Mental Rotation of Three-Dimensional Objects

Abstract. The time required to recognize that two perspective drawings portray objects of the same three-dimensional shape increases with the familiarity function of the angular difference in the portrayed orientations of the two objects and (ii) no shorter for differences corresponding simply to a rigid rotation of one of the two-dimensional drawings in its own picture plane than for differences corresponding to a rotation of the three-dimensional object in depth.

Human subjects are often able to determine that two two-dimensional drawings of the same object are different by rotating one picture plane to bring it into alignment with the other picture plane. This finding appears to be consistent with the fact that subjects require a momentary identity of shape as a function of the angular difference in the portrayed orientations of the two three-dimensional objects.

This angular difference was produced either by a rigid rotation of one of the two-dimensional drawings in its own picture plane or by a much more complex, nonrigid transformation, of one of the pictures, that corresponds to a rigid rotation of the three-dimensional object in depth.

This reaction time is found (i) to increase linearly with the angular difference in the portrayed orientations of the two objects and (ii) to be no longer for a rotation in depth than for a rotation merely in the picture plane. These findings appear to place severe constraints on possible explanations of how subjects go about determining whether or not two pictures of differently oriented objects are, however, consistent with an explanation proposed by the subjects themselves. Although introspective reports must be interpreted with caution, all subjects claimed (i) that to make the required comparison they first had to imagine one object as rotated into the same orientation as the other and that they could carry out this "mental rotation" at no greater than a certain limiting rate; and (ii) that, since they perceived the two-dimensional pictures as objects...
In three-dimensional space, they could imagine the rotation around whichever axis was required with equal ease.  

In the experiment, each of eight adult subjects was presented with 1600 pairs of perspective line drawings. For each pair the object was asked to match and describe a right-hand lever as soon as he determined that the two drawings portrayed objects that were congruent with respect to three-dimensional shape and to pull a left-hand lever as soon as he determined that the two drawings depicted objects of different three-dimensional shapes. According to a random sequence, in half of the pairs (the "same" pairs) the two objects could be rotated into congruence with each other (as in Fig. 1A and B), and in the other half (the "different" pair) the two objects differed by a reflection as well as a rotation and could not be rotated into congruence (as in Fig. 1C). The choice of objects that were mirror images or "twins" of each other for the "different" pairs was intended to prevent subjects from discovering some distinctive feature possessed by only one of the two objects and thereby reaching a decision of noncongruence with certainty, thereby avoiding any mental rotation. As a further precaution, the ten different three-dimensional objects displayed in these perspective drawings were chosen to be relatively unfamiliar and rearrangements of the objects, less likely to occur in three-dimensional space. Each object consisted of ten solid cubes attached face-to-face to form a rigid frame structure with exactly three right-angled "edges" (see Fig. 1). The set of all ten shapes included two subsets of five: within either subset, no object could be transformed into another by any or any other reflection or rotation (about 360°). However, each of these subsets in either subset of the image of one object in the other subset as required for the construction of the "different" pairs. For each of the ten objects, 18 different perspective projections—corresponding to one complete turn around the vertical axis by 360° steps—were generated by digital computer and associated graphical output (1). Seven of the 18 respective views of each object were then selected so as to avoid any views in which some part of the object was not commonly objects in the adjacent field and yet to permit the construction of two pairs that differed in orientation by as little possible, in the range from 0° to 180°. These 70 line drawings were then reproduced by photographic process and were attached to cards in pairs for presentation to the subjects. Half of the "same" pairs (the "depth"-pairs) represented two objects that differed by some multiple of a 20° rotation about a vertical axis (Fig. 1B). For each of these pairs, copies of two appropriately similar perspective views were simply attached to the cards in the orientations in which they were originally generated. The other half of the "same" pairs (the "picture-plane" pairs) represented two objects that differed by some multiple of a 20° rotation in the plane of the drawings themselves (Fig. 1A). For each of these, one of the seven perspectives was used as a "reference" for each object and two copies of this picture were attached to the card in appropriately different orientations. Altogether, the 1600 pairs presented to each individual subject included 800 "same" pairs, which consisted of 400 unique pairs (200 in each of the two different mirror-image groups) at each of the ten angular differences from 0° to 180°, of each which was presented twice. The remaining 800 pairs, randomly intermixed with these, consisted of 400 unique "different" pairs, each of which (again) was presented twice. Each of these "different" pairs corresponded to one "same" pair of either the "depth" or "picture-plane" variety in which, however, one of the three-dimensional objects had been reflected through the plane and terminated the visual display. The line drawings, which averaged between 4 and 6 cm in maximum linear extent, appeared at a viewing distance of about 60 cm. They were positioned, with the center of the card at a distance that subtended a visual angle of 9°, in two circular apertures in a black vertical backboard. 

The subjects were instructed to respond as quickly as possible while keeping errors to a minimum. For each individual, age 3 only 37% of the responses were incorrect (ranging from 0.6 to 3.7 percent for individual subjects). The reaction-time data presented below include only the 96.8 percent correct responses. However, the data for the incorrect responses exhibit a similar pattern. 

In Fig. 2, the overall means of the reaction times for the two sorts of "mental rotation in two-dimensional space," then the slope of the obtained functions indicates that the average rate at which these particular objects can be thus "rotated" is roughly 60° per second. 

In the course of the repeated reaction times necessarily include any times taken by the subjects to decide how to process the pictures in each presented pair as well as the time taken actually to carry out the process, once it was chosen. Moreover, because of the different subjects, the reaction times were still linear and were no more than 20° per second lower in the "same" blocks of presentations (in which the subjects knew both the axis and the direction of rotation) compared with the "rotation" blocks of each presentation) than in the "mixed" blocks (in which the axis of rotation was unpredictable). Tentatively, this suggests that 80 percent of a typical one of these rotation times may represent some process such as "mental rotation" itself, rather than a preliminary process of presentation or search. Furthermore, in no case was there any indication of response time or classification of this point and others. 

Roger N. Shepard
Jacqueline Metzler
Department of Psychology, Stanford University, Stanford, California 94305

Figure 1. Examples of pairs of perspective line drawings. (A) A "same" pair, which differs by an 80° rotation in depth; (B) a "same" pair, which differs by an 80° rotation in depth; and (C) a "different" pair, which cannot be brought into congruence by any rotation.

Figure 2. Mean reaction times to two perspective line drawings portraying objects of the three-dimensional shape. Times are plotted as a function of angular difference in portrayed orientation. (A) For the subject differing by a rotation in the picture plane only; and (B) for the same subject differing by a rotation in depth. (The center of the circles indicate the means; and the lines extending far enough to show outside these circles, the vertical bars around each circle indicate the range of variation from linearity. No significant quadratic or higher-order effects were found for these functions.) The angle through which different three-dimensional shapes are to be rotated to achieve congruence is not, of course, determined by the principles in depth perception, but the one like those plotted in Fig. 2 cannot be constructed in any straightforward manner for the "different" pairs. The overall mean reaction time for these pairs was found, however, to be 3.8 seconds—"for the same" objects, as the responses in the various conditions of the following presentation times, as the time singleshown as the synchronous onset of a timer. The leaver-pulling response stopped the timer, and the interval recorded was the time between the onset of the visual display and the current display. The line drawings, which averaged between 4 and 6 cm in maximum linear extent, appeared at a viewing distance of about 60 cm. They were positioned, with the center of the card at a distance that subtended a visual angle of 9°, in two circular apertures in a black vertical backboard. 

The subjects were instructed to respond as quickly as possible while keeping errors to a minimum. For each individual, age 3 only 37% of the responses were incorrect (ranging from 0.6 to 3.7 percent for individual subjects). The reaction-time data presented below include only the 96.8 percent correct responses. However, the data for the incorrect responses exhibit a similar pattern. 

In Fig. 2, the overall means of the reaction times for the two sorts of "mental rotation in two-dimensional space," then the slope of the obtained functions indicates that the average rate at which these particular objects can be thus "rotated" is roughly 60° per second. 

In the course of the repeated reaction times necessarily include any times taken by the subjects to decide how to process

Neural Pathways Associated with Hypothalamically Elicited Attack Behavior in Cats

Abstract. Small electrolyte lesions were made in cats through electrodes, which, when stimulated, elicited either quiet sitting or aggressive paw strike attack upon rats. The Neura method for impregnating degenerating axons was used to reveal that degeneration resulting from lesions in frontal attack sites followed largely along the course of the medial forebrain bundle, while the degeneration after lesions of attack sites was concentrated more heavily in the periventricular system.

Although it is now firmly established that the hypothalamus is intimately involved in the expression of aggressive behavior (1), little is known about the neural pathways through which such behavior is mediated. To attempt to trace out the circuits which may be associated with an attack on a rat we have employed neuroanatomical techniques in cats. The development of silver stains capable of selectively impregnating degenerating axons by Neura (2) has revolutionized the studies of neuroanatomists to determine the pathways of conduction and the areas of termination of brain structures and unmyelinated fiber systems. The first step in tracing out degeneration by this technique consists in destroying a small amount of neural tissue in a selected anatomic target area and then permitting the animal to survive for a