

## The effect of speed-accuracy tradeoff on sex differences in mental rotation

DAVID F. LOHMAN

*The University of Iowa, Iowa City, Iowa*

This experiment examined the effects of sex differences in the form of speed-accuracy curves on sex differences in rate of mental rotation. Eighty-nine subjects attempted 1,200 rotation problems similar to those used by Shepard and Metzler (1971). Stimulus exposure was varied systematically over a wide range, and response accuracy was determined at each exposure. Speed-accuracy curves were then fit using an exponential function similar to one proposed by Wickelgren (1977). Results showed that apparent differences between males and females in rate of rotation are explained by sex differences in the shape of the speed-accuracy curves, with females reaching asymptote sooner on trials requiring more rotation. Similar effects were obtained in a comparison of subjects high and low in spatial ability.

There is controversy in the literature on sex differences in spatial abilities. Some have argued that the male advantage is large and consistent (Harris, 1978; Maccoby & Jacklin, 1974; McGee, 1979), whereas others have claimed that the difference is usually small and is often not observed (Caplan, MacPherson, & Tobin, 1985; Hyde, 1981; Sherman, 1978).

Tasks that require subjects to determine whether one stimulus can be rotated into congruence with another stimulus are among those that frequently show a large sex difference favoring males (Linn & Petersen, 1985). More importantly, several studies now suggest that the locus of sex differences in this task resides in the speed with which subjects mentally rotate stimuli, with males on average rotating stimuli faster than females (Kail, Carter, & Pellegrino, 1979; Tapley & Bryden, 1977).

In the standard mental rotation task, the subject is shown two stimuli that differ in orientation. The subject must determine whether the stimuli can be rotated into congruence or whether one stimulus is a mirror image of the other. Time to verify stimulus identity is ordinarily found to be a linear function of the angular separation between figures. According to the standard Shepard and Metzler (1971) model, the slope of this line estimates rate of rotation, whereas the intercept estimates time for encoding, comparison, and response processes. More complex models are sometimes proposed (e.g., Just & Carpenter, 1976; Yuille & Steiger, 1982), especially for the less well-behaved negative trials (e.g., Kail, Stevenson, & Black, 1984).

Metzler (1973) noted that both slopes and intercepts were larger for women than for men. However, Metzler's small sample sizes ( $N = 4$  women in each of two experiments) precluded statistical tests of the significance of the difference. Furthermore, the variance of the distribution response latencies for each subject was highly correlated with rate of rotation in one study ( $r = .78$ ), with the slowest rotators showing the most variability in response latency.

Tapley and Bryden (1977) administered 200 trials of the same stimuli used by Shepard and Metzler (1971) to 20 male and 20 female undergraduate volunteers. Regressions of response latencies on angle of rotation showed a significant sex difference in slopes, with females having the steeper slope. Intercepts were higher for females, but not significantly so. Furthermore, men committed fewer errors than women on 40° and 80° *same* trials, whereas women committed fewer errors than men on the 120° and 160° *same* trials. This suggests that some part of the sex difference in rate of rotation may be due to sex differences in speed-accuracy tradeoff.

Kail et al. (1979) also found a significant sex difference in the slope parameter. They administered 144 rotation trials to 51 male and 55 female undergraduates. Half of the trials used alphanumeric stimuli; half used less familiar letter-like stimuli from the Primary Mental Abilities (PMA) Space test (Thurstone, 1958). Slopes for *same* trials were significantly steeper for women on both types of problems, but particularly on the more difficult PMA stimuli. Intercepts showed no significant sex difference. Overall, women committed 1% fewer errors than men; however, the difference was not significant. Kail et al. (1979) also computed slopes and intercepts for each subject. The distribution of these slopes was more variable for women than for men. In fact, the distribution of slopes for females on the PMA stimuli was bimodal, with 30% of the females having slopes outside the range of the male distribution.

This research was conducted under Grant No. 84R1P80 from the Research Initiation Program of the Air Force Office of Scientific Research and a grant from the Iowa Measurement Research Foundation. I am indebted to Paul Nichols for collecting the data, to H. D. Hoover and Steve Dunbar for assistance in deriving the model for the speed-accuracy curves, and to Jim Pellegrino for copies of the slides used as stimuli. Requests for reprints should be sent to D. Lohman, 366 Lindquist Center, The University of Iowa, Iowa City, IA 52242.

In a similar study, Kail et al. (1984) administered 100 rotation problems using PMA stimuli to 36 male and 32 female undergraduates. Once again, females showed a steeper slope than males. No sex differences were noted in solution strategy.

However, not all rotation tasks show significant sex differences. Berg, Hertzog, and Hunt (1982) administered 480 rotation problems on each of 4 days to 66 subjects. The subjects were divided into four age groups, with average ages of 20.9, 32.4, 50.9 and 63.3 years. The stimuli were 12 letter-like forms developed using Palmer's (1977) methods. Sex differences in rotation slope were significant only for the oldest subjects. Females in this group were significantly slower than males, especially on Day 1. There were no significant age or sex differences in error rate. Berg et al. (1982) suggested that consistent sex differences may appear only when the task is more difficult.

Other evidence suggests that sex differences in rate of rotation may be reduced or eliminated with practice. Alderton, Pellegrino, and Lydiatt (1984) administered 192 rotation problems, using either alphanumeric or PMA characters, to 76 subjects on each of four sessions. Subjects were selected on the basis of extreme scores on pretests of spatial ability. Alderton et al. found that sex differences in rate of rotation were significant only on the first session. Regian and Pellegrino (1984), using asymmetric polygons as stimuli, found a similar effect. Thirty-one females and 29 males matched for pretest scores on the Card Rotations Test (Ekstrom, French, & Harman, 1976) solved 280 rotation problems on each of seven sessions. Sex differences in slope were significant only for the first two sessions.

In summary, males and females often differ in their performance on rotation tasks. On paper-and-pencil versions of the sort of three-dimensional figures used by Shepard and Metzler (1971), differences are usually on the order of one standard deviation in number of problems correctly solved in a fixed time (Linn & Petersen, 1985). In experiments in which trials are self-paced, sex differences are usually confined to the slope of the regression of response latency on angular separation between figures. These differences are generally larger when stimuli are complex and unfamiliar, and are often attenuated or eliminated with practice, especially on simpler figures. Women also sometimes show slightly greater intercepts and commit fewer errors overall than men; however, differences are usually small and not statistically significant.

Unambiguous comparisons of response latencies for different subjects or groups of subjects presume that the subjects or groups being compared have been equated for speed-accuracy tradeoff. The hypothesis that a significant difference in latencies between two groups of subjects is in part attributable to group differences in speed-accuracy tradeoff is often tested by comparing raw or adjusted error rates (e.g.,  $d'$ ) for the groups. If groups do not differ significantly in error rate, it is concluded that differences in latencies are genuine and cannot be attributed to differences in speed-accuracy tradeoff. This argument is flawed

in three respects. First, error rates for different trial types may differ even though overall error rates are equivalent. For example, error rates for 150° rotation problems may differ between groups even when overall error rate is equivalent. Second, even small, statistically nonsignificant differences in error rate can produce large differences in response latencies, especially when subjects are operating near asymptote on the speed-accuracy curve (Pachella, 1974; Wickelgren, 1977). Third, even if error rates are identical, groups will be equated on speed-accuracy tradeoff only if asymptotes of the speed-accuracy curves are identical for the two groups. If asymptotes differ, then groups may be at quite different levels of speed-accuracy tradeoff even when error rates are equivalent. This is most likely to be the case when the experimental task is complex and groups differ in average error rate on relatively unspeeded tasks similar to the experimental task. For example, comparison of latencies for subjects high and low in spatial ability on a complex mental rotation task may be suspect even when error rates are identical, since, for any error rate, low-spatial subjects will be performing nearer asymptote than will high-spatial subjects.

The purpose of the present study was to determine whether sex differences in rate of rotation may be attributable, in part, to sex differences in speed-accuracy tradeoff adopted or to sex differences in the shapes of speed-accuracy curves for rotation problems of increasing complexity. Speed-accuracy curves were estimated by presenting pairs of stimuli similar to those used by Shepard and Metzler (1971) for various fixed exposures and then determining probability of a correct response at each exposure. In addition, because some theorists (e.g., Furneaux, 1961; Pachella, 1974; Thurstone, 1937) have hypothesized that motivation also exercises a substantial influence on response latencies, motivation was manipulated by offering subjects a monetary bonus for correct responses on some trials, but not on others.

## METHOD

### Subjects

The subjects were 89 paid volunteers recruited through newspaper advertisements. Their ages ranged from 16 to 20 years. All subjects were required to attend four sessions that lasted approximately 2 h each. During the first hour of the first session and the last hour of the last session, all subjects were administered four spatial tests: Cards, Figures, Paper Folding, and Form Board (Ekstrom et al., 1976). The Cards and Figures tests estimated the spatial factor called spatial relations (SR), and the Paper Folding and Form Board tests estimated the factor called visualization (Vz). These are the two most frequently reported spatial factors in the literature (Lohman, 1979). During the remaining time, subjects attempted 1,200 mental rotation problems similar to those used by Shepard and Metzler (1971). Six subjects failed to complete all sessions or all tests. The final sample consisted of 30 males and 53 females.

### Stimuli

The stimuli were 120 slides that contained pairs of figures similar to those used in the original Shepard-Metzler (1971) study. Stimuli differed in the configuration of blocks used to construct the stimulus pair and in the orientation of the arms. Examples of the

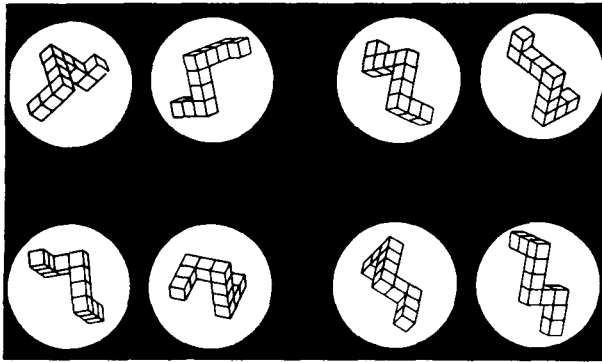


Figure 1. Examples of the four types of stimulus figures used in the experiment. All stimuli are constructed from 10 cubes organized into four arms, with three right angles. Stimuli differ in the number of elements in each arm (top row vs. bottom row) and in the orientation of the arms (first column vs. second column). In this example, all require 60° rotations about the *x* axis.

four types of stimuli are shown in Figure 1. Stimuli were rotated about either the *x*, *y*, or *z* axis in increments of 30°. The full set of 120 stimuli was defined by the crossing of stimulus type (1, 2, 3, or 4) × angular separation (30°, 60°, 90°, 120°, or 150°) × axis of rotation (*x*, *y*, or *z* axis) by trial type (positive or negative). Each pair of stimuli was presented 10 times.

### Procedure

The subjects were randomly divided into two groups. Subjects in Group 1 ( $n = 45$ ) saw each stimulus for five different exposures, once in a high-incentive condition and once in a low-incentive condition. Incentive was manipulated by paying subjects a one-cent bonus for each correct answer in blocks of 40 trials designated "high incentive." Blocks of high- and low-incentive trials were alternated throughout the experiment. Blocks also varied in the axis about which stimuli were rotated. Since we were not interested in estimating the effect of this factor, all stimuli in a block of 40 trials were rotated about the same axis. Furthermore, the subjects were told at the beginning of each block the incentive value and axis of rotation for all trials in the block.

The subjects in Group 2 ( $n = 38$ ) saw each stimulus for 10 different exposures; 5 exposures were the same as those used for Group 1. The additional 5 exposures were generally much shorter, and were included to ensure that even the most able subjects would be confronted with trials they could not easily solve, even after hundreds of trials. Blocks of 40 normally paced and 40 rapidly paced trials were alternated throughout the experiment. Equations used to determine exposure are given in the first column of Table 1. Since negative trials generally take longer than positive trials, all negative trials were exposed for an additional 100 msec. Subjects responded to each trial on a 5-point scale with values of 1 (definitely no), 2 (probably no), 3 (uncertain), 4 (probably yes), and 5 (definitely yes).

## RESULTS AND DISCUSSION

Our primary goal in these analyses was to estimate five speed-accuracy curves for each of the five values of angular separation between figures, separately for males and females, and then to use these curves to estimate rate of rotation with accuracy held constant. Dependable estimation of each curve requires many observations. In order to avoid problems that could result from presenting the

same set of stimuli on all trials, stimuli were varied along two dimensions—axis of rotation and configuration of cubes in each stimulus—that would make them nonidentical, yet not radically different. Depth rotations are sometimes found to be more difficult than picture-plane rotations, and are sometimes found to be of equivalent difficulty (see Shepard & Cooper, 1982). We attempted to minimize effects due to axis of rotation by presenting trials in blocks of 40 in which axis of rotation was the same for all trials, and by forewarning subjects of the axis of rotation. Responses were then averaged over these facets, reducing the data set from 1,200 to 100 observations for each subject.

A simple analysis of errors can be misleading because subjects may be biased in the frequency with which they say "same" or "different." One way to correct for this bias is to compute  $P(A)$ , a nonparametric measure of sensitivity often used in signal detection analyses. A simple approximation to  $P(A)$  is given by  $P(C)$  (McNichol, 1972).  $P(C)$  is the probability of a subject's responding "definitely yes" to positive trials and "definitely no" to negative trials, with partial credit given for intermediate scale values. It is a more sensitive measure of response accuracy than the usual binary error score. We have used  $P(C)$  in the analyses we report here because it preserves the original score metric and thus is easily compared with  $P(S)$ , the probability of a correct "same" response.<sup>1</sup>

Although  $P(A)$  and  $P(C)$  adjust scores for response bias, the correction is based on the assumption that positive and negative trials are psychologically equivalent. For exam-

Table 1  
Mean Accuracy Scores, by Level of Stimulus Exposure and Angular Separation Between Stimuli, Separately for Males and Females

Exposure*	Angular Separation				
	30°	60°	90°	120°	150°
Males					
5R+1	.811	.770	.728	.711	.673
9R+1	.796	.780	.754	.725	.710
12R+1	.783	.779	.772	.756	.739
18R+1	.854	.791	.815	.787	.763
35R+1	.856	.882	.862	.860	.811
11R+40	.837	.811	.784	.761	.762
18R+100	.857	.850	.802	.802	.777
29R+130	.876	.847	.844	.839	.800
40R+500	.870	.895	.830	.826	.832
51R+1400	.889	.902	.881	.848	.824
Females					
5R+1	.803	.734	.707	.684	.644
9R+1	.805	.778	.735	.695	.680
12R+1	.827	.772	.755	.751	.683
18R+1	.852	.803	.785	.745	.741
35R+1	.864	.854	.833	.801	.752
11R+40	.801	.781	.746	.729	.694
18R+100	.826	.814	.777	.767	.747
29R+130	.859	.829	.808	.808	.752
40R+500	.858	.858	.821	.802	.792
51R+1400	.874	.861	.831	.827	.792

\*Stimulus exposure in milliseconds, where R is the angular separation between figures. The first five equations in each set were used to generate rapidly paced trials, and the second five equations were used to generate normally paced trials.

ple, if it is sometimes easier to detect a foil than to confirm that stimuli are the same, then positive and negative trials should not be combined as they are in  $P(A)$  or  $P(C)$ . In fact, regression of response latencies on angle of separation between figures is often shallower and less linear for negative than for positive trials, probably because negative trials may be solved differently from positive trials (Kail et al., 1984; Pellegrino & Kail, 1982). In the present experiment, 22% of the subjects reported solving problems by rotating one stimulus only if an initial comparison did not indicate that the stimuli were mirror images. Therefore, in addition to the analyses on  $P(A)$  and  $P(C)$ , we have also conducted analyses on the probability of a correct "same" response, or  $P(S)$ . In general, effects were larger and speed-accuracy curves more divergent for different levels of rotation in the analysis of  $P(S)$ , but they were more likely to be statistically significant for  $P(C)$  than for  $P(S)$ , since  $P(C)$  is based on twice as many observations as  $P(S)$ . However, all three analyses yielded the same conclusions.

### Preliminary ANOVAs

Our attempt to manipulate incentive by giving bonuses for correct responses on half of the trials for subjects in Group 1 showed no important effects. A mixed-model ANOVA on  $P(C)$  with sex as the between-subjects factor and incentive (yes, no), type of trial (positive, negative), angle of rotation (30°, 60°, 90°, 120°, or 150°), and stimulus exposure (five levels) as within-subjects factors showed no effect for incentive condition [ $F(1,43) = 2.37, p > .10$ ]. Either our bonuses and instructions did not motivate subjects or, more likely, motivation exerts less influence when trials are experimenter-paced than when they are self-paced. Therefore, we averaged over incentive condition for subjects in Group 1.

For subjects in Group 2, five exposure durations were the same as the exposures used for Group 1. The two groups did not differ on these trials [ $F(1,81) = 1.18, p > .05$ ], and so data were combined. Thus, the basic data for fitting speed-accuracy curves consisted of average scores at each of 10 different exposure durations. Means for the five shortest exposures came from the subjects in Group 2. Means for the five longest exposures came from subjects in Group 1 and Group 2.

### Speed-Accuracy Curves

**Sex differences.** Speed-accuracy curves were determined by fitting exponential functions of the form  $P(C) = \gamma + \alpha(1 - e^{-\beta t})$ , where  $P(C)$  is average acceptance rating (see Table 1),  $\gamma$  is the intercept of the curve at  $t = 0$ ,  $\alpha$  is the scale factor,  $\beta$  is the rate of increase in accuracy with increases in stimulus exposure, and  $t$  is the exposure latency of the stimulus.  $\gamma$  was fixed at .5 in these models, and so the asymptote of each curve is simply  $\alpha + .5$ . This model is similar to the one proposed by Wickelgren (1977). Parameters were estimated by the method of least squares, using the Marquardt method. Separate analyses were performed for males and females at each

of the five levels of angular separation between figures. Each curve was based on the 10 means reported in Table 1. The data were quite regular, and thus were well described by these curves. Median R-squares were .78 for males and .82 for females. As might be expected, R-squares were smallest for the 30° rotation trials, due to restriction of range on exposure latencies, and increased systematically with increases in angular separation and the range of exposure latencies. Root-mean-square errors were quite small, averaging .026 points for males and .021 points for females on the  $P(C)$  scale. These speed-accuracy curves are plotted in Figure 2. The asymptote ( $\alpha + .5$ ) and curvature ( $\beta$ ) parameters for these curves are reported in Table 2. There were no significant sex differences in the  $\beta$  parameters. Therefore, curves for males and females do not differ in the rate at which they reach asymptote. However, the  $\alpha$  parameter showed a significant difference favoring males in 90°, 120°, and 180° trials. In each case, the curve for females reached asymptote sooner. Furthermore, the magnitude of the difference in asymptote was systematically related to the amount of rotation required, with 30° trials showing the smallest difference and 150° trials showing the largest difference. These differences in asymptote on larger rotation trials translate into even larger differences in response latencies as accuracy approaches asymptote in the lower-scoring group.

The speed-accuracy curves in Figure 2 can be used to estimate rate of rotation at different levels of accuracy. Rotation latencies for a fixed level of accuracy can be estimated graphically by drawing a horizontal line through the speed-accuracy curves at the chosen accuracy level and then determining the latency that corresponds to each point of intersection. Alternatively, the equation can be reversed, and latency estimated from accuracy. Latencies for different levels of accuracy can then be plotted against angular separation between figures.

Figure 3 shows a family of slopes constructed in this manner for four levels of accuracy for males and females. Solid lines represent data for males; dashed lines, data for females. For a given level of accuracy, females were slightly slower than males. However, slopes were virtually identical over most of the plot, indicating that rate of rotation is the same for males and females. Latencies diverged only at the highest level of rotation, as accuracy for females approached asymptote. For example, as shown in Figure 3, latencies for 150° trials were significantly elevated for females at an accuracy score of .74. On average, females had great difficulty solving these trials, perhaps making several attempts or double-checking their answers to achieve this level of accuracy.

At higher levels of accuracy, say .82, some curves do not even intersect the line for fixed accuracy (e.g., females on 90° to 150° trials, and males at 150° trials). On average, females were unable to achieve this level of accuracy on the 150° rotation trials, even when allowed 9,050 msec. Such extreme latencies rarely occur when trials are self-paced, rather than experimenter-paced as

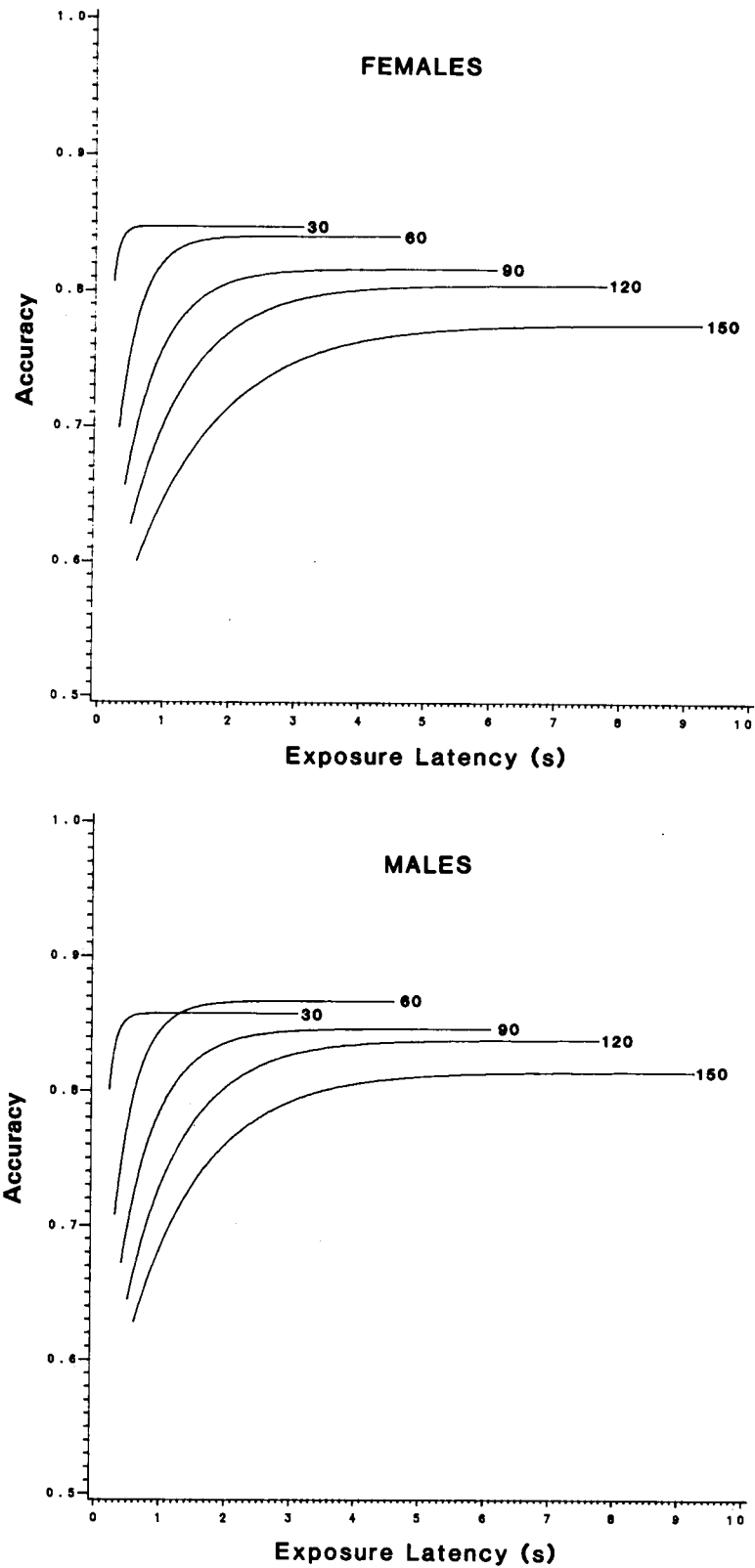


Figure 2. Speed-accuracy curves for females (top panel) and males (bottom panel) at each of five levels of angular separation between stimuli. Numbers on each curve represent the angular separation between figures. The length of each curve indicates the range of exposure latencies for trials requiring a given amount of rotation.

**Table 2**  
**Sex Differences in the Asymptote and Curvature Parameters**  
**for Speed-Accuracy Curves at Each Level of Rotation**

Rotation	Parameter					
	Asymptote ( $\alpha + .5$ )			Curvature ( $\beta$ )		
	Males	Females	$t(18)$	Males	Females	$t(18)$
30°	.857	.847	.67	9.26	10.77	-.47
60°	.866	.839	1.33	2.80	2.93	-.19
90°	.846	.816	2.03*	1.72	1.71	.04
120°	.839	.805	2.14*	1.12	1.08	.21
150°	.815	.777	3.03†	.87	.74	1.06

\* $p < .05$ . † $p < .01$ .

in this study. However, when trials are self-paced, subjects impose a deadline and respond, but at the price of an increased error rate (Pellegrino & Kail, 1982). Error rate typically increases with angle of rotation, especially when stimuli are complex (e.g., Lansman, 1981).

Note in Figure 3 that the line showing rate of rotation for males at an accuracy of .74 parallels the line for females at an accuracy of .70. In each case, the form of

the line for males at one level of accuracy is anticipated by the line for females at the previous level of accuracy. Asymptotes of the speed-accuracy curves for males and females also differed by .04 points on the 120° and 150° trials. Thus, sex differences in rate of rotation appear to be well predicted by sex differences in asymptotes of speed-accuracy curves.

In summary, our results support three conclusions. First, speed-accuracy curves for males and females differed significantly for 90°, 120°, and 150° trials, with females' curves reaching asymptote sooner. Second, when accuracy was controlled, rate of rotation did not differ significantly between the sexes, although females were generally slower than males. However, at a given level of accuracy, problems were relatively more difficult, and thus took longer to solve, for females. On trials of the same relative difficulty, rate of rotation was the same for males and females. Third, sex differences in rate of rotation found in most experiments may reflect the fact that small group differences in asymptotes on speed-accuracy curves have pronounced effects on latencies to solve the

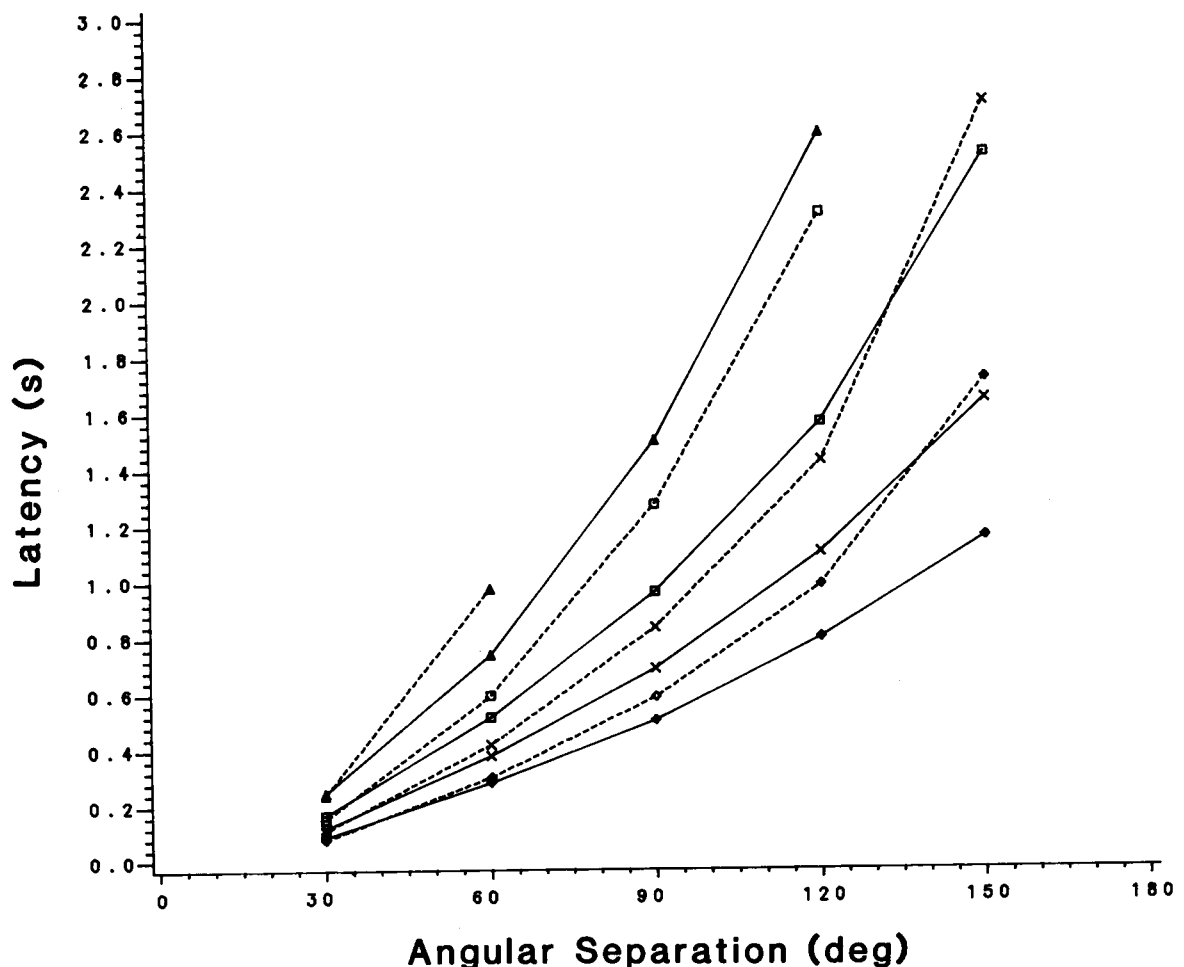


Figure 3. Rotation latency versus angular separation between stimuli with accuracy [ $P(C)$ ] fixed at .70 (diamond, .74, (cross), .78 (square), and .82 (triangle). Solid lines represent data for males; dotted lines, data for females. Fewer than five points are plotted when accuracy is greater than the asymptote of the corresponding speed-accuracy curve (see Figure 2).

more complex problems. These effects will then be compounded when the slope of the regression of latency on angular separation is computed, since the slope will be most strongly influenced by those problems requiring the most rotation.

**Spatial ability differences.** The same patterns emerged when we compared high- and low-spatial males, high- and low-spatial females, and groups of high- and low-spatial subjects formed without regard for sex. Results were similar in all cases. Figure 4 shows speed-accuracy curves for each level of rotation for high- and low-spatial subjects. Dashed lines represent data for high-spatial subjects; solid lines, data for low-spatial subjects. Each curve was based on the 10 means reported in Table 3. High-spatial subjects were here defined as the 28 subjects above the median on both the SR composite (sum of standard scores on Cards and Figures) and the Vz composite (sum of standard scores on Paper Folding and Form Board); low-spatial subjects were the 28 subjects below the median on both the SR and Vz composites. Curves for the two off-diagonal groups were intermediate and thus are not reported here. The most consistent difference between

high- and low-spatials was in the asymptote of the speed-accuracy curves. For all five curves, the asymptote for high-spatials was significantly higher than the asymptote for low-spatials (see Table 4). Differences in asymptotes between high- and low-spatials were smallest for 30° trials and largest for 150° trials. In other words, low-spatial subjects showed significantly greater declines in asymptote than high-spatial subjects with increases in the amount of rotation required [ $F(1,90) = 6.82, p < .05$  for the spatial group by rotation linear interaction]. Thus, high spatial ability means, in part, superior accuracy in solving complex rotation problems. Differences between groups in the  $\beta$  parameter were much smaller. Curvatures were greater for high-spatials; however, the difference was significant only for 90° trials. This suggests that low- and high-spatial subjects differ primarily in the level of performance eventually reached ( $\alpha$ ), and only secondarily in the rate at which they improve when given additional processing time ( $\beta$ ). Furthermore, it is possible that the one significant difference in  $\beta$  at 90° merely reflects the fact that  $\alpha$  and  $\beta$  are correlated in these models (median  $r = .68$ ).<sup>2</sup>

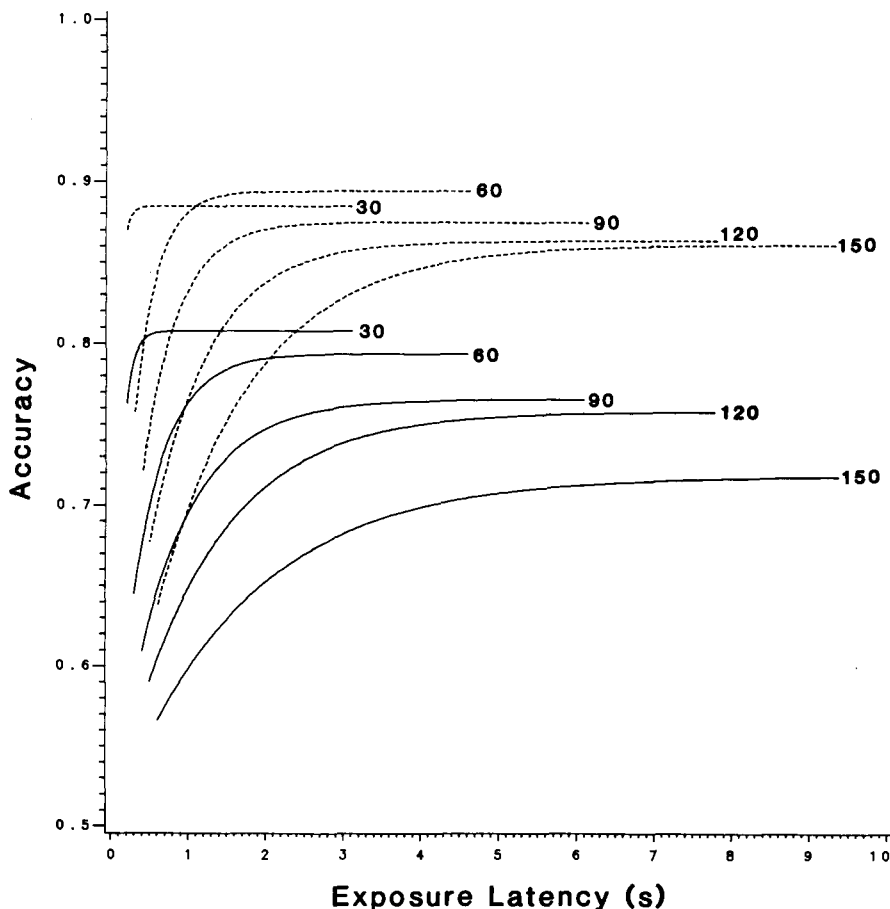


Figure 4. Speed-accuracy curves for low-spatial subjects (solid lines) and high-spatial subjects (dashed lines) at each of five levels of angular separation between stimuli. Numbers on each curve represent the angular separation between figures. The length of each curve indicates the range of exposure latencies for trials requiring a given amount of rotation.

**Table 3**  
Mean Accuracy Scores, by Level of Stimulus Exposure and Angular Separation Between Stimuli, Separately for High- and Low-Spatials

Exposure*	Angular Separation				
	30°	60°	90°	120°	150°
Low-Spatials					
5R+1	.763	.695	.645	.626	.597
9R+1	.774	.731	.664	.628	.609
12R+1	.785	.698	.702	.712	.651
18R+1	.826	.742	.751	.701	.690
35R+1	.833	.820	.785	.744	.697
11R+40	.744	.724	.695	.674	.630
18R+100	.787	.748	.710	.722	.673
29R+130	.816	.776	.749	.762	.694
40R+500	.805	.813	.761	.748	.743
51R+1400	.837	.806	.770	.769	.710
High-Spatials					
5R+1	.864	.807	.767	.754	.686
9R+1	.853	.810	.819	.770	.734
12R+1	.831	.837	.821	.781	.753
18R+1	.881	.840	.823	.813	.780
35R+1	.902	.900	.894	.886	.857
11R+40	.873	.849	.821	.795	.792
18R+100	.891	.899	.854	.836	.826
29R+130	.908	.886	.886	.872	.847
40R+500	.895	.923	.871	.854	.865
51R+1400	.912	.924	.908	.886	.873

\*Stimulus exposure in milliseconds, where R is the angular separation between figures. The first five equations were used to generate rapidly paced trials, and the second five equations were used to generate normally paced trials.

Figure 5 shows plots of estimated response latency versus angular separation between figures for various levels of accuracy, separately for high- and low-spatial subjects. As before, slopes of the lines increase as accuracy is increased. Furthermore, some lines contain fewer than five points, which occurs when the chosen level of accuracy is greater than the asymptote of the curve, and others show a distinct elevation in the last point plotted, which occurs when the chosen level of accuracy is near the asymptote of the curve.

For a given level of accuracy, the slope of the line is generally steeper for low-spatials than for high-spatials. However, when accuracy is fixed, low-spatials are always operating at a different point on the speed-accuracy curve than are high-spatials, and so problems are relatively more difficult for low-spatials. On average, asymptotes for high- and low-spatials differed by .11. Thus, low- and high-spatials should be at approximately the same points on their respective speed-accuracy curves if we compare the line for low-spatials at  $P(C) = .70$  with the line for high-spatials at  $P(C) = .82$ . Figure 5 shows these lines to be virtually coincident. The same relationship holds for low-spatials at  $P(C) = .74$  and high-spatials at  $P(C) = .86$ . In both cases, the last point plotted is significantly elevated in one group or the other, depending on which curve was approaching asymptote.

Speed-accuracy curves were also estimated for each subject at each level of angular separation. However, plots of these curves revealed that data for some individuals

were simply too variable to allow estimation of five separate curves. The most frequent problem was that one speed-accuracy curve was seriously distorted. Therefore, we estimated one asymptote and one curvature parameter for each subject by taking the median of the five curvature ( $\beta$ ) and the median of the five asymptote ( $\alpha$ ) parameters for each subject. The distribution of individual  $\beta$ s was markedly skewed, and therefore this distribution was normalized by rank ordering the  $\beta$ s and then transforming these ranks into normal scores using Blom's (1985, p. 145) method. The  $\alpha$  and normalized  $\beta$  parameters were then correlated with the SR composite (Figures plus Cards) and the VZ composite (Paper Folding plus Form Board). The results are shown in Table 5. The SR composite showed the same correlation with both parameters, whereas the VZ composite correlated higher with the  $\alpha$  parameter than with the  $\beta$  parameter [ $r(80) = 2.57, p < .05$ ].

These correlations concur with the previous analysis, in which spatial ability was treated as a single two-level variable, in indicating that asymptotes of the speed-accuracy curves were higher for high-spatial subjects than for low-spatial subjects. However, they diverge somewhat from the previous analysis in suggesting that speed-accuracy curves reach asymptote faster for subjects who score well on the SR or speeded rotation tests. Furthermore, this pattern of correlations supports the hypothesis that VZ tests estimate level or power, whereas SR tests estimate speed of executing the rotation transformation (Lohman, 1979). However, these conclusions must be tentative, given the extreme variability in individual  $\beta$ s and the fact that these correlations are based on a transformed  $\beta$  score.

Certainly, subjects differ in the speed with which they execute different spatial transformations (Shepard & Cooper, 1982); such differences may signal important differences in the ways stimuli are processed (e.g., Cooper, 1976) or may be important indicators of those sorts of spatial abilities measured by perceptual speed tests, closure speed tests, or tests for other spatial factors. Individual and group differences in the spatial abilities measured by the more complex spatial tasks, such as the Shepard-Metzler (1971) task, or by other tests that load on the factor called visualization (Lohman, 1979), are probably better described by differences in asymp-

**Table 4**  
Differences Between Groups of High- and Low-Spatial Subjects in the Asymptote and Curvature Parameters for Speed-Accuracy Curves at Each Level of Rotation

Rotation	Parameter					
	Asymptote ( $\alpha + .5$ )			Curvature ( $\beta$ )		
	High	Low	$t(18)$	High	Low	$t(18)$
30°	.885	.808	5.39†	16.28	9.64	1.09
60°	.894	.794	4.90†	3.53	2.27	1.60
90°	.875	.766	7.39†	2.21	1.33	2.41*
120°	.864	.759	6.07†	1.34	.86	2.05
150°	.862	.720	10.40†	.80	.60	1.92

\* $p < .05$ . † $p < .001$ .



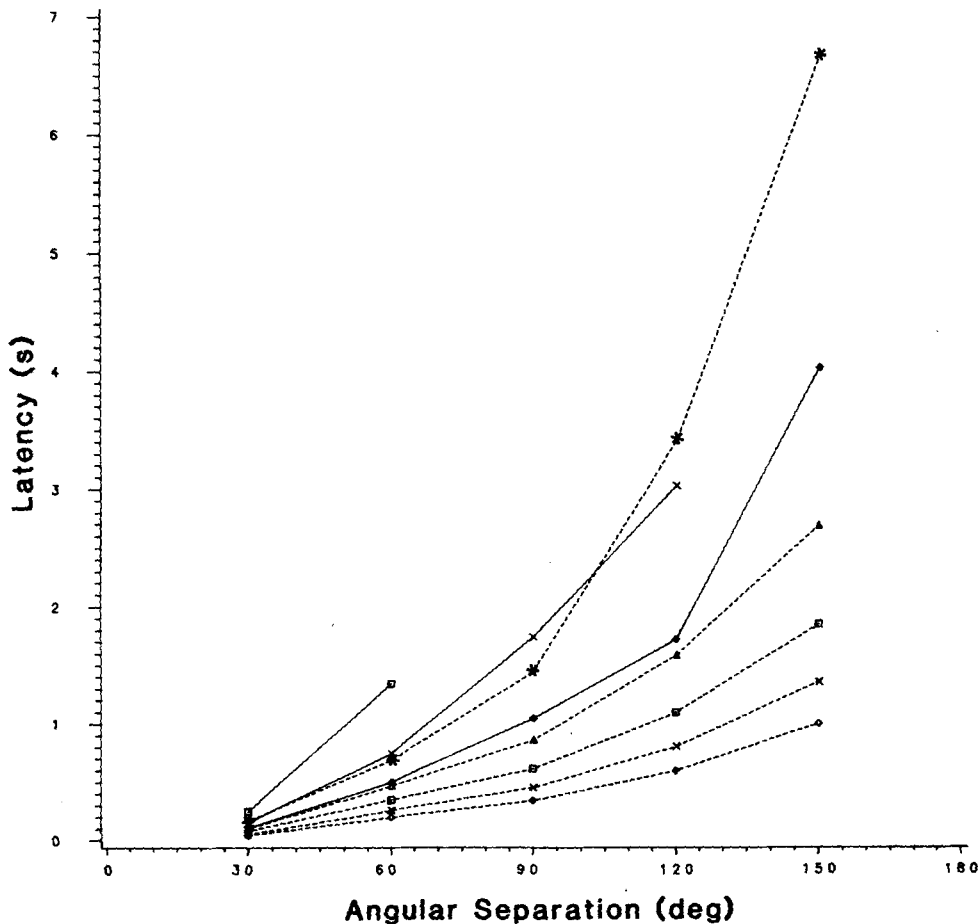


Figure 5. Rotation latency versus angular separation between stimuli with accuracy [ $P(C)$ ] fixed at .70 (diamond), .74 (cross), .78 (square), .82 (triangle), and .86 (asterisk). Dashed lines represent data for high-spatial; solid lines, data for low-spatial. Fewer than five points are plotted when accuracy is greater than the asymptote of the corresponding speed-accuracy curve (see Figure 4).

tote of the speed-accuracy curves for problems of a given level of complexity, or, inversely, by differences in the complexity of problems subjects can solve at a given level of accuracy. This implicates quality of representation as a crucial factor in spatial ability. High-spatial subjects of either sex seem particularly able to construct, maintain, and transform complex but systematically structured visual images.

In summary, the large sex difference often found in rate of rotation may be a predictable consequence of small sex differences in asymptotes of speed-accuracy curves for problems requiring the most rotation. These differences are then magnified a second time when response latencies are regressed on angular separation between figures, since trials requiring the most rotation influence the slope more than trials requiring the average amount of rotation. Comparisons based on spatial ability rather than on sex produced similar effects. Our results are congruent with the finding of Kail et al. (1979) that the distribution of

slopes for females in their sample was bimodal, with 30% of the females having slopes outside of the male distribution. Excessively long slopes would be produced by subjects operating at or near asymptote of the speed-accuracy curve. These results also demonstrate why equating groups on accuracy equates them on speed-accuracy tradeoff only if speed-accuracy curves are identical for both groups at all levels of trial complexity.

Table 5  
Correlations Between Median  $\alpha$ , Median Normalized  $\beta$ , and Composite Scores for the SR and Vz Spatial Factors ( $N=83$ )

Composite	Parameter	
	$\alpha$	$\beta$
SR	.44*	.43*
Vz	.54*	.29*

Note—SR is the sum of standard scores on the Figures and Cards tests; Vz is the sum of standard scores on the Paper Folding and Form Board tests (all tests from Ekstrom et al., 1976).  $\beta$ s were first ranked and then normalized using Blom's (1958, p. 145) method. \* $p < .01$ .

## REFERENCES

- ALDERTON, D. L., PELLEGRINO, J. W., & LYDIATT, S. (1984). *Effects of extended practice on spatial information processing ability*. Unpublished manuscript, University of California, Santa Barbara.
- BERG, C., HERTZOG, C., & HUNT, E. (1982). Age differences in the speed of mental rotation. *Developmental Psychology*, **18**, 95-107.
- BLOM, G. (1958). *Statistical estimates and transformed Beta variables*. New York: Wiley.
- CAPLAN, P. J., MACPHERSON, G. M., & TOBIN, P. (1985). Do sex-related differences in spatial abilities exist? A multilevel critique with new data. *American Psychologist*, **40**, 786-799.
- COOPER, L. A. (1976). Individual differences in visual comparison processes. *Perception & Psychophysics*, **19**, 433-444.
- EGAN, J. P. (1958). *Recognition memory and the operating characteristic* (Tech. Rep. No. AFCRC-TR-57-50). Bloomington: Indiana University, Hearing and Communication Lab.
- EKSTROM, R. B., FRENCH, J. W., & HARMAN, H. H. (1976). *Manual for the kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.
- FURNEAUX, W. D. (1961). Intellectual abilities and problem-solving behavior. In H. J. Eysenck (Ed.), *Handbook of abnormal psychology* (pp. 167-192). New York: Basic Books.
- HARRIS, L. J. (1978). Sex differences in spatial ability: Possible environmental, genetic, and neurological factors. In M. Kinsbourne (Ed.), *Asymmetrical function of the brain* (pp. 405-522). New York: Cambridge University Press.
- HYDÉ, J. S. (1981). How large are cognitive gender differences? A meta-analysis using  $\omega^2$  and  $d$ . *American Psychologist*, **36**, 892-901.
- JUST, M. A., & CARPENTER, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, **8**, 441-480.
- KAIL, R., CARTER, P., & PELLEGRINO, J. W. (1979). The locus of sex differences in spatial ability. *Perception & Psychophysics*, **26**, 182-186.
- KAIL, R., STEVENSON, M. R., & BLACK, K. N. (1984). Absence of a sex difference in algorithms for spatial problem solving. *Intelligence*, **8**, 37-46.
- LANSMAN, M. (1981). Ability factors and the speed of information processing. In M. P. Friedman, J. P. Das, & N. O'Connor (Eds.), *Intelligence and learning* (pp. 441-458). New York: Plenum Press.
- LINN, M. C., & PETERSEN, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. *Child Development*, **56**, 1479-1498.
- LOHMAN, D. F. (1979). *Spatial ability: A review and reanalysis of the correlation literature* (Tech. Rep. No. 8). Stanford, CA: Stanford University, School of Education, Aptitude Research Project. (NTIS No. AD-A075 972)
- MACCOBY, E. E., JACKLIN, C. N. (1974). *Psychology of sex differences*. Stanford, CA: Stanford University Press.
- MCGEE, M. G. (1979). *Human spatial abilities: Sources of sex differences*. New York: Praeger.
- McNICHOL, D. (1972). *A primer of signal detection theory*. London: George Allen & Unwin.
- METZLER, J. (1973). *Cognitive analogues of the rotation of three-dimensional objects* (Doctoral dissertation, Stanford University, 1973). *Dissertation Abstracts International*, **34**, 2973B.
- PACHELLA, R. G. (1974). The interpretation of reaction time in information processing research. In B. Kantowitz (Ed.), *Human information processing: Tutorials in performance and cognition* (pp. 41-82). Hillsdale, NJ: Erlbaum.
- PALMER, S. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology*, **9**, 441-474.
- 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.
- REGIAN, J. W., & PELLEGRINO, J. W. (1984). *Practice and transfer effects in two-dimensional mental rotation*. Unpublished manuscript, University of California, Santa Barbara.
- SHEPARD, R. N., & COOPER, L. A. (1982). *Mental images and their transformations*. Cambridge, MA: MIT Press.
- SHEPARD, R. N., & METZLER, J. (1971). Mental rotation of three-dimensional objects. *Science*, **171**, 701-703.
- SHERMAN, J. A. (1978). *Sex-related cognitive differences: An essay on theory and evidence*. Springfield, IL: Charles C. Thomas.
- TAPLEY, S. M., & BRYDEN, M. P. (1977). An investigation of sex differences in spatial ability: Mental rotation of three-dimensional objects. *Canadian Journal of Psychology*, **31**, 122-130.
- THURSTONE, L. L. (1937). Ability, motivation, and speed. *Psychometrika*, **2**, 249-254.
- THURSTONE, T. G. (1958). *Manual for the primary mental abilities*. Chicago, IL: SRA.
- WICKELGREN, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, **41**, 67-85.
- YUILLE, J., & STEIGER, J. (1982). Nonholistic processing in mental rotation: Some suggestive evidence. *Perception & Psychophysics*, **31**, 201-209.

## NOTES

1. A method that is widely used in signal detection analyses is to ask subjects to rate each response on something more than a 2-point (yes/no) scale, and then to estimate a receiver operating curve (ROC) using these ratings for each subject, separately for each condition in the experiment (Egan, 1958).  $P(A)$ , which is a nonparametric measure of response sensitivity that is independent of response bias, may then be computed by finding the area under each ROC curve (McNichol, 1972).  $P(A)$  was computed for each subject at each level of stimulus exposure for each level of rotation. As frequently happens, the distribution of  $P(A)$  was markedly skewed and so a  $2(\arcsin \sqrt{P(A)})$  transformation was performed before further analyses (see McNichol, 1972). However, this triply transformed score was highly correlated with a simple accuracy score,  $P(C)$ , which is the probability of responding "definitely yes" to positive trials and "definitely no" to negative trials. The median correlation between  $2(\arcsin \sqrt{P(A)})$  and  $P(C)$  within the 50 cells of design was .92. Therefore, at the suggestion of an early reviewer and with some disappointment, we report our results in terms of the simpler score,  $P(C)$ .

2. This difference in  $\beta$  was not significant in the analyses of  $P(A)$  (see Note 1).

(Manuscript received November 21, 1985;  
accepted for publication May 12, 1986.)