

Mental rotation and perceptual uprightness

HOWARD S. HOCK and CHERYL L. TROMLEY
Florida Atlantic University, Boca Raton, Florida 33431

Performance in Cooper and Shepard's (1973) mental rotation task was examined in the context of a model that defined the extent to which alphabet letters could be tilted from their normal orientation and still be perceptually upright. For letters with a broad range of orientations for which they remain perceptually upright, a nonlinear effect of orientation on reaction time was obtained (as in Cooper and Shepard). However, for letters with a narrow range of orientations for which they remain perceptually upright, reaction time was linearly related to orientation. The results supported the hypothesis that subjects in the Cooper and Shepard task would mentally rotate alphabet letters only when they were presented in orientations for which they were not perceptually upright.

In their classic study of mental rotation, Shepard and Metzler (1971) required subjects to make same-different responses to pairs of perspective line drawings depicting unfamiliar, three-dimensional objects. "Same" responses were required when the two objects were the same, regardless of whether they were in the same or different orientations. "Different" pairs also comprised objects in the same or different orientations, different responses being required when the pair of objects were mirror-image reversals of each other. Shepard and Metzler found that the time required for same-different judgments increased *linearly* with the angular displacement between the two objects. This was the case for rotations in the picture plane, as well as rotations in depth. More recently, Shepard and Judd (1976) obtained a similar result using an apparent rotation paradigm. They found, with the same kind of stimuli used by Shepard and Metzler, that the minimum interstimulus duration required to establish apparent motion increased linearly with the angular displacement between the depicted objects.

Cooper and Shepard (1973) have used a different experimental paradigm to study mental rotation. They presented single alphabet letters or numerals in orientations varying from 0° to 360°, and required subjects to determine whether the stimuli were normal or mirror-image reversed. Subjects' reaction times increased with the angular displacement of the stimuli from upright. Unlike Shepard and Metzler's results, however, the relationship between reaction time and orientation was *nonlinear*. The reaction time function was generally quadratic; an angular increment of 60°, which had a relatively small effect on reaction time at small disorientations

from upright, had a relatively large effect on reaction time at large disorientations from upright. Of further interest in Cooper and Shepard's results was evidence that the function relating reaction time to orientation was symmetrical with respect to the 180° orientation. This indicated that mental rotation was counterclockwise for orientations between 0° and 180° and clockwise for orientations between 180° and 360°. The symmetrical reaction time function indicated that the stimuli were rotated through the minimum angle necessary to reach upright, which implied that subjects identified the alphanumeric characters and determined their orientation prior to the onset of mental rotation (Cooper & Shepard, 1973).¹

The nonlinear quadratic relationship between reaction time and orientation in Cooper and Shepard's study provided the impetus for the present research. Cooper and Shepard (1973) have suggested that the nonlinearity may have been due to mental rotation not being required for stimuli presented at relatively small disorientations from upright. Implicit in this explanation is the idea that a familiar stimulus can be perceptually upright even though it is not in its physically upright, or normal, orientation. The research reported in this paper examined the influence of perceptual uprightness on performance in a task similar to Cooper and Shepard's. The rationale for the study was that mental rotation would be required only at orientations in which familiar stimuli were not already perceptually upright.

The model we adopted to define perceptual uprightness was closely related to Rock's (1973) model for explaining the recognition of disoriented stimuli. According to our model, the memory representation for a familiar stimulus is spatially coded by the assignment of "top" and "bottom" to the highest and lowest points of the stimulus in its normal, objectively upright orientation [Attneave & Olson (1967) and Rock & Heimer (1957) have shown that

The results of this study were reported at the annual meeting of the Psychonomic Society, Washington, D.C., 1977. Requests for reprints should be sent to Howard S. Hock, Department of Psychology, Florida Atlantic University, Boca Raton, Florida 33431. Cheryl L. Tromley is currently at Yale University.

memory representations of familiar stimuli are usually coded on the basis of their objectively upright orientation]. The perceptual representation of a stimulus is similarly coded, "top" and "bottom" being assigned to the highest and lowest points of the stimulus (whether this is based on an objective or retinal frame of reference could only be ignored in this study since subject's heads were always upright). If the "top" and "bottom" assigned to the perceptual representation for a stimulus is consistent with the "top" and "bottom" assigned to the memory representation for that stimulus, the stimulus is perceptually upright and subjects can judge whether it is normal or reversed. If, however, the spatial codes for the perceptual and memory representations are not consistent, mental rotation is required before the subject can reach a normal-reversed decision. That is, it is necessary to continuously reassign "top" and "bottom" to the perceptual representation until the spatial code for the perceptual representation is consistent with the spatial code for the memory representation.

In applying this model to alphabet letters it was observed that the shape of the letter was an important factor influencing perceptual uprightness. The letters selected for the present study were grouped into three sets on the basis of their shape (see Figure 1). The letters used were either circular (e,G), elongated (L,J), or rectangular (R,F). Their orientation varied in 36° increments from 0° to 324°.

The effect of orientation on the perceptual uprightness of the letter G (stylized to be perfectly circular) is illustrated in the top panel of Figure 2. For the G, the spatial codes for the perceptual and memory representations are inconsistent at 36°, 72°, 108°, and all other orientations. Even at the 36° orientation, which is closest to objective upright, the circular G is no longer perceptually upright and mental rotation is required. Hence, in a Cooper and Shepard mental rotation task, a linear effect of orientation on reaction time was expected for the circular letters (e,G).

SET I	SET II	SET III
G	R	J
e	F	L

Figure 1. The three sets of letters used in this study.

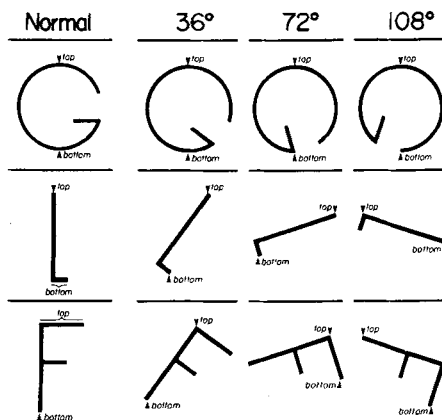


Figure 2. Illustration of spatial coding for the memory (normal orientation) and perceptual representations (at 36°, 72°, 108°) of some of the letters used in this study.

The effect of orientation on the letter L (stylized to be highly elongated) is illustrated in the center panel of Figure 2. For the L, the spatial codes for the perceptual and memory representations are consistent at 36° and 72°. It is not until 108° that the L is no longer perceptually upright and mental rotation is required. Hence, in a Cooper and Shepard mental rotation task, a pronounced nonlinear effect of orientation on reaction time was expected for the elongated letters (L,J).

The circular (e,G) and elongated letters (L,J) described above were stylized to emphasize perceptual characteristics that would minimize or maximize nonlinearity in the effect of orientation on reaction time. Alphabets, however, are usually designed so that the letters are either rectangular in shape or fit a rectangular frame. The alphanumeric characters used by Cooper and Shepard are typical in this regard, and the third set of letters (F,R) used in this experiment had the more standard rectangular shape characteristic of Cooper and Shepard's stimuli. The effect of orientation on the letter F is illustrated in the bottom panel of Figure 2. For the F, the spatial codes for the perceptual and memory representations are consistent at 36°, but not at 72° or 108°. For orientations of 72° or greater, the F is no longer perceptually upright, and mental rotation is required. In a Cooper and Shepard mental rotation task, a nonlinear effect of orientation on reaction time was expected for the F and R, but the nonlinearity was not expected to be as pronounced as that expected for the elongated letters, L and J.

To summarize, a model of perceptual uprightness was proposed to explain the nonlinearity obtained in the Cooper and Shepard mental rotation function. It was hypothesized that mental rotation would be required only when familiar stimuli were disoriented to the degree that they were no longer perceptually upright. On this basis, it was predicted that a linear

mental rotation function would be obtained for letters that have a narrow range of orientations for which they are perceptually upright (e,G), while a nonlinear mental rotation function would be obtained for letters that have a wide range of orientations for which they are perceptually upright (L,J and, to a lesser extent, F,R).

METHOD

Subjects

Nine subjects, eight male and one female, participated in three daily experimental sessions lasting about 1 h each. Seven subjects were right-handed and two were left-handed. The subjects, who were undergraduate and graduate students at Florida Atlantic University, were paid \$6 for their participation.

Design

The stimuli used in the experiment comprised the six letters presented in Figure 1. Each letter was presented equally frequently at each of 10 orientations, varying in 36° steps from 0° to 324°. At each orientation, half the presentations of each letter were normal and half were mirror-image reversed. The orthogonal combination of letters (e, G, L, J, R, F), orientations (10), and responses (normal vs. reversed) generated blocks of 120 trials. Each subject received three blocks of 120 trials in each experimental session. The order of stimuli was randomized in units of 60 trials within each block. Each subject received a different order of the three blocks for each of the three daily experimental sessions (Latin square design). Finally, subjects were divided into three groups, each of which received a unique order of blocks (Latin square design).

The stimuli were back-projected onto a translucent screen by a Carousel projector. The subjects sat at a distance of 274 cm from the screen, and the visual angle intercepted by the letters was approximately .4°. Each stimulus was presented for 2.5 sec, the exposure duration being controlled by an electronic shutter. Responses that did not occur within the 2.5-sec exposure duration were counted as errors of omission. Subjects were instructed to press one button when the stimulus was a normal letter and another button when the stimulus was a mirror-image reversed letter. Four subjects pressed the "normal" button with a finger of their preferred hands and the "reversed" button with a finger of their nonpreferred hands. For the remaining five subjects, the assignment of preferred and nonpreferred hands to the two response buttons was reversed. The subjects were instructed to respond as quickly as possible, while keeping their errors to a minimum. They were told that errors were expected. At the start of the first day of experimentation, each subject received 60 practice trials comprising stimuli from all the experimental conditions.

RESULTS

The subjects responded within the required 2.5-sec interval following stimulus presentation on all but .20% of the experimental trials. Errors of commission varied from 3.4% at 0° to 11.8% at 180°. Reaction times and commission errors at each orientation were highly correlated for both "normal" ($r = +.96$) and "reversed" ($r = +.91$) responses. The high positive correlations indicated that the differences in reaction time obtained at each orientation were not the result of differences in speed-accuracy criteria. While reaction times and errors decreased over the

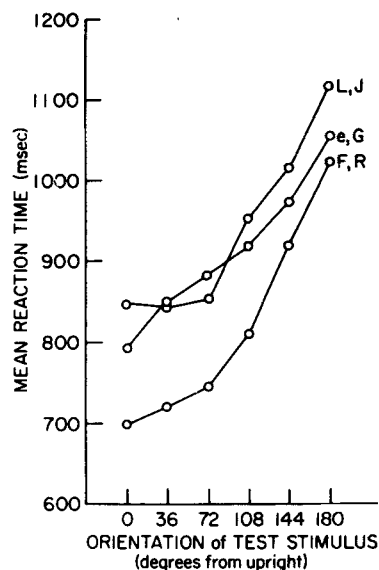


Figure 3. The effect of angular displacement from upright on mean reaction time (msec) for the three sets of letters (e,G; L,J; F,R).

3 days of the experiment, practice had no apparent effect on the slope of the rotation function. "Normal" responses were an average of 125 msec faster than "reversed" responses, and the rotation function was somewhat steeper for "normal" than for "reversed" responses. In addition, the rotation function was symmetrical with respect to the 180° orientation (as in Cooper and Shepard). As a result, the reaction times for stimuli that were equally disorientated but in opposite directions from upright could be averaged together (e.g., 36° with 324°) in the analysis that follows.

The effect of orientation on mean reaction time for each set of letters is presented in Figure 3. These data are averaged for the 3 days of the experiment, for orientations in opposite directions from upright, and for the "normal" and "reversed" responses. Analysis of variance indicated that the effect of type of letter (e,G; F,R; L,J) [$F(2,16) = 10.96$], orientation (0°, 36°, 72°, 108°, 144°, 180°) [$F(5,40) = 65.79$], and the interaction between the type of letter and orientation [$F(10,80) = 2.82$] were all significant ($p < .01$). The significant interaction was of particular importance, since it supported our prediction that the shape of the letter would influence the nature of the rotation function. Tests on trends indicated that the linear interaction was nonsignificant [$F(2,16) = 1.68, p > .05$]. That is, there was no indication of a difference in linear trend between the three types of letters. The quadratic interaction, however, was significant [$F(2,16) = 4.22, p < .05$]. That is, there was a significant difference in quadratic trend between the three types of letters. The quadratic trend was significant for the elongated (L,J) letters [$F(1,24) = 26.82, p < .01$] and rectangular (F,R)

letters [$F(1,24) = 15.93, p < .01$], but was nonsignificant for the circular (e,G) letters [$F(1,24) = 2.19, p > .05$].

The significant quadratic nonlinearity obtained for L,J and F,R, but not for e,G, supported the hypothesis that the linearity of the rotation function would depend on the range of perceptual uprightness for the letters. An examination of Figure 3 indicated that the "flattening" of the rotation functions for orientations close to upright was not due to ceiling effects. If there was a ceiling on how fast subjects could respond, it would have had a similar "flattening" effect on both L,J and e,G for which reaction times were comparable.

According to our model, the quadratic nonlinearity obtained for L,J and F,R was due to subjects' having to mentally rotate these letters only when they were presented at orientations in which they were no longer perceptually upright. As can be seen in Figure 3, the rotation functions for all three types of letters were linear (and almost parallel) between 108° and 180° , orientations outside the range of perceptual uprightness for all the letters. It also appears, from Figure 3, that when subjects did mentally rotate, they rotated the disorientated letters into perceptual upright, not into objective upright. If subjects had rotated the letters into objective upright, the linear portions of the rotation functions for L,J and F,R would have been aligned with the mean reaction time at 0° for those letters.

The final analysis performed for this experiment focused on the letter F. According to our model, the range of perceptual uprightness for this letter is asymmetrical. The "normal" F rotated counterclockwise (normal, ccw) as well as the "reversed" F rotated clockwise (reversed, cw) should have a broad range of perceptual uprightness, while the "normal" F rotated clockwise (normal, cw) and the "reversed" F rotated counterclockwise (reversed, ccw) should have a narrower range of perceptual uprightness. Evidence supporting this prediction is presented in the left panel of Figure 4. The rotation function obtained for the Fs that were expected to have a relatively broad range of perceptual uprightness (normal, ccw/reversed, cw) was flatter than the rotation function for the Fs that were expected to have a relatively narrow range of perceptual uprightness (normal, cw/reversed, ccw). The interaction assessing the difference between the two rotation functions over angular displacements of 36° , 72° , 108° , and 144° fell just short of significance at the .05 level [$F(3,24) = 2.61, p \cong .07$].

Of further interest in the normal, cw/reversed, ccw graph of Figure 4 was the difference in reaction time between 144° and 180° . While responses in all other conditions were faster for 144° than for 180° , subjects in these conditions responded slightly more

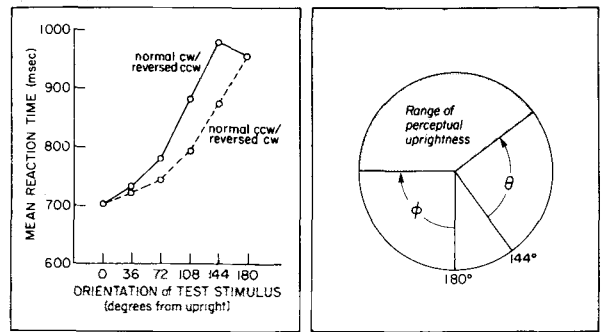


Figure 4. Left panel: The effect of angular displacement from upright on mean reaction time (msec) for the letter F. The lower graph represents the rotation function for the normal F rotated counterclockwise (ccw) combined with the reversed F rotated clockwise (cw). The upper graph represents the rotation function for the normal F rotated clockwise (cw) combined with the reversed F rotated counterclockwise (ccw). Right panel: Illustration of the predicted asymmetry in the range of perceptual uprightness for the F, and the minimum angular distance from the perceptual uprightness for the 144° and 180° orientations.

rapidly to the 180° orientation than to the 144° orientation [the difference between 144° and 180° was nonsignificant, $t(8) < 1$, but the cubic trend of the rotation function was significant, $F(1,42) = 8.45, p < .01$]. This result was consistent with the predicted asymmetry in the range of perceptual uprightness for the F, which is represented graphically in the right panel of Figure 4. In this graph, θ represents the minimum angular distance from perceptual uprightness for the 144° orientation and ϕ represents the minimum angular distance from perceptual uprightness for the 180° orientation. On the basis of these angular distances, the similar reaction times for 180° and 144° could be explained by ϕ , the minimum rotational distance for the 180° orientation, being approximately the same as θ , the minimum rotational distance for the 144° orientation.

DISCUSSION

The experimental results provided evidence for the role of perceptual uprightness in the Cooper and Shepard mental rotation paradigm. Perceptual uprightness may be of special importance in this paradigm because of the requirement that subjects discriminate between normal and mirror-image versions of alphanumeric characters. Mirror-image discrimination involves coding the perceptual representation of a stimulus in terms of left-right spatial relations (e.g., the parallel lines of the F are to the right of the line they both touch), and comparing this coded representation to the memory representation for the stimulus. Left-right coding of this kind, which depends on the formation of an internalized vertical reference axis,² can lead to accurate mirror-image

discrimination only when the stimulus is perceptually upright or the perceptual representation of the stimulus is mentally rotated into perceptual upright.

The results supported Cooper and Shepard's suggestion that the nonlinearity of the mental rotation function in their task was due to rotation not being required when the orientation of their stimuli was close to 0° (i.e., when they were perceptually upright). In the present research, stimuli predicted to have a narrow range of perceptual uprightnes (e,G) produced a linear rotation function with no significant quadratic non-linearity. In addition, stimuli predicted to have a relatively broad range of perceptual uprightnes (F,R; L,J) also produced linear rotation functions, but only at orientations outside their range of perceptual uprightnes. Thus, the Cooper and Shepard paradigm, with the effects of perceptual uprightnes controlled, meets Shepard and Metzler's (1971) requirement that linearity be obtained in order to interpret reaction time results as indicative of an underlying process of mental rotation.

REFERENCES

ATTNEAVE, F., & OLSON, R. K. Discriminability of stimuli varying in physical and retinal orientation. *Journal of Experimental Psychology*, 1967, **74**, 149-157.

COOPER, L. A., & SHEPARD, R. N. Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press, 1973.

CORBALLIS, M. C., ZBRODOFF, N. J., SHETZER, L. I., & BUTLER, P. B. Decisions about identity and orientation of rotated letters and digits. *Memory & Cognition*, 1978, **6**, 98-107.

HOCK, H. S., & HILTON, T. Spatial coding and oblique discrimination by children. *Journal of Experimental Child Psychology*, in press.

ROCK, I. *Orientation and form*. New York: Academic Press, 1973.

ROCK, I., & HEIMER, W. The effect of retinal and phenomenal orientation on the perception of form. *American Journal of Psychology*, 1957, **70**, 493-511.

SHEPARD, R. N., & JUDD, S. A. Perceptual illusion of rotation of three-dimensional objects. *Science*, 1976, **191**, 952-954.

SHEPARD, R. N., & METZLER, J. Mental rotation of three-dimensional objects. *Science*, 1971, **171**, 701-703.

NOTES

1. Corballis, Zbrodoff, Shetzer, and Butler (1978) have substantiated this conclusion in experiments involving identity and orientation decisions for disoriented alphanumeric characters.
2. Hock and Hilton (in press) have shown that children's difficulty in discriminating between mirror-image obliques was due, in part, to their difficulty in establishing an internalized vertical reference axis for spatially coding the obliques.

(Received for publication June 20, 1978;
accepted September 29, 1978.)