

Using the fold point to analyze mental rotation data: A second look

DIANE L. DAMOS

University of Southern California, Los Angeles, California

This experiment was designed to examine some of the characteristics of the estimated fold point in a mental rotation task. Thirty males performed a mental rotation task for approximately 1 h on each of 5 consecutive days. Each day was divided into an early and late session. Fold points were estimated for each session for both standard and mirror responses on a subject-by-subject basis. Standard and mirror responses had significantly different estimated fold points, which were unaffected by practice. The test-retest reliability for successive sessions was moderate for both standard and mirror responses and also appeared unaffected by practice.

In one variation of the mental rotation paradigm developed by Cooper and Shepard in 1973, the subject must decide whether the stimulus being displayed is the standard version of a previously learned figure or its mirror image. For two-dimensional rotations (rotations in the picture plane), the results of numerous experiments demonstrate that subjects mentally rotate the displayed stimulus to the upright position before making a standard/mirror decision. This rotation appears to be analogous to the way in which a physical object would be rotated. Thus, the relation between the reaction time and the absolute degrees of rotation can be expressed as:

$$RT = a + bD, \quad (1)$$

where RT is the reaction time, D is the absolute degrees of rotation of the stimulus from upright, and a and b are the parameters of the best-fitting linear equation.

Although Equation 1 provides a good general description of the data obtained from the mental rotation paradigm, several investigators (e.g., Cooper & Shepard, 1973; Hock & Tromley, 1978; Koriat & Norman, 1985) have noted consistent discrepancies between the predicted and the obtained reaction times. Generally, the discrepancy is a shorter-than-predicted reaction time when the rotation is less than approximately 60° . Hock and Tromley (1978), among others, have argued that this discrepancy is caused by familiarity with the stimuli; typically, the discrepancy occurs when the stimulus is either a letter or a number. Their data indicate that familiar stimuli have a range of orientations that are perceptually upright and, consequently, that they do not require mental rotation to true upright. The faster-than-predicted reaction times occur because the stimuli, although objectively tilted, are within the perceptually upright range.

The idea of a perceptually upright range cannot, however, be applied to abstract stimuli. Steiger and Yuille

(1983) noted that the relation between reaction time and the degrees of rotation given in Equation 1 assumes that the subject remembers the stimulus as being exactly in the upright position. If this assumption is not correct, the number of degrees the subject must rotate the stimulus will differ from the true number of degrees to be rotated by the difference between the subject's 0° (vertical) point and true vertical.

Steiger and Yuille (1983) referred to the subject's 0° point as the *fold point*, which was expressed in degrees from true vertical. These investigators developed a technique for identifying the fold point mathematically, which they used to analyze data from 8 subjects who rotated abstract shapes in the picture plane. The subjects performed 288 trials on each of 2 consecutive days. Each subject's estimated fold point appeared remarkably stable; the between-sessions correlation was 0.88. Large individual differences in the estimations of the fold point were also apparent, ranging from 0° to 95° . As anticipated, the average percentage of variance (R^2) accounted for by Equation 1 increased from 0.782 to 0.913 when the data were adjusted for the fold point.

The concept of a fold point could be very useful for interpreting mental rotation data when the stimulus is an abstract shape. As Steiger and Yuille (1983) demonstrated, using a fold point can reduce the amount of "unexplained" variance. This reduction is accomplished essentially by correcting for a constant error. In addition, differences in the location of the fold point may explain some of the individual differences observed in both the rate of mental rotation and the shape of the function relating reaction time to degrees of rotation.

Because the concept of a fold point is relatively new, its characteristics must be explored more extensively before it can be routinely adopted. One of the most important characteristics concerns its test-retest reliability. Although Steiger and Yuille (1983) found a high test-retest reliability, their subjects received relatively little practice on the task. Examining the reliability of the fold point when the subjects receive more practice is im-

Correspondence may be addressed to Diane Damos, Department of Human Factors, Institute of Safety and Systems Management, University of Southern California, Los Angeles, CA 90089-0021.

perative. Another important characteristic concerns the relation between the fold points for the standard and mirror versions of the stimulus. Presumably, the fold points for the standard and mirror versions of a stimulus should be identical. Because Steiger and Yuille did not calculate the fold points for the standard and mirror response trials separately, no information concerning this relation is available.

The data presented below are from one of a series of experiments designed to examine the effect of extensive practice on mental rotation skills and the generality of these skills. The data will be analyzed to determine the test-retest reliability of the estimated fold point when the subjects receive extensive practice. In addition, the relation between the estimated fold points for the mirror and standard versions of an abstract shape will be determined.

A brief description of the experimental procedure is given below. Interested readers should contact the author for more details.

METHOD

Subjects

Thirty right-handed males between 18 and 35 completed the experiment. All of the subjects were native English speakers, and none had any flight training. The subjects were paid \$5.50/h for participating, and they were informed that they could win performance bonuses, as described below.

Apparatus

A DEC PDP-11 microcomputer generated all stimuli, recorded and processed the subjects' responses, and timed all trials. The stimuli were displayed on a Tektronix 4125 color graphics terminal. The subjects responded by pressing keys on two identical 4 × 4 matrix keypads. Each subject sat approximately 110 cm from the display screen.

Task

The task required each subject to distinguish between a shape and its mirror image as quickly as possible. If the standard shape was presented, the subject pressed a key under his right index finger. If the mirror image was presented, he pressed a key under his left index finger. The stimuli could be presented at one of six orientations: 0° (upright), 60°, 120°, 180°, 240°, and 300° of clockwise rotation. Three different stimuli were used: uppercase F, uppercase G, and a 24-point abstract shape (Stimulus 29; see Vanderplas and Garvin, 1959). The stimulus presentation was unpaired.

Design

The subjects were randomly assigned to one of three groups. Group 1, an experimental group, received training using the letter G and transferred to the 24-point figure. Group 2, also an experimental group, trained using the letter F and transferred to the 24-point figure. Group 3, a control group, practiced only with the 24-point figure.

Procedure

All testing began on a Monday. The subjects in Groups 1 and 2 completed one session on each of 10 consecutive week days. The subjects in Group 3 completed one session on each of 5 consecutive week days. Thirty-six trials were administered in each session. Each trial consisted of 5 presentations of the standard shape and 5 of the mirror image at each of the six orientations for a total of 60 stimulus presentations/trial. The order of presentation of the stimuli was random within each trial. All subjects saw the same order of stimuli.

At the beginning of each week, the subjects were shown the stimuli to be used during that week. Both the mirror and the standard versions of the stimulus were displayed. The subjects were allowed to view the

stimuli as long as they desired. At the beginning of the second day, the subjects were asked if they remembered the stimulus. If they did, they began the experiment immediately. If they did not, they were allowed to view the standard and mirror versions of the stimulus again.

The percentage of correct responses and the average correct reaction time (RT) was presented as feedback to the subject at the end of each trial. The subject received a \$0.25 bonus for each trial in which he exceeded his own previous best performance. All instructions were taped and immediately preceded the relevant condition. After the subjects had completed the experiment, they were asked to draw the abstract shape from memory.

RESULTS

All of the data discussed below are from the second week of the experiment, when the subjects were viewing the abstract shape. For purposes of analysis, each day was divided into two sessions, each of which comprised 18 trials.

Error Rates

The error rates of Groups 1, 2, and 3 for standard responses were 5.1%, 5.6%, and 4.9%, respectively. The corresponding values for mirror responses were 5.2%, 4.5%, and 4.3%.

Fold Point Calculation

Two methods were used to estimate each subject's fold point. One method involved a statistical approach that consisted of a mathematical computation of the fold point. The second involved subjective estimations based on drawings the subject produced at the end of the experiment. Each method is described below.

Statistical computation. On the basis of the description given in Steiger and Yuille (1983), a computer program was developed to estimate the fold point for each subject during each session. Thus, two estimates of the fold point for standard responses and two estimates for mirror responses were obtained for each subject on each day. All estimates were based only on correct responses.

Basically, the program followed an iterative approach to identify the fold point from a given session. The program calculated the best-fitting linear equation for each of 500 fold points. These fold points occurred every 0.36°, beginning at 0.0°. The fold point that resulted in the lowest standard deviation of the residuals was identified as the subject's fold point for that session.

Drawings. An attempt was made to identify each subject's fold point during the last session from the drawing he made at the end of the experiment. Two judges independently estimated the fold point for each subject. Although the judges' estimates correlated significantly ($r = .66, p < .01$), both of the judges' estimates correlated poorly with the computed estimate ($r = .38$ and $r = .14, p > .05$ for both).

R^2

The method used to select the fold point ensured that the percentage of variance accounted for by Equation 1

Table 1
Median R^2 for the Best-Fitting Equations,
Based on Raw Data versus Data Corrected for the Fold Point

Session	Group					
	1		2		3	
	Standard	Mirror	Standard	Mirror	Standard	Mirror
Raw Data						
1	.965	.890	.915	.690	.950	.890
5	.955	.885	.830	.850	.955	.895
10	.935	.815	.840	.745	.925	.840
Corrected Data						
1	.975	.970	.970	.980	.985	.960
5	.990	.980	.960	.930	.985	.975
10	.965	.990	.920	.935	.980	.980

Table 2
Correlations between Successive Sessions

Session	Group		
	1	2	3
Standard Responses			
1 with 2	-.001	.353	.625
2 with 3	-.019	.057	.804*
3 with 4	.846*	.644*	.968*
4 with 5	.420	-.077	.525
5 with 6	-.063	.722*	.060
6 with 7	-.399	.318	.133
7 with 8	.054	.308	.706*
8 with 9	.450	.892*	.642*
9 with 10	.624	-.164	.361
Mirror Responses			
1 with 2	.133	-.126	.648*
2 with 3	.533	.199	.331
3 with 4	.628	.891*	-.177
4 with 5	-.052	.508	.113
5 with 6	.299	.700*	.209
6 with 7	-.040	.266	.607
7 with 8	.639*	.450	.572
8 with 9	.453	.943*	.819*
9 with 10	-.073	.780*	.501

* $p < .05$.

when the data were corrected for the fold point was at least equal to that obtained when Equation 1 was calculated on the raw data. Space limitations prohibit an exhaustive comparison of R^2 for the raw data versus the data corrected for the fold point. Nevertheless, the median R^2 is provided in Table 1 for Sessions 1, 5, and 10, for both standard and mirror responses.

Test-Retest Reliabilities

Table 2 shows the correlation between the estimated fold points for successive sessions, for standard and for mirror responses by group. Table 3 shows the correlation between the estimated fold point for standard responses and the estimate for mirror responses obtained during the same session.

Because the correlations in both tables were relatively low, a three-way (group \times trial block \times standard/mirror response) analysis of variance (ANOVA) was conducted to determine whether the fold points were affected by

practice. The main effect of response type was significant [$F(1,27) = 13.14, p = .0012$]. The average estimated fold point for standard responses was 55.67° ; the average for mirror responses was 78.22° . The group \times session interaction was also significant [$F(18,243) = 1.84, p = .0218$]. This interaction appears to be caused by differences in trends among the three groups; the estimated fold points for Groups 2 and 3 decrease with practice, whereas those for Group 1 increase. No other main effects or interactions were significant.

DISCUSSION

Generally, the results of this experiment concur with those of Steiger and Yuille (1983) although the benefits obtained from using a fold point are not as large. As Steiger and Yuille reported, using a fold point to calculate Equation 1 did increase R^2 (see Table 1). These increases are of approximately the same magnitude as those reported by Steiger and Yuille.

The magnitude of the test-retest reliabilities given in Table 2 is considerably less than the 0.88 reported by Steiger and Yuille (1983). Nevertheless, 15 of the correlations in Table 2 are significantly different from 0.00, whereas only 2.7 would be expected by chance. Practice apparently had no effect on the magnitude of these correlations; no trends are evident.

Steiger and Yuille (1983) did not report the correlation between the estimated fold points for standard and mirror responses emitted during the same session. Consequently, the results of this experiment cannot be compared to any existing data. By chance, 1.5 of the correlations reported in Table 3 should have been significant; 2 were significant. Apparently, then, for any given session, the estimated fold point for standard responses had no relation to the estimate for mirror responses. This finding is substantiated by the results of the ANOVA, which revealed a significant difference between the estimated fold points for standard and mirror responses.

This lack of relation between the estimated fold points may have been caused by the experimental procedures. Most subjects saw the drawings of the mirror and standard versions of the abstract shape only once before their first trial using the shape. The subjects probably encoded two different images, resulting in two different "uprights."

The estimated fold points for both standard and mirror responses did not change significantly with practice. This lack of change again may be caused by the training procedures. Because most of the subjects saw the drawings only once, any discrepancies between the encoded orientation of the abstract shape and its true orientation would not be detected during the experiment.

In summary, the results of this experiment indicate that the concept of a fold point may be helpful in interpreting data from a mental rota-

Table 3
Correlation between the Estimated Fold Points for Standard
and Mirror Responses during Each Session

Session	Group		
	1	2	3
1	-.132	-.308	.233
2	.032	-.174	.303
3	.362	.505	-.026
4	-.335	.459	-.008
5	.367	.242	.137
6	.239	.366	.698*
7	.654*	.360	.480
8	.345	.240	-.017
9	.578	.275	.448
10	.152	-.143	.202

* $p < .05$.

tion paradigm by decreasing the amount of "unexplained" variance. Although the test-retest reliabilities are somewhat low, they may be sufficient for many types of laboratory and applied research. The lack of relation between the estimated fold points for standard and mirror responses and the insensitivity of the fold point to extended practice may be caused by the experimental procedures used in this study. The effect of such procedures on the characteristics of the fold point should be investigated systematically.

REFERENCES

- 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.
- HOCK, H. S., & TROMLEY, C. L. (1978). Mental rotation and perceptual uprightiness. *Perception & Psychophysics*, *24*, 529-533.
- KORIAT, A., & NORMAN, J. (1985). Mental rotation and visual familiarity. *Perception & Psychophysics*, *37*, 429-439.
- STEIGER, J. H., & YUILLE, J. G. (1983). Long-term memory and mental rotation. *Canadian Journal of Psychology*, *37*, 367-389.
- VANDERPLAS, J. M., & GARVIN, E. A. (1959). The association value of random shapes. *Journal of Experimental Psychology*, *57*, 147-154.

(Manuscript received May 29, 1989.)