on the transformation list.
17	The complete match between the two <i>B</i> s is noticed, the <i>B</i> s are marked as done, and the number of completely matching letters is set to 1, and all indications to rotate or turn are removed.
18	The match between the two <i>A</i> s is noticed, and they are marked as focused. The same thing happens to the two <i>G</i> s.
19	The focus is removed from the two <i>G</i> s and a working copy of the transformation list is made.
20-31	All the rotations that had previously been applied to the <i>B</i> are now applied to the <i>A</i> , one rotation step at a time.
32	The complete match between the two <i>A</i> s is noticed, the focus is removed from them, they are marked as done, the number of completely matching letters is incremented to 2.
33-47	The same processes that were applied to the <i>A</i> s in Cycles 18-32 are now applied to the <i>G</i> , except that only the <i>G</i> s become focused, so there is no need to remove a focus marker from any other pair.
48	The fact that three letters have been completely matched is noticed, and the statement <i>SAME</i> is made and processing stops.

trajectory with the same initial pair of letters, or they try restarting with an alternative pair of matching letters. According to this explanation, the relatively unsystematic variation in the *different* response times is largely due to fluctuations in the criterion for judging two figures as being *different*. The response times for the various different problem types have not been analyzed because the appropriate trajectories (i.e., ones that will bring all the faces of the two cubes into congruence) are not well defined.

*Summary.* The cognitive coordinate system used by the low-spatial subjects consists of axes that are parallel to the main axes of the cubes, which in turn correspond to the main horizontal, vertical, and depth axes of the visual world. The cube was represented as upright and aligned with the major horizontal and vertical axes of the visual environment. Each letter's orientation and location was coded relative to an external upright. To determine if the two figures depict the same cube, these subjects mentally manipulated one cube in order to equalize these representations with respect to the extrinsic axes. The representation included a viewing point defined by the perspective in the drawing, such that the front, right, and top faces of the cube are visible.

What determined this cognitive coordinate system—gravitational or retinal frames of reference, or the standard axes of the visual environment, some local cues (like the edge of the display screen), or the cube's internal structure? The influence of gravity alone may be minor because the interpretations of the cubes seem unchanged if they are viewed while lying on one's back and looking upwards. Disparity between other frames of reference seems not so easily ignored. Tilting one's head to disassociate the retinal frame of reference from the vertical axis defined by the cube or the visual environment seems to introduce a conflict that has no dominant resolution. However, examining what people do when faced with conflicting frames of reference does not unambiguously indicate how they use each frame when it is congruent with the others. Even if viewers ignore a particular frame of reference when it conflicts with others, this does not mean that they do

not attend to it when it is congruent. If the various frames do not coincide, then people may be flexible in choosing a coordinate system that is congruent with the greatest number of possible frames or with one that is dominant in some way. In the problems we presented, the frames of reference defined by the cube, the visual environment, gravity, and the retina all coincided so any one of them or any combination of them could have been the determinants.

### *Sources of Individual Differences*

A priori, some processes seem more likely to differentiate subjects of differing spatial ability. Computing and executing the trajectory necessary to bring the first pair of matching letters into congruence (Process 2) could differentiate subjects because there have been previous reports that low-ability subjects take longer in mental rotation tasks (Snyder, cited in Posner, 1973), although contrary results also have been reported (Egan, 1978). Confirmation (transforming the remaining letters to determine that they match their counterparts) might also be expected to differentiate subjects. By contrast, finding a pair of matching letters should not differentiate subjects because this can be done without any spatial manipulation. The analysis of the gaze durations into component processes, shown in Figure 5, suggests two major sources of individual differences, initial rotation (Panel B) and confirmation (Panel C). There is essentially no difference between the two groups with respect to the search process (Panel A).

*Search.* Panel A indicates that the two groups do not differ in the time they spend in the search for the first pair of matching letters. The search time differs across the six conditions,  $F(5, 75) = 9.81, p < .01$ , primarily because the  $0^\circ$ -3 Matches condition requires almost no search. There is no interaction between the six conditions and spatial ability.

*Rotation and confirmation.* Panel B shows that the low-spatial subjects take longer (663 ms longer, on average) in the initial rotation of matching letters,  $F(1, 15) = 14.87, p < .01$ . There are differences among the 6 conditions,  $F(5, 75) = 3.28, p < .01$ , such that

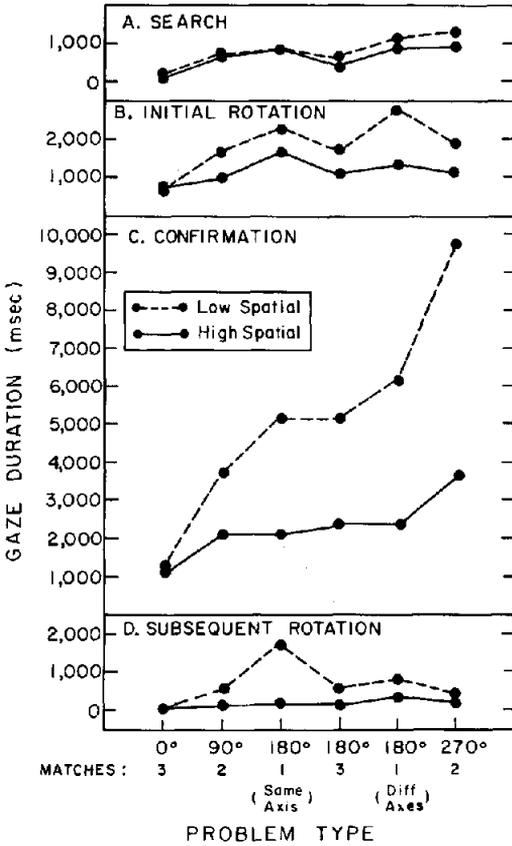


Figure 5. Mean gaze duration of various processes in same trials of the Cube Comparisons task for low-spatial and high-spatial subjects.

the initial rotation time is generally longer in problems that entail longer rotation trajectories. The interaction between condition and spatial ability is not significant.

The largest individual differences appear in confirmation, shown in Panel C. The low-spatial subjects take much longer than the high-spatial subjects (2,928 ms longer, on average),  $F(1, 15) = 35.89, p < .01$ . There are large differences among the problems,  $F(5, 75) = 6.76, p < .01$ , and the interaction is significant,  $F(5, 75) = 2.50, p < .05$ .

Panel D shows the time spent in subsequent rotation episodes, that is, rerotation of a cube face that had already been rotated into congruence previously. The high-spatial subjects spent very little time in subsequent rotation (216 ms), whereas the low-spatial subjects

spent considerably more time (705 ms), a 489 ms difference,  $F(1, 15) = 9.00, p < .01$ . The low-spatial subjects may rotate the same cube face for a second time because they have forgotten the product of their original computation. The six problems differ in the amount of subsequent rotation they required,  $F(5, 75) = 3.36, p < .01$ , with the most occurring for the low-spatial subjects in the 180°-1 Match(same) condition, resulting in an interaction between problems and groups,  $F(5, 75) = 4.33, p < .01$ .

The difference between the rotation strategies used by the high- and low-spatial subjects can be viewed in terms of a difference in cognitive coordinate systems. The subjects who rotated around a nonstandard axis were using that nonstandard axis as one of the axes of their cognitive coordinate system, at least during the time when they were rotating. The low-spatial subjects almost never used a cognitive coordinate system that did not closely correspond to the cubes' axes or to the main axes of the visual environment.

*Rotation rate.* The initial rotation times can indicate whether the low spatial subjects mentally rotate at a slower rate than high spatial subjects. The comparison can be made by considering only those three trial types in which the high and low spatial subjects all used the same trajectories, namely those that do not permit rotation around a nonstandard axis—the 0°-3 Matches, 90°-2 Matches, and 180° (same)-1 Match conditions. The slope is 826 ms/90° for the low spatial subjects, and 435 ms for the high spatial subjects,  $F(2, 108) = 2.95, p < .06$ , indicating the low spatial subjects may mentally rotate at half the rate of the high spatial subjects.

*Processing Model for High-Spatial Subjects*

The simulation model of the high-spatial subjects differs only in that it rotates in 30° steps and is not restricted to rotations around the standard axes. As we have described, three of the trial types (180°-3 Matches, 180° (different axes)-1 Match, and 270°-2 Matches), have alternative, shorter trajectories that are illustrated in the right-hand column of Figure 1d, 1e, and 1f. During the rotation process, rather than going through two sepa-

rate phases that equalize first the location and then the orientation of the two matching letters, the model equalizes both aspects of the letters with a single rotation phase, due to the ability to rotate the cube around arbitrary axes. In other words, the axes of the cognitive coordinate system need not correspond to any of the standard axes. Any axis that is useful for performing a rotation can become a temporary axis of the cognitive coordinate system. The task-defined axis for rotation is selected by specifying both the orientation and location of the source and destination letters. For example, in Figure 4, the two *B*s can be completely aligned by rotating the right *B* 120° around the oblique axis.

The processes that precede and follow initial rotation have the same structure for the high- and low-spatial subjects. The search process is identical in the two cases. The confirmation process has the same control structure in the two cases, but the actual rotation performed during confirmation mirrors the initial rotation in its step size and choice of axes.

*Rotation steps and step sizes.* Several considerations suggest that the rotation around a given axis is done in steps, rather than by a single transformation. First, the eye-fixation behavior indicates a sequence of alternating fixations between the parts that are being rotated into each other, with the number of switches monotonically related to the rotation angle (Just & Carpenter, 1976). We have interpreted these switches in fixation as indices of a boundary between successive rotation steps, although there need not be a switch in fixation after each step. The theoretical factor associated with these switches is that the rotation process in the proposed model is guided rather than ballistic. So the subject need not know ahead of time how far to rotate. He or she can rotate in steps and apply a comparison operation after each step to determine whether the two representations are sufficiently similar in orientation. Another empirical consideration suggesting that rotation is in steps is that the rotated object is represented at orientations intermediate between the initial and final orientation in the course of a rotation (Cooper & Shepard,

1973). In the proposed model, such intermediate representations are produced by each step of the stepwise rotation. A final theoretical factor is that stepwise (rather than continuous) rotation is more congenial to a production system's operation, which inherently segments the processing into discrete recognize-act cycles. In such a system, each rotation step corresponds to the firing of a rotation production that changes the represented orientation of some object by a given number of degrees in each cycle.

Step sizes of 15° and 30° were used to model the low- and high-spatial subjects, respectively, a ratio of 1:2 in the amount of rotation per CAPS cycle. The reason for this ratio was that the low-spatial subjects rotated approximately half as far per unit time as the high-spatial subjects, as indicated by the ratio of their initial rotation durations on problems that did not permit nonstandard axes. A second factor determining step sizes of 15° and 30° was the response time in the identity condition, which involved no rotation, compared to the response times in the other conditions. The identity condition took subjects approximately 2 s to perform, and required nine CAPS cycles, providing an estimate of a little over 200 ms per cycle during this interval. The theoretical assumption is that all CAPS cycles take the same amount of time, so if the duration of some other interval is known, the number of elapsed CAPS cycles can be inferred. The particular intervals of interest here were the extra amounts of time required (relative to the identity condition) to perform the rotation in the various trial types. For example, if a trial type required 7,200 ms more than the identity condition, then we can infer that the rotation process consumed approximately thirty-six 200-ms cycles. Moreover, for most of the trial types, we know how many degrees of rotation are required. If the trial in the example required that each of three cube faces be rotated through 180°, for a total of 540° of rotation, we can infer that the size of the rotation step in each of the 36 cycles was 15° (ignoring the few cycles in the interval that did not include rotation). Using this method, we estimated that the rotation step sizes for the low- and high-spatial subjects were ap-

proximately  $15^\circ$  and  $30^\circ$ , respectively. These estimates are not precise, but are the largest common divisors of all the required rotations (90, 120, 180). A slightly better fit to the data can be obtained by allowing rotation steps that are not necessarily divisors of the final trajectory, and then rotating the letter until it gets close enough.

There are several possible theoretical explanations of the high-spatial subjects' faster rotation rate. One possibility is that both groups use the same representations and that high-spatial subjects simply execute a basic mental operation faster. A second possibility is that the speed of the operation depends on the nature of the operand, and that high-spatial subjects use a more economical code to represent the figure that permits faster execution of the rotation and comparison operations. A third possibility is that the rotation is performed incrementally, in steps, and that both groups take the same amount of time for each step, but that the steps (i.e., the rotation angle per increment) is larger for the high-spatial subjects. The simulation models instantiate the third alternative, namely that the rotation is stepwise and high-spatial subjects rotate faster because they have a larger step size.

Almost all the extra cycles of the low-spatial model (compared to the high spatial) are accounted for by the rotation productions that are involved in initial rotation and confirmation. Over 96% of the low-spatial model's penalty is paid in rotation time, because the rotation and comparison is slower and because in some cases the trajectories are longer. This echoes the result of the human performance that most of the time difference between the low- and high-spatial subjects was attributable to initial rotation and to confirmation.

A few productions that select rotation trajectories also distinguish the low- and high-spatial strategy. The high-spatial model has a single production that notices two matching focused letters that mismatch in location or orientation, or in both, and takes the action of computing the axis and direction of rotation that would eliminate the mismatch. By contrast, the low-spatial model has one production that notices a mismatch in location and a separate production that notices mis-

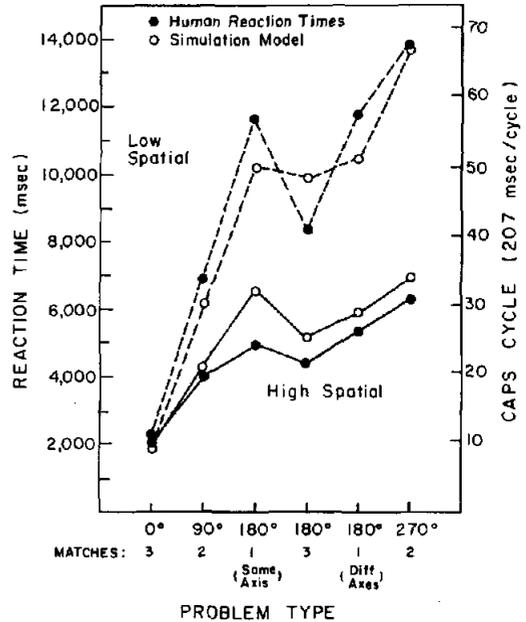


Figure 6. Human reaction times (left axis) and simulation model cycles (right axis) in the Cube Comparisons task.

match in orientation, given that the locations are the same. Otherwise, the two production systems are almost identical. The actual rotation productions are the same in the two models, except for the size of the rotation step.

*Quantitative comparison between models and data.* There is a close correspondence between the number of CAPS cycles and the response times for each of the six trial types, for both the low-spatial and high-spatial subjects, as shown in Figure 6. Making the comparison between the model and the human data was slightly complicated by the fact that the data obtained from the high-spatial subjects represents a mixture of two strategies on some of the problem types. On those problems that permitted nonstandard trajectories, the high-spatial subjects rotated around nonstandard axes approximately 81% of the time and around standard axes about 19% of the time, as indicated by the relative frequencies of the retrospective reports. The cycle count plotted for the high-spatial model for these problems consists of a corresponding mixture of two models. The mixture is a

weighted average of the high-spatial and low-spatial models, with weights of 81:19, and with both models using 30° step sizes. The actual cycle counts of the pure high-spatial model for the problem types represented in Figure 1d, 1e, and 1f, were 23, 26, and 30 cycles, respectively.

To obtain a quantitative comparison between the model and the data, a linear regression analysis was run in which the dependent variable was the human response time and the independent variable was the number of CAPS cycles used, as plotted in Figure 6. This regression accounted for 94.2% of the variance among the 12 means. When a zero intercept was forced, the analysis produced a regression weight of 207 ms per CAPS cycle (and 211 ms without the forced zero intercept).

#### *Orientation-Free Description Strategy*

The single high-spatial subject who reported a nonmanipulative strategy said that he always encoded the relations between letters on the same cube (e.g., the bottom of the *P* points toward the top of *L*), compared the codes for the two cubes, and coded a second relation (e.g., the back of the *G* points to the front of the *L*). This representation is generated within an object-defined cognitive coordinate system and so it will be invariant with the object's orientation in space. Consequently, no mental transformation is required to equalize the orientations of the two cubes in any of the problems. Thus it is not surprising that this subject's response times showed relatively little effect of problem difficulty as defined by the amount of rotation required. His response times in the five nonidentity conditions lay between 8,000 and 10,000 ms, considerably slower than the high-spatial subjects but still slightly faster than the low-spatial subjects. His error rates were 5.6% and 11.1% for *same* and *different* trials, respectively. The existence of this strategy illustrates that tasks ostensibly requiring spatial manipulation can sometimes be effectively performed without manipulation if the appropriate cognitive coordinate system is used.

#### *Perspective-Change Strategy*

In addition to the use of orientation-free descriptions and the rotation strategies, another strategy, perspective change, can be used to solve the problems in the Cube Comparisons test, the Vandenberg Mental Rotation test, and in similar tasks. Even though this strategy happened not to be observed among our subjects, it is a theoretical possibility that some subjects might use it and we can list some of the factors that govern its use. In the perspective-change strategy, the object's orientation in space is kept constant, but there is a change in the representation of the viewing point, and hence the represented view of the object. In this case, the object's position and the observer's position are both coded within a cognitive coordinate system that includes both the observer and the object, and whose origin corresponds to the object's position. In the Cube Comparisons task, for example, one can imagine how the right-hand cube in Figure 1f would look when viewed from directly below. That view is consistent with the view depicted on the left, and so the correct response is *same*. The rotation axis (the *x* axis) in this example is one of the three standard ones. Future experiments will have to tell us whether any subjects can mentally change perspective around an arbitrary task-defined axis.

Although mental rotation and perspective change are algebraically equivalent, there are several ways in which the two psychological processes seem to differ. First, they appear to be used selectively for different types of stimulus objects. If the object is small, mobile, and manipulable, (like a child's alphabet block), then a mental rotation strategy is more likely to be evoked. By contrast, if the object is large and immobile, like a building or a room, then people are more likely to mentally keep it stable and imagine their own position changing. A common demonstration of this phenomenon is that people who are asked to mentally count the number of windows in their house consistently report taking a mental walk around or through the house, rather than imagining the house rotating while they remain stationary. Perspective change may be more prevalent in navi-

gation, which requires manipulation of one's own position relative to stable parts of the environment (Kuipers, 1978). A second difference is that mental rotation is sometimes accompanied by an imagined manipulation of the object with one's hands. By contrast, perspective change involves an imagined transformation in body position that is sometimes accompanied by reports of proprioception of such a change (Carpenter & Just, 1982). A third distinction is that children of a particular age can perform a mental rotation task but cannot perform an equivalent perspective-change task (Huttenlocher & Presson, 1973). A fourth possible distinction is that mental rotation produces intermediate representations that correspond to intermediate orientations of the rotated object that lie between the initial and final orientation (Cooper, 1976; Cooper & Shepard, 1973). By contrast, it seems possible to take opposite perspectives without passing through intermediate stages (Hintzman, O'Dell, & Arndt, 1981).

This account of strategy differences can be generalized to other spatial processes besides rotation, such as size scaling (Bundesen & Larsen, 1975). In the size-scaling paradigm, the subject is shown two figures that differ in size and is asked to judge if they are the same or different. The response time increases with the ratio in size difference and this has been interpreted as reflecting a mental size-scaling operation analogous to mental rotation. There is, in addition, the possibility of a size-free representation, analogous to the orientation-free representation, that would permit direct comparison without regard to size. Finally, it is possible to perform the task using a process analogous to perspective change, by having the viewer imagine a change in his distance from the object, moving either nearer to or farther from one of the objects, until the mental visual angle subtended by the two objects is similar. Thus the theory developed in the domain of mental rotation may provide a more general framework that appears applicable to size scaling, and perhaps to other spatial processes as well.

### *Spatial and Linguistic Processing Systems*

In the widespread discussion of the diversity of mental processes (e.g., verbal-pictorial,

analytic-Gestalt, left hemisphere-right hemisphere), there has been much emphasis on the distinctions between various families of processes, and relatively little consideration of the commonalities. Within almost any processing system, it is possible to categorize the basic processes into families, all of which share some characteristic. For example, in a standard digital computer, one can distinguish between arithmetic operations and logical operations. But they work in concert within a common architecture, can communicate with each other, and can collaborate on performing tasks that require the participation of both kinds of operations. Although it is certainly important to categorize the types of operations available to the human processing system, it is equally important to consider the larger system that can embrace different types of operations. The simulation model presented here, along with the model of human reading (Thibadeau et al., 1982), provides a demonstration that both spatial and linguistic processes of considerable complexity can be accommodated within a single processing environment. Mental rotation of a cube and comprehension of an embedded clause can both be accomplished within a CAPS framework and still comfortably conform to human performance characteristics. The particular properties of the CAPS framework that lend themselves to embracing different kinds of processes are its use of procedural knowledge that is completely modularized (in the form of productions) and a representational scheme capable of dealing with semantic, logical, and metric information.

### Generalizing the Theory to Other Tasks

The next section of the article describes two studies that generalize the approach in two respects. The first study shows that the model applies to the performance of a larger group of subjects performing a spatial psychometric test. The second study examines the generality of our characterization of high and low spatial subjects, by analyzing their performance in a spatial manipulation task that focuses on the process of rotation itself, namely the Shepard-Metzler (1971) task.

### *Comparison With Psychometric Test Performance*

To verify that the production system models provide satisfactory explanations of psychometric test performance, the performance in the laboratory task and the psychometric test were directly compared in the study reported below. The possibility exists that the processes in the laboratory task (and hence the models) are different from those in the psychometric test. Psychometric tests are usually paper and pencil tests, with a large number of problems presented for solution within an overall time limit, rather than individual problems presented one at a time under speed and accuracy instructions. Below, we briefly report a study that provides the desired verification, and shows that the models apply to the psychometric tests and hence the criterion tasks against which the tests are traditionally validated.

The experiment was run analogously to Experiment 1, except that eye fixations were not recorded. Also, the design was changed so that two thirds of the problems had matching letters that were ambiguous in orientation (e.g., O, S, N), as they are in a similar proportion of problems in the psychometric test. Ambiguity in orientation may influence the decision of whether to rotate or how far to rotate. For example, a subject could decide that two faces, each containing a perfectly round O, have corresponding orientations when, in fact, the faces differ by 90° or 180°. The subjects were 23 students who had not participated in the preceding experiment and who were not preselected for spatial ability. In addition to this laboratory experiment, two psychometric tests were administered, the Cube Comparisons test and the Vandenberg Mental Rotation test. Scores on the two psychometric tests were correlated,  $r(21) = .56, p < .01$ , indicating that the two tests tap some shared, as well as some nonoverlapping processes. The sum of their standardized scores on the two psychometric tests was used to group the subjects into three categories: 8 high-, 8 medium-, and 7 low-spatial subjects.

Subjects did tend to perform similarly in the laboratory experiment and in the Cube Comparisons psychometric test. Subjects who had a higher proportion of errors in the

psychometric test also tended to make errors in the experiment,  $r(21) = .58, p < .01$ . Subjects who attempted more problems in the psychometric test also tended to respond faster to problems in the experiment,  $r(21) = -.79, p < .01$ . (The speed measure in the experiment was obtained by computing the average response time for the nonidentity *same* problems.) The speed measure was also correlated with the Cube Comparisons total score,  $r(21) = -.69, p < .01$ , and the proportion of errors,  $r(21) = .46, p < .05$ .

The major contributor to the correlation between speed in the experimental task and performance in the psychometric test appears to be the speed of manipulating the cube, rather than the speed of nonmanipulative processes, such as encoding, response selection, and execution. The slope of the response time for the three problems without alternative trajectories (identity, 90 Degrees-2 Matches, and 180 Degrees(same)-1 Match) correlated with the psychometric score,  $r(21) = -.46, p < .05$ . By contrast, there was no significant correlation between psychometric scores and the response times in the identity condition (0 Degree-3 Matches), which requires only encoding, letter matching, and response selection and execution,  $r(21) = -.15, ns$ . Thus, the probable reason for the correlation between the mean time spent per problem in the experiment and the psychometric score is that the latter reflects the variability between subjects in how much time they take on those problems that require mental manipulation.<sup>5</sup>

Not only do the results show a convergence between the experimental and psychometric tasks, but the experiment provides a replication of Experiment 1. The response times, shown in Figure 7, follow the pattern found in Experiment 1. As the graph suggests, high-spatial subjects had a larger advantage in the nonidentity problems (because they can rotate

<sup>5</sup> These results differ from those of Egan (1978), who found no correlation between the psychometric score and the slope on mental rotation tasks, and a very slight correlation between the score and intercept. However, Egan's subjects were Navy pilot trainees, a group that may already have been selected for a high level of spatial ability, and may have shown less variability in manipulation time and strategies than did our unselected subjects.

faster) and in problems that permitted shorter, nonstandard trajectories,  $F(10, 100) = 1.93$ ,  $p < .05$ . The presence of the orientation ambiguity increased the response times, especially for the low-spatial subjects on the more difficult problems,  $F(10, 100) = 2.20$ ,  $p < .02$ . The error rates for the high-, medium-, and low-spatial subjects were 7.8%, 9.5%, and 13.5%, respectively. As in Experiment 1, the retrospective reports indicated that the confirmation process was the major source of errors for all three groups of subjects. Also replicating Experiment 1, high-spatial subjects were more likely to report nonstandard trajectories. On those trials in which the trajectory could be categorized (using the same criteria as in Experiment 1), the high-spatial subjects reported nonstandard trajectories 49% of the time, compared to 24% and 6% for the medium- and low-spatial subjects,

$F(2, 20) = 9.09$ ,  $p < .01$ . Mental rotation was the most frequently reported strategy. Comparison of orientation-free descriptions was reported as the sole strategy on less than 5% of the same trials, but it was reported as the sole strategy on one third of the *different* trials, and the percentage was similar for each of the three ability groups. Several of the subjects first compared orientation-free descriptions to detect some types of inconsistencies, and if they found no inconsistency, they proceeded to use the mental rotation strategy.

This study confirmed the results of Experiment 1, that high spatial ability is associated with the use of shorter, nonstandard trajectories, faster rotation, and lower susceptibility to error. The convergence between the psychometric and the experimental measures suggests that the models developed for the experimental task generalize to the psychometric test.

#### *Individual Differences in the Shepard-Metzler Task*

Unlike the Cube Comparisons task, the Shepard-Metzler task is less open to alternative strategies. The rotations are always around a single axis in any one trial, so there are no short-cut trajectories. Although it is possible to perform the Shepard-Metzler task by using orientation-free descriptions and doing no spatial manipulation, naive subjects seldom develop the appropriate descriptions in the course of one or two experimental sessions. Thus this task is likely to evoke the same strategy in all subjects. The prediction of the model is that low-spatial subjects should rotate at a slower rate than do high-spatial subjects, and should have more difficulty keeping track of their intermediate products, resulting in reinitializations of various processes.

The dimensions of variation of the stimuli included seven angular disparities (varied from 0° to 180° in 30° steps), three figure types, and the same-different variable. Due to an error in stimulus construction, there were four exemplars of stimuli at 30° and only two at 150°. The *different* trials were constructed by replacing one of the two figures with its mirror-image isomorph. The participants were the 4 high-spatial and 3 of the

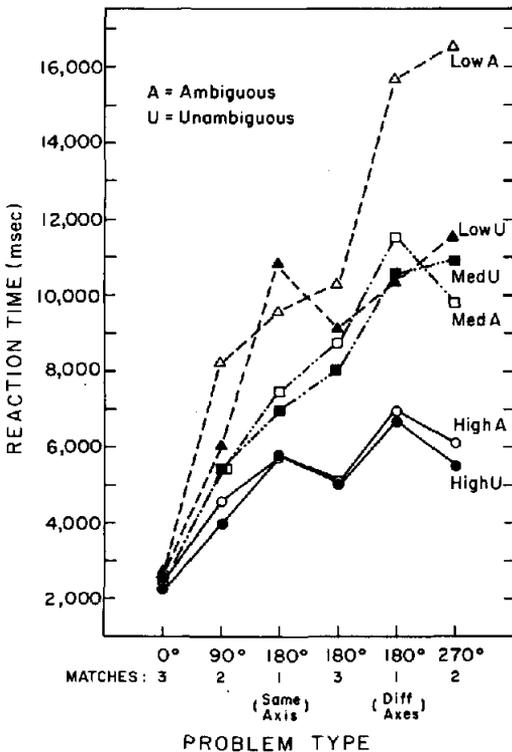


Figure 7. Reaction times for *same* trials for subjects classified as low, medium, or high spatial, for the unambiguous orientation problems (filled symbols) and the ambiguous orientation problems (unfilled symbols) in the Cube Comparisons task.

low-spatial subjects from the Cube Comparisons study reported above; the 4th low-spatial subject from Experiment 1 was unavailable for testing.

The eye-fixation protocols were divided into episodes associated with three main processes:

1. *Search* for potentially matching ends (terminal arms) of the figures,
2. *Rotation* of one of these parts until its orientation was similar to its mate's, and
3. *Confirmation* that the remaining parts of the figures were related by the same transformation that related the initially rotated pair.

Initial rotation was identified as the first pair or series of consecutive fixations between matching ends of the figures. Fixations that occurred before this episode were identified as search. Occasionally, subjects systematically looked back and forth between nonmatching ends of the figure prior to the initial rotation stage. In previously reported research (Just & Carpenter, 1976), this was categorized with the search behavior. In the current experiment it was categorized separately as incorrect initial rotation. After the initial rotation, subjects looked between the other two ends or sometimes scanned the entire figure. This was categorized as initial confirmation. Subsequent fixations between the ends that had been involved in the initial rotation were categorized as subsequent rotation. Subsequent fixations between ends involved in the confirmation stage were categorized as subsequent confirmation. The initial and subsequent episodes of a stage had to be separated by more than one fixation that did not fit the definition for that stage. Fixations that could not be categorized were tallied separately, but constituted a very small proportion of the data. Of the 147 *same* trials, only 5 could not be analyzed, 3 from high-spatial subjects and 2 from the low-spatial subjects. Another 22 trials were error trials or trials on which data were lost due to machine error.

*Results and discussion.* The pattern of response times and error rates for the *same* trials, shown in Figure 8, indicates that the performance of the low-spatial subjects was poorer, as one would expect. The low-spatial subjects' response times increased faster with angular disparity, and they had a higher

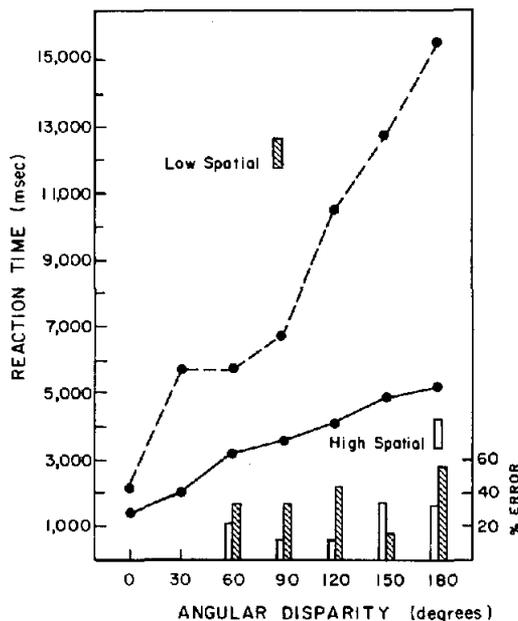


Figure 8. Reaction times and error rates for the *same* trials for the low-spatial and high-spatial subjects in the Shepard-Metzler task.

intercept, as indicated by the reliable difference between the best-fit lines for the high- and low-spatial groups,  $F(2, 113) = 57.82$ ,  $p < .01$ . The gaze durations discussed below help to localize these differences. The error rates for the low-spatial subjects were 26.1% and 18.6% for the same and different trials, respectively, and for the high-spatial subjects, 15.7% and 18.6%.<sup>6</sup>

<sup>6</sup> Two statistical analyses were performed on the response times and gaze duration measures from only those trials that had correct responses and scorable eye-fixation protocols. One analysis was a multiple linear regression, with angular disparity as the independent variable. This procedure is applicable because the rotation angle increases linearly across trial types for both groups of subjects. Separate regression analyses were performed on the high-spatial subjects, the low-spatial subjects, and the two groups combined, hence deriving the reduction in the residual sum of squares due to grouping by ability. The second analysis was a standard ANOVA on the means of the three or fewer usable observations of each subject in each cell. The independent variables were ability level and angular disparity. The results from the two analyses were generally similar, and we will report only on the first analysis.

The analysis of the gaze durations shown in Figure 9, indicated that the two groups of subjects differed primarily in the time they spent on initial rotation and initial confirmation. As Panel C of Figure 9 indicates, the slope of the low-spatial subjects in initial rotation was twice as steep (more precisely, 2.3 times as steep) as for the high-spatial subjects, and the intercept was slightly higher, producing a reliable difference between the two groups,  $F(2, 113) = 18.63, p < .01$ . This replicates the result from Experiment 1 that this group of low-spatial subjects mentally rotates half as fast as the high-spatial subjects, and generalizes it to a slightly different task. Also, the times for initial rotation increased reliably as a function of angular disparity for both the high- and low-spatial groups,  $F(1, 72) = 13.16, p < .01$ , and  $F(1, 41) = 6.53, p < .02$ , respectively.

Initial confirmation (Figure 9, Panel D) produced a very similar pattern of results,

with a reliable difference between the two groups,  $F(2, 113) = 18.13, p < .01$ . The time in initial confirmation increased reliably as a function of angular disparity for both the high- and low-spatial subjects,  $F(1, 72) = 5.62, p < .02$ , and  $F(1, 41) = 9.77, p < .01$ , respectively.

Similarly, subsequent rotation time and subsequent confirmation (Figure 9, Panels E and F) increased with angular disparity for both the high- and low-spatial subjects,  $F(1, 72) = 11.00$  and  $12.81, p < .01$ , and  $F(1, 41) = 11.61$  and  $11.66, p < .01$ , respectively. For each category, the difference between the high- and low-spatial subjects was also significant,  $F(2, 113) = 10.51$ , and  $F(2, 113) = 15.38$ , respectively, both  $p < .01$ .

The gaze duration attributable to search (Figure 9, Panel A) increased with angular disparity for the high-ability group,  $F(1, 72) = 34.08, p < .01$ , and for the low-ability group,  $F(1, 41) = 12.21, p < .01$ . However, consistent with the theoretical analysis, the two groups did not differ significantly from each other in the search process,  $F(2, 113) = 2.92, p > .05$ . The gaze duration attributable to incorrect rotation did not significantly increase with angular disparity for either group of subjects, nor was the difference between the two groups significant,  $F(2, 113) = 2.45$ . Incorrect rotation occurred when a subject repeatedly looked between noncorresponding ends of the figure. One reason that this analysis indicates relatively little time spent on this process and no reliable group difference is that the data here are based only on correct responses. Often when subjects looked between noncorresponding ends of the figure, they eventually responded incorrectly, as the analysis of errors shows.

In a subsequent follow-up study, we obtained very similar results with 5 subjects who were high spatial, as defined by the psychometric battery. Their response times and gaze durations followed the same function of angular disparity as did the high-spatial subjects described above, even to the values of the slopes. The close similarity in the parameters suggests that the results, although based on relatively few subjects, are generalizable to other subjects of similar psychometric skill.

It is interesting to note that the durations of initial rotation at  $0^\circ, 90^\circ$ , and  $180^\circ$  (Figure

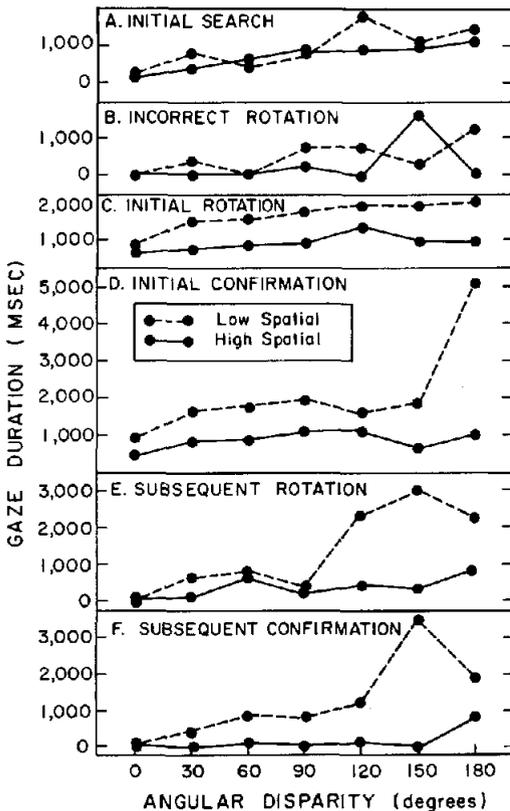


Figure 9. Mean gaze duration of various processes in same trials of the Shepard-Metzler task for low-spatial and high-spatial subjects.

9, Panel C) resemble the corresponding durations observed in the Cube Comparisons task for the 0°, 90°, and 180° (same) problems (Figure 5, Panel B), particularly for the low-spatial subjects. This resemblance is consistent with the hypothesis that subjects rotate just a part of a skeletal representation, so that rotation times should be similar across figure types for a given subject. This result must be interpreted with caution because the data are too sparse to provide a sensitive test of the hypothesis of no difference in rotation times between the two experiments. Of course, among the many reaction time experiments in the literature there is a great deal of variability in rotation rates, variability that may largely be due to subject differences, strategy differences, practice differences, and the inclusion of processes other than initial rotation in the slopes of the total reaction times.

In summary, the eye-fixation results indicate that low-spatial subjects take longer to perform a mental rotation task (increasingly longer at greater angular disparities) because their rotation rates are slower and because they are less efficient at mentally keeping track of their work in more demanding problems. Their poor bookkeeping forces them to do extra work, occurring in the episodes we have called *subsequent rotation* and *subsequent confirmation*.

*Analysis of errors.* Figure 8 shows the distribution of errors in the *same* trials. An analysis of the eye-fixation protocols suggested that many of the errors occurred when a subject initially chose to rotate two ends that did not match and never discovered which ends did match. We counted the number of trials in which subjects looked only between matching ends, only between nonmatching ends, or between both, and then cross tabulated this factor with response accuracy, as shown in Table 5. Both high- and low-spatial subjects generally responded correctly when they looked only between matching ends, but they generally responded incorrectly when they looked only between nonmatching ends, pseudo  $\chi^2(1) = 66.40, p < .01$ . Thus a major source of errors on *same* trials appears to be the incorrect pairing of nonmatching ends during the search process.

Experimental analyses of individual differences, such as the present one, are typically

Table 5  
*Frequency of Correct and Incorrect Same Trials With Fixations Between Matching and Nonmatching Ends*

Subjects and response	Matching ends	Both pairings	Nonmatching ends
Low spatial			
Correct	37	5	1
Error	5	2	10
High spatial			
Correct	68	3	1
Error	4	0	6

based on many fewer subjects than are traditional psychometric investigations because data collection and analysis is so much more demanding, particularly in eye-fixation studies. Although the eye-fixation studies reported here are based on only 8 subjects, we have independently replicated the major results of the Cube Comparisons study in several pilot studies and those of the Shepard-Metzler study in a follow-up experiment. The reliability is also confirmed by the convergence between studies, reported previously. Part of the reason for the replicability is that we chose subjects at known points on a psychometrically determined dimension. Finally, it is not essential to study large groups of subjects to document different strategies, although larger groups could indicate the relative frequency of the strategies with more precision.

What this experiment indicates is that a very similar account of individual differences applies to both the Cube Comparisons task and the Shepard-Metzler task. Although the Shepard-Metzler task is not as open to alternative strategies, the high- and low-spatial subjects did differ in rotation rate, in having to reexecute parts of the process, and in error patterns, much as they did in the Cube Comparisons study. The results are also entirely consistent with our previously described model for the Shepard-Metzler task (Just & Carpenter, 1976). According to this model, subjects use a skeletal representation, consisting of pipe-cleaner-like vectors that correspond to the major axes of each segment of the figure. The cognitive coordinate system within which the figures are represented is the standard environmentally defined one.

The axis of rotation always corresponded to the environmentally defined depth axis. Thus there was no opportunity for the task to define some alternative arbitrary axis that high-spatial subjects might use for rotation.

### Coordinate Systems and Strategies in Spatial Thinking

The type of mental operation performed in spatial tasks is intertwined with the cognitive coordinate system that is used to code the object. The three different coordinate systems observed in our experiments led to three different processes: mental rotation around standard axes, mental rotation around task-defined axes, and comparison of orientation-free descriptions. In addition, one other possible strategy that was not observed could have led to a solution by imagining a change in perspective. In this section of the article, we briefly examine the differences among the different processes, focusing on the differences in how spatial information is treated.

#### *Orientation-Free Descriptions Versus Mental Rotation*

In all three experiments the subjects' task is to determine whether two drawings depict the same object. In all strategies, subjects construct a representation of the object depicted by the two drawings, and compare them. The strategy of comparing orientation-free descriptions is different from the other strategies, because it seems to allow a subject to perform a spatial task while circumventing the need for spatial transformation. The subject in Experiment 1 who used orientation-free descriptions in the Cube Comparisons task coded the orientation of one letter relative to another on the same cube, without reference to any larger frame of reference external to the cube. In other words, the cognitive coordinate system was defined entirely by the cube itself. A representation developed within an object-defined coordinate system will be invariant under object rotation. Consequently, the representations for the left and right cubes can be directly compared without any mental rotation.

The relationships among the parts of an object must be very completely understood

if they are to be used as the basis of an object-defined cognitive coordinate system. The subject in Experiment 1 who compared orientation-free structural descriptions often gave evidence of such understanding, indicating that he had integrated the information from the two drawings of the cube to completely infer the structure of the cube. For example, in the course of solving items like the one shown in Figure 1b, he would often say "so the *J* could be opposite the *P*." By contrast, the subjects who used mental rotation did not make such comments. The rotators seemed to be using an algorithm that was effective in this task, but it did not necessarily require or produce a complete knowledge of the cube's structure. Thus the two kinds of coordinate systems may be associated with differences in how well the representation of the object is integrated.

Orientation-free representations also exist for the Shepard-Metzler figures (Metzler & Shepard, 1974). For example, one can construct an orientation-free description by taking an imaginary walk through the interior corridors formed by a Shepard-Metzler figure, assigning some local orientation (e.g., marking one of the four sides of the corridor as the floor) and coding each bend in the corridor as a turn to the left, right, up, or down, as one mentally walks from one end of the figure to the other. (The analogy of a mental walk is used here only to indicate the nature of the resulting representation, and is not meant to imply that subjects who form this type of code imagine themselves taking a mental walk. In particular, we suggest that the process by which the representation is formed requires no spatial transformation.) This kind of representation appears to be difficult to construct for Shepard-Metzler figures. Subjects seldom report representing Shepard-Metzler figures with orientation-free descriptions unless they have been instructed in how to construct the representation or have been given many hours of practice in the task.

The relative difficulty of constructing orientation-free representations for Shepard-Metzler figures suggests why mental rotation is often the preferred strategy. Mental rotation allows subjects to compare the structure of two objects in considerable detail without

completely understanding the structure of either one. The mental rotation strategies permit an approach of divide and conquer, by picking an object apart, representing each component within a coordinate system defined by the environment or by the task, and dealing with the object's components one at a time. The two cubes in Cube Comparisons are represented and compared one face at a time, without ever explicitly representing the relation between letters on adjoining faces. Within the rotation strategy, there is no necessity to encode the interpart relations. The difficulty in representing an entire cube or Shepard-Metzler figure would explain not only why many subjects choose to mentally rotate, but also why they would rotate only one part of the figure at a time. If they have difficulty in representing the structure of the entire figure at one time, then they would also have difficulty in rotating it all at one time.

Although the comparison of orientation-free descriptions allows spatial transformation to be circumvented, it does not necessarily detract from good performance in spatial tasks. The single subject in Experiment 1 who compared orientation-free descriptions had been classified as high spatial on the basis of his performance on a battery of spatial ability tests, so there is not much doubt about his ability to handle spatial information. In fact, one might expect primarily people of high-spatial ability to be able to construct complete orientation-free structural representations because this requires a more complete appreciation of an object's structure.

### *Psychometric Accounts of Spatial Ability*

The account of spatial ability that we propose can provide an alternative interpretation of previous psychometric results, as well as clarify a few mysteries within the psychometric literature. Psychometric research successfully established the existence of a spatial factor, by documenting significant individual differences in people's success in solving spatial problems of intermediate difficulty and distinguishing this factor from verbal and numerical factors (Smith, 1964). Beyond this, the psychometric literature on

spatial ability has been preoccupied with a controversy of whether spatial ability consists of a single unitary ability or several distinct component abilities. The controversy exists in part because some of the factors are not stable across populations or across tests, and because different researchers have sometimes used different labels to describe a factor arising from similar tests.

*Earlier descriptions.* Those psychometricians who have searched for separate components of spatial ability typically distinguish among two and sometimes three factors (see McGee, 1979, for a summary that is adapted from Michael et al., 1957). The first and clearest factor is often called *spatial visualization*. This factor is usually associated with tasks that elicit mental rotation, although the descriptions given by different psychometricians have varied somewhat. Of course, we must qualify this to take into account our own results showing that such tasks are typically performed with more than one strategy.

A second factor, sometimes called *spatial orientation*, has been described very differently by different psychometricians. We interpret this factor to be a mixture of using orientation-free descriptions and using perspective-change processes, and we attribute the disparate descriptions to the impurity. First consider those psychometricians who have regarded this factor in terms of perspective-change processes. Some of these researchers have suggested that the body orientation of the observer is an essential part of the problem (Thurstone, cited in Michael et al., 1957), consistent with our analysis of the perspective-change process. A typical marker test for this factor is the Guilford-Zimmerman Spatial Orientation test (Guilford & Zimmerman, 1947). In this test, the subjects are shown two photographs of a shoreline taken from a boat and are asked to imagine themselves looking over the prow of the boat. They are then asked what changes in the boat's orientation have occurred between the time the two photographs were taken. This format encourages some subjects to represent the perspective of the shoreline within a cognitive coordinate system defined by the visual world as seen from the boat, and to compute the transformation that caused a given change in perspective (Carpenter & Just, 1982).

Other psychometric investigators have described the spatial orientation factor in terms that are similar to the use of orientation-free representations. The descriptions of this factor imply the ability to assess the similarity of two objects that differ in orientation without mentally manipulating the representation of either one. For example, French (1951) described this factor as the ability to perceive spatial patterns accurately and to compare them with each other. Guilford and Lacey (1947, cited in Michael et al., 1957) described it as an ability to determine the relationships between different spatially arranged stimuli and responses and the comprehension of the arrangement of elements within a visual stimulus pattern. These are apt descriptions of the orientation-free description strategy used in the Cube Comparisons test. The possibility of performing this test with this strategy may explain why the test is sometimes thought to tap the spatial orientation factor.

*Alternate interpretation.* Our theory suggests that the varying psychometric descriptions of these factors may refer to three distinct processes engendered by the use of different coordinate systems. The visualization (rotation) factor may result from mental manipulation within a coordinate system defined extrinsically to the object. The object in this case is represented with respect to an axis that is usually provided by the visual environment or the retinal upright. The factor described as spatial orientation seems to be a mixture of two distinct processes—using orientation-free descriptions and perspective change. The orientation-free descriptions are generated within an object-referenced coordinate system, whereas the perspective-change strategy may result from a coordinate system that includes both the object and the observer, with the object at the origin.

*Other spatial tasks.* The proposed framework can also account for performance in seemingly unrelated spatial tests, like the surface development test. In this test, subjects are shown a two-dimensional unfolded layout of a hollow, three-dimensional object. Their task is to decide which of several drawings of three-dimensional foils matches the two-dimensional layout. The depicted object generally has one or more sides that contain a distinguishing feature, such as a figure, some

shading, or a notch. This test appears at first glance to require constructing a three-dimensional image, a sort of mental paper folding. But contrary to the first-glance analysis, the mental paper-folding process itself is probably not an important source of individual differences because the foils do not differ much in the structure of the three-dimensional object, so no difficult paper folding need be done.

We propose that performance in the surface development test depends largely on using orientation-free descriptions and on mental rotation, precisely the processes used in the Cube Comparisons test and the Vandenberg Mental Rotation test. Consistent with this proposal, we found that the surface development score was highly correlated with the Cube Comparisons test,  $r(28) = .82$ , and with the Vandenberg test,  $r(28) = .75$ , in a new group of 30 unselected subjects. The use of mental rotation is called for because the foils often differ in orientation from the unfolded layout, and so the subject has to mentally rotate the foils or the layout in order to compare their structure. Orientation-free descriptions are used to discriminate among the foils, which differ with respect to the presence and location of the distinguishing features on the sides of the layout and foils. Another group of subjects that gave think-aloud protocols while solving such problems clearly used orientation-free descriptions for this purpose. Thus mental rotation and orientation-free descriptions are used in the surface development test, the same strategies that occur in Cube Comparisons.

*Strategy variation.* The factor analysis methodology assumes that all subjects use the same general processes and structures on a test, and that differences among individuals arise because some people have more of the ability or because they use it more effectively. But the differences are construed as quantitative rather than qualitative. This assumption is incorrect, and its violation may account for many of the confusions in the psychometric literature. For example, French (1965) showed that different self-reported strategies (loosely characterized as global or analytic) in some psychometric tests resulted in different factor loadings.

Many spatial tests allow for more than one strategy, as we have demonstrated for the

Cube Comparisons test and for the Guilford-Zimmerman boat task (Carpenter & Just, 1982). Moreover, the test items sometimes systematically vary with respect to which strategy they evoke. For instance, in the Cube Comparisons test, certain kinds of *different* trials were less likely to evoke mental rotation than did *same* trials. This was possible because subjects could sometimes determine that two cubes were different using a simple feature-matching strategy, and they could then make the different response without mentally rotating. Similarly, in the Guilford-Zimmerman boat task, different strategies were used depending on whether the shoreline was tilted. Barratt (1953) also found variation in strategies in a number of spatial tests, particularly for more difficult items. Thus, there is likely to be both within-subject and between-subject contamination of the single-strategy assumption in all but the simplest tests. This contamination could cause a test to sometimes load on one factor and sometimes on another if the two populations tested had different strategy preferences. Many previous psychometric results are susceptible to these problems.

Even different versions of the same test can elicit different strategies. There exists a version of the Cube Comparisons test that uses simple geometric forms (such as arrows, circles, and pluses) in place of letters to distinguish the sides of the figures (Thurstone, 1938), that encourages greater use of orientation-free descriptions. A protocol analysis of 5 subjects solving problems from the Thurstone version indicated that the dominant strategy was the use of orientation-free descriptions. By contrast, in the lettered version we used in our main experiments, mental rotation was the dominant strategy and the use of orientation-free descriptions was a secondary strategy. Two versions of a test, which elicit different strategies, may still be described as essentially identical in the psychometric literature (cf. Karlins, Schuerhoff, & Kaplan, 1969).

The existence of multiple strategies may explain why it has been difficult to convincingly demonstrate the discriminant validity of the visualization and spatial orientation factors (i.e., that they are independent components of spatial ability). The correlations

between tests that are assumed to tap the two different factors are sometimes higher than those between tests assumed to tap the same factor (Borich & Bauman, 1972). In the psychometric tradition, this would suggest that the two factors are actually one. But a more likely interpretation, in view of our results, is that the strategies used in the visualization tests may overlap with those in the spatial orientation tests. Moreover, two tests of the same factor could encourage somewhat different processes. To determine whether the two factors are discriminable requires a more detailed analysis of the processes used in the individual tests, as well as a theory of what underlies the factors.

*Task complexity.* The degree of possible variation in strategies is closely related to the complexity or difficulty of the test. In a very simple spatial test that requires shape comparison (same or different) of two figures of the same size and orientation, there is not much opportunity for multiple strategies. The judgments are usually made without error; individual differences in the test reflect the speed of the comparison process (Ekstrom, French, & Harman, 1979). However, the individual differences in speed in such tests are not correlated with performance on the more difficult tasks that require more complex strategies and processes (Lohman, 1979). In much more difficult tests having a spatial format, like the Raven (1962) Progressive Matrices test, there are many possible strategies, and some items that are too difficult for most subjects. The spatial format of the Raven test is quite secondary to the induction processes used in the problem-solving aspects of this intelligence test. It is not surprising that scores in the extremely difficult tests correlate with other reasoning tests, rather than with other spatial tests. Even within a single type of test, item difficulty can affect which processes are elicited (cf. Lohman, 1979; Zimmerman, 1954). Zimmerman found that a Visualization of Maneuvers test composed primarily of simple items correlated with tests of perceptual speed, whereas a version composed of more difficult items correlated with tests of visualization and spatial orientation. Thus item and test difficulty may be major determinants of what strategies and processes will be evoked in a task that

appears to tap spatial ability. It would seem worthwhile to experimentally determine what stimulus characteristics govern the choice of strategy and then construct psychometric tests that systematically vary these characteristics.

### Summary

In summary, this analysis of spatial test performance has considered the nature of spatial representations and processes, as well as differences among individuals in how they are used. First, we have provided a theoretical account of the individual differences in spatial tasks, explaining in what way the high-spatial subjects are faster in their manipulation processes and more flexible in the cognitive coordinate systems they adopt. The CAPS production system framework was also used to consider a number of ways of construing individual differences in spatial cognition, as well as relating spatial cognition to other kinds of thinking. Second, we have documented two types of strategies that commonly occur in such tasks, using orientation-free descriptions and mental rotation, and described a third type, perspective change, that is used in spatial orientation tasks. We have suggested that these different processes arise from coding objects with respect to different coordinate systems. Third, we have suggested that these different coordinate systems, and the concomitant processes they engender, can help reconcile some of the traditional controversies in the psychometric literature on spatial ability.

### References

- Anderson, J. R. (1978). Arguments concerning representations for mental imagery. *Psychological Review*, 85, 249-277.
- Barratt, E. S. (1953). An analysis of verbal reports of solving spatial problems as aid in defining spatial factors. *The Journal of Psychology*, 36, 17-25.
- Baylor, G. (1971). *A treatise on the mind's eye: An empirical investigation of visual mental imagery*. Unpublished doctoral dissertation, Carnegie-Mellon University.
- Borich, G. D., & Bauman, P. M. (1972). Convergent and discriminant validation of the French and Guilford-Zimmerman spatial orientation and spatial visualization factors. *Educational and Psychological Measurement*, 32, 1029-1033.
- Bundesen, C., & Larsen, A. (1975). Visual transformation of size. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 214-220.
- Carpenter, P. A., & Eisenberg, P. (1978). Mental rotation and the frame of reference in blind and sighted individuals. *Perception & Psychophysics*, 23, 117-124.
- Carpenter, P. A., & Just, M. A. (1978). Eye fixations during mental rotation. In J. W. Senders, D. F. Fisher, & R. A. Monty (Eds.), *Eye movements and the higher psychological functions* (pp. 115-133). Hillsdale, NJ: Erlbaum.
- Carpenter, P. A., & Just, M. A. (1982). *Processes in solving the Guilford-Zimmerman Spatial Orientation test*. Pittsburgh, PA: Carnegie-Mellon University.
- Carpenter, P. A., & Just, M. A. (in press). *Spatial ability: An information processing approach to psychometrics*. In R. J. Sternberg (Ed.), *Advances in psychology of human intelligence* (Vol. 3). Hillsdale, NJ: Erlbaum.
- Carroll, J. B. (1976). Psychometric tests as cognitive tasks: A new "structure of intellect." In L. B. Resnick (Ed.), *The nature of intelligence* (pp. 27-56). Hillsdale, NJ: Erlbaum.
- Cooper, L. A. (1976). Demonstration of a mental analog of an external rotation. *Perception & Psychophysics*, 19, 296-302.
- Cooper, L. A., & Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance*, 2, 503-514.
- Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 76-176). New York: Academic Press.
- Corballis, M. C., Zbrodoff, J., & Roldan, C. E. (1976). What's up in mental rotation? *Perception & Psychophysics*, 19, 525-530.
- Egan, D. E. (1978). *Characterizing spatial ability: Different mental processes reflected in accuracy and latency scores*. Unpublished manuscript, Bell Laboratories, Murray Hill, NJ.
- Ekstrom, R. B., French, J. W., & Harman, H. H. (1979). Cognitive factors: Their identification and replication. *Multivariate Behavioral Research Monographs*, No. 19-2.
- French, J. W. (1951). The description of aptitude and achievement tests in terms of rotated factors [Special issue]. *Psychometric Monographs* (No. 5).
- French, J. W. (1965). The relationship of problem-solving styles to the factor composition of tests. *Educational and Psychological Measurement*, 15, 9-28.
- French, J. W., Ekstrom, R. B., & Price, L. A. (1963). *Kit of reference tests for cognitive factors*. Princeton, NJ: Educational Testing Service.
- Funt, B. V. (1983). A parallel-process model of mental rotation. *Cognitive Science*, 7, 67-93.
- Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York: Wiley.
- Ghiselli, E. E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psychology*, 26, 461-477.
- Guilford, J. P., Fruchter, B., & Zimmerman, W. S. (1952). Factor analysis of the Army Air Force's Shepard Field Battery of experimental aptitude tests. *Psychometrika*, 17, 45-68.
- Guilford, J. P., & Zimmerman, W. S. (1947). *Guilford-Zimmerman Aptitude Survey: Part V. Spatial orientation*. Beverly Hills, CA: Sheridan Supply.

- Hayes-Roth, F. (1979). Distinguishing theories of representation: A critique of Anderson's "Arguments concerning mental imagery." *Psychological Review*, 86, 376-382.
- Hinton, G. (1979). Some demonstrations of the effects of structural descriptions in mental imagery. *Cognitive Science*, 3, 231-250.
- Hintzman, D. L., O'Dell, C. S., & Arndt, D. R. (1981). Orientation in cognitive maps. *Cognitive Psychology*, 13, 149-206.
- Humphreys, G. W. (1983). Reference frames and shape perception. *Cognitive Psychology*, 15, 151-196.
- Huttenlocher, J., & Presson, C. (1973). Mental rotation and the perspective change problem. *Cognitive Psychology*, 4, 277-299.
- Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, 8, 441-480.
- Just, M. A., & Carpenter, P. A. (1979). The computer and eye processing pictures. *Behavioral Research Methods and Instrumentation*, 11, 172-176.
- Karllins, M., Schuerhoff, C., & Kaplan, M. (1969). Some factors related to architectural creativity in graduating architecture students. *The Journal of General Psychology*, 81, 203-215.
- Keenan, J. M., & Moore, R. E. (1979). Memory for images of concealed objects: A reexamination of Neisser and Kerr. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 374-385.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- Kosslyn, S. M. (1981). The medium and the message in mental imagery: A theory. *Psychological Review*, 88, 46-66.
- Kuipers, B. (1978). Modeling spatial knowledge. *Cognitive Science*, 2, 129-153.
- Levelt, W. J. (1982). Cognitive styles in the use of spatial direction terms. In R. J. Jarvella & W. Klein (Eds.), *Speech, place, and action* (pp. 251-268). New York: Wiley.
- Linde, C., & Labov, W. (1975). Spatial networks as a site for the study of language and thought. *Language*, 51, 924-939.
- Lohman, D. F. (1979). *Spatial ability: A review and reanalysis of the correlational literature* (Tech. Rep. No. 8). Stanford, CA: Stanford University, School of Education.
- Marr, D. (1982). *Vision*. San Francisco: W. H. 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.
- McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin*, 86, 889-918.
- Metzler, J., & Shepard, R. (1974). Transformational studies of the internal representation of three-dimensional objects. In R. Solso (Ed.), *Theories of cognitive psychology: The Loyola Symposium* (pp. 147-201). Potomac, MD: Erlbaum.
- Michael, W. B., Guilford, J. P., Fruchter, B., & Zimmerman, W. S. (1957). The description of spatial-visualization abilities. *Educational and Psychological Measurement*, 17, 185-199.
- Newell, A. (1973). Production system: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 463-526). New York: Academic Press.
- Pellegrino, J. W., & Kail, R. (1982). Process analyses of spatial aptitude. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol. 1, pp. 311-366). Hillsdale, NJ: Erlbaum.
- Posner, M. I. (1973). Coordination of internal codes. In W. G. Chase (Ed.), *Visual information processing* (pp. 35-73). New York: Academic Press.
- Pylyshyn, Z. W. (1973). What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Review*, 80, 1-24.
- Pylyshyn, Z. W. (1979). Validating computational models: A critique of Anderson's indeterminacy of representation claim. *Psychological Review*, 86, 383-394.
- Raven, J. C. (1962). *Advanced progressive matrices: Sets 1 and 2*. London: H. K. Lewis.
- Rock, I. (1973). *Orientation and form*. New York: Academic Press.
- Shepard, R., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703.
- Smith, I. M. (1964). *Spatial ability: Its educational and social significance*. San Diego, CA: Robert R. Knapp.
- Snow, R. E. (1980). Aptitude processes. In R. E. Snow, P. A. Fredericko, & W. E. Montague (Eds.), *Aptitude, learning and instruction* (Vol. 1, pp. 27-64). Hillsdale, NJ: Erlbaum.
- Snow, R. E., & Lohman, D. F. (1984). Toward a theory of cognitive aptitude for learning from instruction. *Journal of Educational Psychology*, 76, 347-376.
- Sternberg, R. J. (1981). Testing and cognitive psychology. *American Psychologist*, 36, 1181-1189.
- Thibadeau, R., Just, M. A., & Carpenter, P. A. (1982). A model of the time course and content of reading. *Cognitive Science*, 6, 157-203.
- Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University of Chicago Press.
- Yuille, J. C., & Steiger, J. H. (1982). Nonholistic processes in mental rotation: Some suggestive evidence. *Perception & Psychophysics*, 3, 201-209.
- Vandenberg, S. G. (1971). *Mental Rotation test*. Boulder: University of Colorado.
- Zimmerman, W. S. (1954). The influence of item complexity upon the factor composition of a spatial visualization test. *Educational and Psychological Measurement*, 14, 106-119.

## Appendix A

### Data Acquisition Procedures

#### *Display Graphics*

The line drawings of the stimulus figures were transformed into a computer representation by a digitizer that converted the output of a standard video camera into a  $256 \times 256$  gray-scale raster (Just & Carpenter, 1979). The stimuli were displayed to the subjects on a standard video monitor at a distance of 61 cm. The front face of each cube subtended approximately  $5.5^\circ$  of visual angle and the center-to-center distance between the cubes was  $10.5^\circ$ .

#### *Eye-Fixation Data Acquisition*

During the experiment, the subject's eye fixations were monitored by a Gulf + Western corneal-reflectance and pupil-center eye tracker. Readings of the  $x$  and  $y$  coordinates were taken every 16.7 ms, and if both the  $x$  and  $y$  coordinates were within  $1^\circ$  of the preceding observation, they were aggre-

gated with that observation. If either the  $x$  or  $y$  coordinate was not within  $1^\circ$ , the aggregation of the preceding set of readings was ended. The location of the aggregate was attributed to the modal  $x$ - and  $y$ -coordinate value of the readings contributing to the aggregate. The result was a series of fixations, usually over 200 ms in duration, separated by readings of 16.7 or 33 ms that could not be aggregated into either the preceding or subsequent fixations. These isolated readings of 33 ms or less reflected saccades and occasional noise, and were ignored in further analyses. Blinks that were preceded and followed by fixations at the same locus were included in the duration of the gaze at that locus. Blinks that occurred immediately before, during, or after a saccade, and the duration of the saccade itself, were not attributed to any locus. In the next step of analysis, fixations on the same face of a cube were aggregated into gazes attributed to that face.

## Appendix B

### Determining the Axis of Rotation

The three locations that were used to define the plane perpendicular to the axis of rotation (call them Locations I, J, and K) are locations at the midpoints of cube edges, chosen as follows. One of the locations, I, was the midpoint of the cube edge that was shared by the source and destination faces of the letter being rotated. Second, the ultimate destination of that point after rotation defines another location, J. The third location, K, is the current location of the point that will ultimately end up at Location I. In most cases, I, J, and K defined a plane whose normal provided the direction vector for the rotation axis. In other cases, all three locations coincided, and in those cases the rotation axis passed through that point. In all cases, the rotation axis also passed through the point at the center of the cube. This approach to axis finding can be generalized to apply to objects of any shape by computing the moments of inertia (Funt, 1983).

The axis-finding process can be further illustrated by working through an example, namely, in equating the locations of the  $B$ s in Figure 1d. The rotation will take the  $B$  from the top to the front face. First, the model uses the midpoint of the

shared edge between the top and front faces as Location I, and it notes that the part of the  $B$  that is nearest to I is the  $B$ 's right side, (where right side happens to be coded as  $270^\circ$  clockwise from the bottom). Then it determines where (i.e., near which edge midpoint of the destination face) the right side of the  $B$  will end up. Because changing the  $B$ 's location is not supposed to change its relative orientation, the right side of the  $B$  should remain near the midpoint of the bottom edge of the front face, which defines Location J. Similarly, the third location, K, is determined by finding the location of the point that will end up at Location I after rotation. Locations I, J, and K turn out to be the midpoints of the top edge of the front face, the bottom edge of the front face, and the top edge of the top face. These three locations define a plane parallel to the visible side of the cube, and the normal is parallel to the  $x$  axis. The normal that passes through the center is the  $x$  axis itself, and this is the axis of rotation.

Received February 2, 1984  
Revision received July 3, 1984 ■