

## Mental Rotation of Random Two-Dimensional Shapes

LYNN A. COOPER

*University of California, San Diego*

Two experiments are reported in which *Ss* were required to determine whether a random, angular form, presented at any of a number of picture-plane orientations, was a "standard" or "reflected" version. Average time required to make this determination increased linearly with the angular departure of the form from a previously learned orientation. The slope and intercept of the reaction-time (RT) function were virtually constant, regardless of the perceptual complexity of the test form and the orientation selected for initial learning.

When *Ss* were informed in advance as to the identity and the orientation of the upcoming test form and, further, were permitted to indicate when they were prepared for its external presentation, RT for determining the version of the form was constant for all test-form orientations. However, the time needed to prepare for the test-form presentation increased linearly with the angular departure of the form from the learned orientation.

It is argued that the processes both of preparing for and of responding to a disoriented test form consist of the mental rotation of an image, and that both sorts of mental rotation (pre-stimulus and post-stimulus) are carried out at essentially the same constant rate.

During the past several years, experimental and theoretical investigation of nonverbal internal representation, particularly mental imagery, has proliferated. The primary focus of this renewed experimental effort has been directed toward questions concerning the functional significance of mental imagery (e.g., Bower, 1972; Paivio, 1971). With the exception of evidence concerning the modality or the coded form of internal representations—deriving primarily from the "selective interference" paradigm (cf., Brooks, 1968; Segal, 1971; Segal & Fusella, 1970; Segal & Gordon, 1969) and the selective reduction of reaction

This report is based on a thesis submitted in partial fulfillment for the Ph.D. degree at Stanford University. The author thanks, especially, her thesis advisor, Roger N. Shepard, for his advice, encouragement, and inspiration. Thanks are also due to the other members of the dissertation committee—Herbert H. Clark, Leo Ganz, and Edward E. Smith. This research was supported by National Science Foundation Grant GB-31971X to Roger N. Shepard. Requests for reprints should be sent to Lynn A. Cooper, Department of Psychology, University of California, San Diego, La Jolla, CA 92037.

times in matching tasks (cf., Posner, 1969; Posner, Boies, Eichelman & Taylor, 1969)—little indication of the nature or internal structure of mental images has been provided.

Recently, Shepard and his associates have initiated a series of reaction-time (RT) experiments which explore the relationship between the structure of internal representations such as mental images and the structure of the external objects to which these internal representations correspond (cf., Cooper & Shepard, 1973a, 1973b; Metzler & Shepard, 1974; Shepard, 1975; Shepard & Feng, 1972; Shepard & Metzler, 1971). In an initial experiment, Shepard and Metzler (1971) reported that the amount of time required to determine whether pairs of perspective line drawings were of the same shape increased linearly with the angular difference between the two objects in the pair. The intercept and slope of this linear function were virtually identical for pairs which differed by a rotation in the two-dimensional picture plane and pairs which differed by a rotation about an axis in depth. Shepard and Metzler argued that the task was performed by “mentally rotating” a representation of one object in the pair into congruence with the other object and then checking for a match or a mismatch in shape. For these complex, unfamiliar, three-dimensional stimuli, this process of mental rotation has an average rate of 60°/sec.

In further studies of mental transformations of visual stimuli, Cooper and Shepard (1973a, 1973b) reported that RT for discriminating “normal” from “backward” versions of individually presented, rotated alphanumeric characters increased monotonically with the angular departure of the character from the standard, upright orientation. Despite the consistent nonlinearity of the RT functions, Cooper and Shepard suggested that the version of a tilted test character is determined by mentally rotating an internal representation of the character into congruence with a long-term memory representation of the normal, upright version of the appropriate letter or number.

The Cooper-Shepard experiments also included conditions in which Ss were provided with advance information concerning the identity and the orientation of an ensuing test stimulus, for a variable amount of time, and were instructed to prepare, during the advance information interval, for the presentation of the test character. On the basis of the RT data, Cooper and Shepard argued that the process of preparation consists of first generating a mental image of the predesignated character and then mentally rotating this image into the predesignated orientation. If given enough time to complete this “preparatory” rotation for the orientation indicated in advance, Ss can use the internally generated and pre-rotated internal representation as a “mental template” against which to compare the external test stimulus rapidly and accurately.

While the RT experiments discussed above provide much evidence concerning the nature of mental transformations and of the internal representations being transformed, several problems remain unclarified. The first problem concerns the relationship between pre-stimulus or preparatory mental rotation and post-stimulus rotation (i.e., rotation of a transformed test stimulus in the absence of preparation). Cooper and Shepard (1973b) proposed that both processes involve the rotation of a mental image and, hence, should be carried out in the same manner and at the same rate. The evidence for this proposal was not conclusive, for in the Cooper-Shepard experiments a direct measure of the time required to complete preparatory mental rotations was not obtained. Experiment II reported here was designed to explore further the process of preparing for a rotated test stimulus and its relationship to post-stimulus mental rotation by providing direct estimates of the time needed to carry out both internal processes.

A second unresolved matter concerns the nonlinearity of the RT functions obtained in the experiments with alphanumeric stimuli. Cooper and Shepard (1973b) have suggested several possible explanations for the nonlinear effect of test-stimulus orientation on RT, all of which are consonant with the notion that a mental rotation is carried out in order to determine the version of a tilted test character. Experiment I reported here was designed to evaluate an explanation which attributes the nonlinearity to two related conjectures concerning familiarity. First, alphanumeric characters, which are generally encountered in or close to the conventional upright position, may seem less familiar when viewed at markedly tipped orientations (cf., Egeth & Blecker, 1971). Second, the rate at which an object can be mentally rotated may increase with the familiarity of that object. Under this account, rotation rate should be slowest for a test stimulus close to an unfamiliar orientation and should accelerate as the stimulus approaches a familiar or learned position.

A final unexplored issue concerns the relationship between rate of mental rotation and complexity of the internal representation undergoing the mental transformation. Rotation rates estimated by Cooper and Shepard for alphanumeric characters were some six times faster than the 60°/sec rate estimated by Shepard and Metzler for complex perspective drawings. The possibility that rotation rate decreases with increasing complexity of the test form being rotated was evaluated in both experiments reported here by employing stimuli which differ on a well-defined measure of perceptual complexity. The implications of the complexity manipulation for the nature of the internal processes and representations underlying these tasks will be discussed in connection with the experiments.

## EXPERIMENT I

## Method

*Subjects*

Eight *Ss*, all students and staff at Stanford University, were paid for their participation in six one-hour experimental sessions. Six of the *Ss* were male, two *Ss* were female, and one male *S* was left-handed. Three of the *Ss* had participated in pilot work for this experiment (cf., Cooper, 1973).

*Stimuli*

The stimuli were the eight random shapes illustrated in Fig. 1. These angular forms were generated by Attneave and Arnoult's (1956) Method I for the construction of random nonsense shapes. The particular forms used as stimuli were selected from a set of shapes which Vandeplass and Garvin (1959) reported to be low in verbal association value.

Studies of rated complexity of forms generated by this method indicate that judged perceptual complexity depends strongly upon the number of points which determine inflections on the perimeter of the

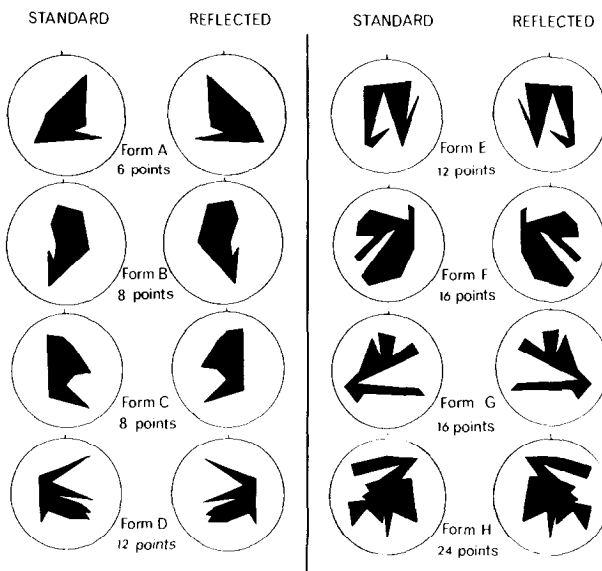


FIG. 1. The eight random forms used in Experiment I, displayed in both standard and reflected versions.

form (Attneave, 1957; Attneave & Arnoult, 1956; Vanderplas & Garvin, 1959). Attneave (1957) found a linear relationship between the logarithm of the number of points and judged complexity and also reported that the number of points (usually identical to the number of angles in the contour) accounts for 80% of the variance of the ratings.

The eight forms depicted in Fig. 1 represent five levels of rated complexity—six, eight, 12, 16, and 24 points. Thus, the stimulus set included one form of the lowest and highest complexity values (forms A and H), and two forms of each of the intermediate levels of complexity (forms B, C, D, E, F, and G).

Different Ss learned to discriminate “standard” versions of the forms from “reflected” or mirror-image versions (cf., Fig. 1) at different previously determined training orientations. The training positions were six equally-spaced orientations about a circle in the picture plane. For standard versions of all forms, the six training orientations consisted of angular departures of  $60^\circ$  steps of clockwise rotation from the orientation depicted in Fig. 1 (including, of course, the depicted orientation itself). In order to preserve the mirror-image relationship between standard and reflected forms, the six corresponding training orientations for reflected versions consisted of the orientation illustrated in Fig. 1 plus angular departures of  $60^\circ$  steps of *counterclockwise* rotation from this orientation.

In order to control for particular characteristics of certain training orientations (e.g., natural alignment of major contours of the forms with respect to a two-dimensional frame of reference), training orientation was varied over Ss. Three Ss had participated in pilot work in which a subset of the forms had been presented in the orientation illustrated in Fig. 1. Hence, all three Ss were assigned to this orientation for initial training. The other Ss were assigned randomly to the five remaining orientations.

The forms were presented in an Iconix three-field tachistoscope and appeared centered within an illuminated circular field with a black surround. The forms themselves subtended a visual angle of about  $2^\circ$ , and the circular aperture in which they appeared subtended an angle of  $4^\circ$ . Luminance of all fields of the tachistoscope was about 20 ft-L.

### *Procedure*

During the first experimental session, Ss learned to discriminate standard from reflected forms at the appropriate training orientation only. Each S was permitted to study a visual display containing drawings of the eight forms, in both standard and reflected versions in the training position only, for about ten minutes. Individual forms were then pre-

sented tachistoscopically (in the training orientation only), and the *S* was required to determine as quickly as possible whether each form presented was a standard or a reflected version.

"Standard" responses were signaled by pushing a right-hand button, and "reflected" responses were signaled by pushing a left-hand button. This response assignment was reversed for the one left-handed *S*; thus, the preferred hand was always used to signal "standard." A two-second gray warning field preceded the presentation of the test form, and the *S*'s button-pressing response terminated the visual display. Between trials the *E* changed stimuli and recorded RT and errors. For each *S*, each of the eight forms was presented in both standard and reflected versions ten times, for a total of 160 training trials.

Sessions two through six were test sessions. At the beginning of each such session, the *S* was familiarized with the forms by means of a small set of training trials consisting of two presentations of both standard and reflected versions of each of the eight forms at the trained orientation. Subjects were required to discriminate standard from reflected versions, and RT's were recorded.

During the remainder of each session, the same eight forms were presented, but each form could appear in any of six possible orientations about the circle. The six orientations were equally spaced and consisted of the trained orientation plus 60°, 120°, 180°, 240°, and 300° angular departures of clockwise rotation from the trained orientation. Regardless of the orientation at which each test form appeared, *S*s always had to determine as quickly as possible whether the form was a standard or a reflected version. As in the training session, choice responses were registered by pushing a right- or left-hand button.

On each of the five test days (in addition to the initial relearning trials) each *S* saw each of the eight forms in both standard and reflected versions *twice* at the trained orientation, *twice* at the orientation departing 180° from the trained orientation, and *once* at each of the four other orientations. These unequal probabilities of appearance were designed to yield an equal number of observations at each angular departure, collapsed over clockwise and counterclockwise directions, from the trained orientation. Thus, each test session consisted of 32 retraining trials and 128 test trials. The order of test trials was randomized anew for each session within each *S*.

Trials on which errors were made were retaken within the same session in order to obtain a complete set of error-free data for each *S*. Session two, the first of the five test sessions, was considered a practice day. The data from this session were not included in the analysis. Consequently, the complete set of test-session data consists of 512 errorless RT's for each of the eight *S*s.

## Results

Figure 2 illustrates mean RT (averaged over *Ss*, forms, and sessions) for correctly determining the version of rotated test forms. Although error RT's are not included in the mean RT's in Fig. 2, error rates are plotted as a function of angular departure from the trained orientation and standard versus reflected versions. Since individual *Ss* learned the forms in different positions, the 0° orientation does not correspond to any unique position of the forms. In addition, clockwise and counterclockwise departures from the trained orientation have been averaged in Fig. 2 and in all of the figures which follow. However, if the RT function is "unfolded" about the 180° point, the shape is remarkably symmetrical (cf., Cooper, 1973).

The most striking features of the data presented in Fig. 2 are the linearity of the increase in RT with angular departure of the test stimulus from the trained orientation and the parallelism between the functions for "standard" and "reflected" responses. The greater speed of the "standard" response and the parallelism of the two functions has also emerged in previous studies using alphanumeric stimuli (Cooper & Shepard, 1973b).

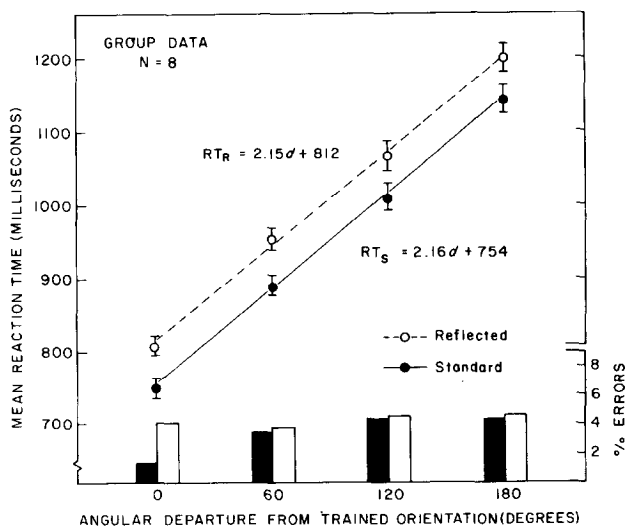


FIG. 2. Mean RT as a function of angular departure of the test form from the trained orientation for the group data from Experiment I. "Standard" and "reflected" RT functions are plotted separately, and equations for the best-fitting straight lines are shown. Vertical bars about each mean RT represent  $\pm$  one standard error of the mean. Error rates are plotted as a function of orientation and version, with solid bars representing "standard" errors and open bars representing "reflected" errors.

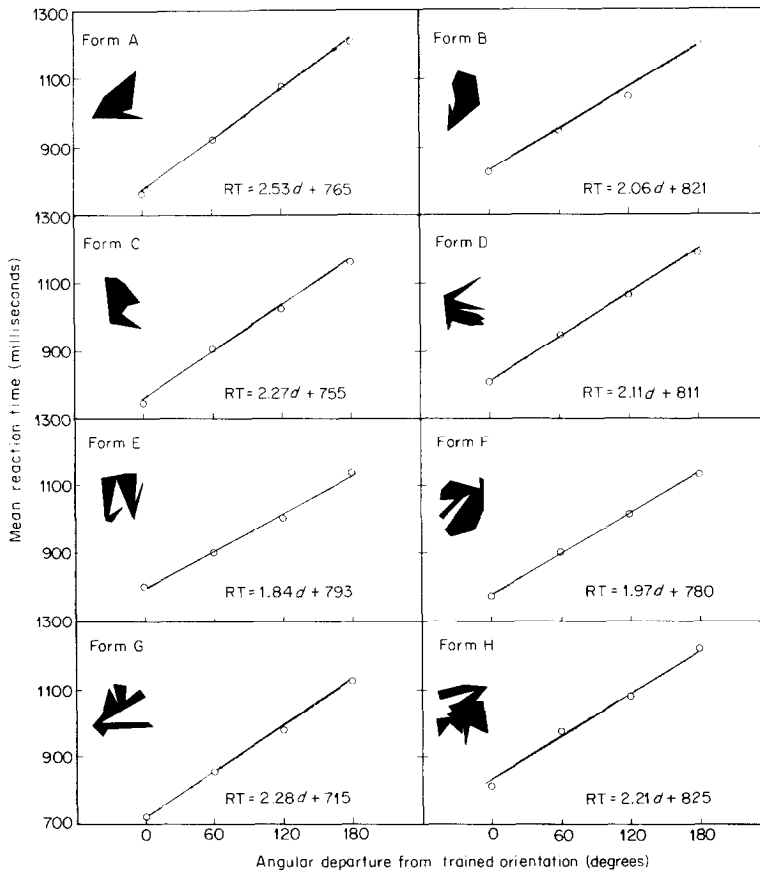


FIG. 3. Mean RT as a function of angular departure of the test form from the trained orientation, plotted separately for each of the eight forms used in Experiment I. Equations for the best-fitting straight lines are shown.

In Fig. 3, group RT functions, averaged over "standard" and "reflected" responses, are plotted separately for each of the eight forms. The striking linearity depicted in Fig. 2 is also apparent when each form is considered separately. The equations for least-square linear fits, shown for each form in Fig. 3, indicate that the complexity of the test form, measured by the number of points, does not produce systematic differences in intercept or slope. This suggestion is confirmed by the fact that correlations between (a) rank order of complexity of the forms and rank order of the corresponding intercepts, and (b) rank order of the complexity of the forms and rank order of the associated slopes, were nonsignificant (for complexity and intercept,  $r_s = +.24$ ; for complexity and slope,  $r_s = -.26$ , with  $N = 8$  for both correlations).



TABLE 1  
SLOPE ESTIMATES AND STANDARD ERRORS, COLLAPSED OVER FORMS,  
SESSIONS, AND VERSIONS, FOR INDIVIDUAL *SS* AND FOR THE  
GROUP DATA FROM EXPERIMENTS I AND II<sup>a</sup>

Subject	Experiment I		Experiment II	
	Slope	Standard error of slope	Slope	Standard error of slope
1 (JC)	1.54	.12	2.15	.12
2 (IF)	1.82	.07	1.82	.04
3 (CH)	2.55	.17	1.99	.06
4 (DK)	2.54	.06	3.10	.07
5 (JK)	3.04	.11	4.23	.13
6 (SR)	2.55	.04	3.75	.25
7 (GL)	1.18	.05	1.77	.05
8 (LC)	2.07	.08	2.85	.12
Group	2.16	.06	2.71	.08

<sup>a</sup> Estimates for Experiment I are based on all eight forms, while estimates for Experiment II are based on only four forms.

The left-hand panel of Table 1 presents the slope estimate and its standard error for the data of each *S* (over forms) and for the group data. In addition, a subset of the data for individual *SS* is plotted in Fig. 6 in connection with Experiment II. For a complete presentation of the individual *S* data, see Cooper (1973).

A five-way analysis of variance (*SS* × sessions × forms × orientations × standard-vs-reflected versions) was performed on the group data.<sup>1</sup> The results of the analysis can be summarized as follows: (a) The main effects of *SS*, sessions, orientations, and versions were all highly significant (*p*-values ranging from .01 to .001). (b) The main effect of forms did not approach statistical significance. (c) All but one of the two-way interactions with the factor "*SS*" were significant (for all, *p* < .001). (The "*SS* × versions" interaction was not significant.) (d) None of the other two-way interactions was statistically significant. (e) Two of the three-way interactions—" *SS* × forms × orientations" and "*SS* × forms × versions"—were significant (for both, *p* < .01). (f) None of the other three-way or higher-order interactions achieved statistical significance. A second analysis of variance, in which just the linear component of the orientation factor was used, revealed the "sessions × orientations"

<sup>1</sup> In the group analysis of variance, both "*SS*" and "forms" were treated as random factors. *F*-ratios and quasi *F*-ratios were computed and degrees of freedom were estimated by the procedure recommended by Clark (1973). See Cooper (1973) for a complete presentation of this analysis and for a presentation of analyses of variance done on the data of individual *SS*.

interaction to be marginally significant ( $p < .05$ ). None of the other two-way interactions tested in this analysis were statistically significant.

Polynomial regressions were computed on the mean RT's for each form within each *S* and for the group data. In all the analyses, no quadratic or higher-order effects were significant. Root mean square deviations (RMSD's) were computed from the residual variance not attributable to linearity. RMSD's for individual *S*s ranged from 23 to 5 msec, for individual forms from 18 to 6 msec, and for the group data the RMSD was 8 msec. None of the RMSD's indicated statistically significant departures from linearity.

Although all error trials were retaken, errors were recorded and the computed error rates were quite low, ranging from 1.2% to 7.9% for individual *S*s, with an average rate of 3.6% (cf., Fig. 2).

### Discussion

The linearity of the RT functions illustrated in Figs. 2 and 3 has implications for (a) the nature of the internal processes involved in determining the version of a rotated form, and (b) the "familiarity" explanation for the nonlinearity obtained with alphanumeric stimuli (cf., Cooper & Shepard, 1973a, 1973b). With respect to the internal processes underlying the standard-reflected discrimination, these data support the following claims: (a) An internal representation of the visual test form is first mentally rotated from its externally presented orientation into the trained orientation.<sup>2</sup> (b) The mentally transformed image of the test form is then compared to the memory representation of the standard version of that form in the learned position. The overall difference between "standard" and "reflected" RT's suggests that a match between the transformed representation of the test form and the standard version of that form is generally tested for first. If this match fails, then extra time is evidently needed to switch to the nonpreferred "reflected" response. This initial comparison of the transformed internal representation with a memory image of the *standard* version is consistent with the "congruence" principles discussed by Clark and Chase (1972) and Trabasso, Rollins and Shaughnessy (1971). (c) For these random two-dimensional forms and these particular *S*s, this process of mental rotation is carried out at an average constant rate of 460°/sec.

The linear RT functions obtained not only for the group data, but also for each individual *S* (cf., Fig. 6), indicate that relatively long-term mem-

<sup>2</sup> These data do not distinguish between a rotation of the test form into the trained orientation and a rotation of the memory representation of the standard version into the orientation of the test form. Either alternative is both theoretically and empirically acceptable; however, the introspective reports of the *S*s suggest that the test form was rotated into congruence with a memory representation of the standard in the trained orientation.

ory representations of these forms were established in whatever orientation they were presented during initial training. It is noteworthy that, after a mere 160 exposures to the random forms at the trained orientation, Ss apparently continued to use a memory representation of each form in the trained orientation as a "mental template" for comparison with the transformed images of the test forms during the 640 subsequent trials.

If Ss either developed "mental templates" for standard and reflected versions at all orientations or learned to recognize these versions on the basis of distinctive features unique to each orientation, then we should expect the RT function to become flat during the later test sessions. Some evidence for a flattening of the RT function over sessions is found in the marginally significant "sessions  $\times$  orientations" interaction. However, such a flattening is not obtained in the experiment to be reported next, and, more importantly, a complete flattening of the RT function, which would result from adopting either of the alternative strategies mentioned above, is not found. (Practice effects are evident in that the average slope decreases monotonically over test sessions. However, the RT function is linear for all test sessions, and even for the final test session the slope is considerable—1.79 compared with 2.57 for the first test session.) Thus, we can tentatively conclude that, while practice *does* affect the slope of the RT function, nonetheless the same basic internal process of mental rotation was used during the entire sequence of test trials.

Perhaps the most significant outcome of the present experiment is the lack of relationship between complexity of the random forms and rate of mental rotation, as measured by the slope of the RT functions (cf., Fig. 3). The six-fold difference between rotation rate for the complex perspective drawings studied by Shepard and Metzler (1971) and rotation rate for the alphanumeric characters studied by Cooper and Shepard (1973a, 1973b) constituted cogent *a priori* evidence for expecting RT to depend upon variations in complexity within the set of stimuli studied here.

One possible explanation for the failure of complexity variations in the random forms to produce systematic RT differences relates to the nature of the Ss' internal representations or mental images of the forms. Note that in the present experiment, complexity was confounded with redundancy of useable cues for making the standard-reflected determination. Thus, Ss could have achieved the correct response by imagining only a few distinctive features of the test form, regardless of its actual perceptual complexity, rotated into the trained orientation. A related possibility is that the internal representation of the visual test form was highly schematic in comparison with the rich detail of the form itself.

If either a schematic representation or an image of one or more salient

features was sufficient for comparing the transformed internal representation with a memory representation of the appropriate test form, then complexity variations in the external stimulus objects were not necessarily incorporated in the internal representations of those objects. Consequently, stimulus complexity would not be expected to produce RT differences. Thus, the data from this experiment do not legislate conclusively on the issue of complexity in that they fail to provide an indication of the degree of specificity and completeness of the transformed internal representation. Further experiments, currently in progress, are addressed specifically to this question.

Finally, the RT functions obtained in the present experiment suggest that the "familiarity" explanation for the nonlinearity obtained with alphanumeric characters is incorrect. If Ss in the present experiment learned the random forms at the trained orientation, then we should expect that the forms would become most familiar in this position and progressively less familiar when shown at larger and larger angular departures from the trained orientation. Thus, if speed of rotation is positively related to familiarity, then RT should increase monotonically but nonlinearly with angular departure from the familiar orientation.

While the predicted nonlinearity is not evident in Figs. 2 and 3, these data do not conclusively infirm the familiarity interpretation for alphanumeric stimuli. First, alphanumeric characters are much more overlearned in the unique upright orientation than were these random forms in the arbitrarily chosen training orientations. Consequently, any familiarity effect in the case of the random shapes may be relatively weak. Moreover, since the forms were presented in only one training orientation (unlike alphanumeric characters which are frequently encountered in orientations departing somewhat from upright), the spread of any familiarity effect resulting from pretraining may have extended only to orientations departing by less than 60°. Thus, the present experiment may not have been sensitive enough to reveal appreciable effects of familiarity. However, the linear RT data do lead us to regard this explanation with much less enthusiasm.

## EXPERIMENT II

Cooper and Shepard (1973b) have proposed that both the time required to *respond* to an externally presented test character and the time required to *prepare* for the external onset of such a test character reflect the amount of time needed to carry out the rotation of a mental image. Although these two sorts of mental rotation (post-stimulus and pre-stimulus) differ in several respects, it was argued that both methods of image transformation should be carried out in essentially the same manner and at essentially the same rate.

In the present experiment, the rate of preparatory mental rotation is

measured directly. This is accomplished by providing the *S* with advance information concerning the identity and orientation of an upcoming test form and, further, by requiring him to indicate when he is prepared for the external onset of that test form. In order to consider this RT a direct measure of the time needed to prepare for the test stimulus, it must also be established that the *S* is, in fact, fully prepared. This is accomplished by presenting the actual test form immediately following the "preparation" response and by requiring a second, discriminative response to the test form.

By comparing the shapes and the slopes of (a) the function relating preparation RT to anticipated test-form orientation, and (b) the function relating choice RT to angular departure of the test form from a trained orientation (cf., Experiment I above), it should be possible to assess the relationship between the nature and rate of both pre-stimulus and post-stimulus mental rotation.

## Method

### *Subjects*

The same eight *Ss* who had previously participated in Experiment I also participated in the present experiment.

### *Stimuli*

The stimuli were the random forms used in Experiment I. In order to reduce the length of the experiment, each *S* was tested with only four of the eight original forms. The stimulus set for each individual *S* consisted of the four forms which (in that *S*'s Experiment I data) had yielded the largest *F*-values for linear components and had accounted for less than 25% of that *S*'s errors. Although only four forms were used per *S*, the eight forms used in Experiment I were represented approximately equally across *Ss*.

In addition to the test stimuli, advance-information cues were presented visually. Identity information consisted of an outline drawing of the standard version of the upcoming test form in the trained orientation. Orientation information consisted of a black arrow, passing through the center of a circular field and pointing to the orientation at which the top of the test form would appear. The specifications of the apparatus and the stimuli were identical to those in Experiment I.

### *Procedure*

The experiment consisted of four one-hour sessions per *S*. During the first half of the initial session, the *S* was refamiliarized with the four random forms which were to be used throughout the experiment by

means of tachistoscopic presentation of individual forms in the trained orientation only. The *S* was required to indicate *vocally* whether each test form was a standard or a reflected version. "Standard" responses were signaled by saying "S," and "reflected" responses were signaled by saying "R." This voice response was registered by a microphone positioned directly below the *S*'s mouth. The response activated a noise-operated relay (Hunter sensitivity relay indicator 3203) which triggered the stop on a clock, permitting the *E* to record the choice RT. After completing a block of 32 retraining trials (four presentations of both standard and reflected versions of each of the four forms), the remainder of the session consisted of practice for the test trials in subsequent sessions.

During the test trials, the *S* was required to discriminate standard from reflected test forms presented at any of the six orientations used in Experiment I. Prior to the onset of the test form, the *S* was given advance information concerning both its identity and orientation. Figure 4 schematically illustrates the structure of the test trials and the appearance of the advance information. At the beginning of each trial, the *S* was instructed to prepare for the upcoming test form by rotating an image of the pre-indicated form in either a clockwise or a counterclockwise direction. Thus, the orientation cue shown in Fig. 4 could indicate either a 60° clockwise rotation or a 300° counterclockwise rotation, depending upon the prior instruction.

Identity information was always presented for a fixed duration of 3000 msec and was immediately replaced by an arrow, indicating the position at which the top of the test form would appear. The duration of the orientation cue was controlled by the *S*. When the *S* felt fully prepared for the presentation of the test form at the indicated orientation, he pushed a right-hand button (a left-hand button for the one left-handed *S*).

The button-pressing response (henceforth called the "preparation"

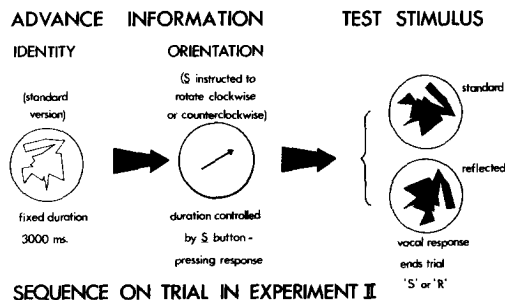


FIG. 4. Schematic illustration of the sequence of visual displays and required responses on a trial in Experiment II.

response) stopped a timer which displayed the total duration of the orientation cue. Immediately following the preparation response, the test form designated in advance appeared in the orientation designated in advance. The *S* was required to determine the version of the test form as quickly as possible by saying "S" for a standard form and "R" for a reflected form. This vocal response stopped a second clock which displayed the total duration of the test stimulus and terminated the visual display. Thus, two RT's, the button-pressing preparation RT and the vocal discriminative RT, were recorded on each test trial.

Following the initial training and practice session, each *S* completed three one-hour test sessions. Within each test session, trials were blocked by direction of rotation such that on half of the trials, clockwise rotations were specified and on the other half of the trials, counterclockwise rotations were specified. The order of clockwise and counterclockwise trial blocks was balanced over sessions. Within each clockwise or counterclockwise block, each of the four forms was presented in both standard and reflected versions at each of the six orientations, for a total of 48 trials per block and 96 trials per test session. Thus, each of the eight *S*s completed 288 test trials (half clockwise, half counterclockwise), and 576 RT's (half preparation, half discriminative) were obtained for each *S*. As in Experiment I, all error trials were retaken.

### Results

The principal results for the group data are illustrated in Fig. 5. The plotted points in the uppermost function, which represent group data from Experiment I, were computed by averaging within each *S* over the four forms subsequently used in Experiment II, and then averaging over all eight *S*s. Thus, the data from Experiments I and II reflect RT's to the same set of forms within each *S* and are perfectly comparable.

The function labeled " $RT_1$ " represents mean preparation time in Experiment II, averaged over clockwise and counterclockwise rotations. The function labeled " $RT_2$ " represents mean discriminative RT, averaged over "standard" and "reflected" responses. Though not shown in Fig. 5, "standard" responses were about 20 msec faster than "reflected" responses for all test-stimulus orientations. Error RT's are not included in the plotted points; however, error rates are plotted as a function of orientation and version.

Average time required to prepare for an upcoming rotated test form ( $RT_1$ ) is a remarkably linear function of the angular departure of the test form from the trained orientation. Furthermore, we can see for the first time that this linear function extends beyond a  $180^\circ$  departure from the trained orientation all the way to  $300^\circ$  (of clockwise or counterclockwise departure, depending upon prior instructions). Average time needed to

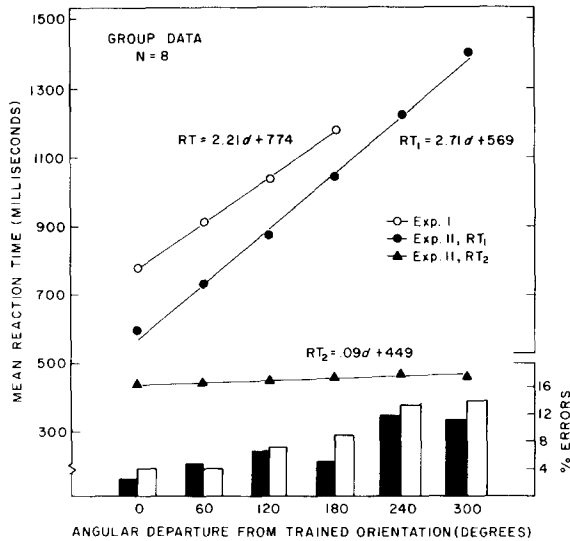


FIG. 5. Mean RT as a function of angular departure of the test form from the trained orientation. The uppermost function illustrates a subset of the group data from Experiment I. The other two functions illustrate group data for both preparatory and discriminative RT's from Experiment II. Equations for the best-fitting straight lines are shown. Error rates are plotted as a function of orientation and version, with solid bars representing "standard" errors and open bars representing "reflected" errors.

make the discriminative (verbal) response (RT<sub>2</sub>) is virtually constant for all angular departures from the trained orientation. This choice RT does increase slightly with orientation; however, the range of mean RT<sub>2</sub> is only 26 msec, compared with the 800-msec range for RT<sub>1</sub>.

This flat RT<sub>2</sub> function demonstrates that Ss were indeed prepared for the onset of the test form when they so indicated. If Ss were not prepared for the onset of the test form when they made the preparatory response, then we should expect discriminative RT to increase with angular departure of the test form from the trained orientation and to reach a maximum value at the 180° departure. Such a relationship is virtually absent in the RT<sub>2</sub> function in Fig. 5. Note, also, that the zero-intercept of the RT<sub>2</sub> function is some 325 msec lower than the post-stimulus RT function from Experiment I.

The rates of pre-stimulus (preparatory) and post-stimulus mental rotation can be estimated directly from the slopes of the RT functions in Fig. 5. Average rate of post-stimulus rotation for the set of data from Experiment I is about 450°/sec, and average rate of pre-stimulus rotation for the same Ss and the same random forms, estimated from the slope of the RT<sub>1</sub> function, is about 370°/sec. For the group data, post-stimulus rotation appears to be somewhat faster than preparatory rota-



tion; however, this relationship between rotation rates is evident for only six of the individual *Ss*. For two of the *Ss*, pre-stimulus rotation is faster than post-stimulus rotation.

In order to assess the comparability between the rates of these two sorts of mental rotation, *t*-tests for differences between the slopes of the  $RT_1$  functions from the present experiment and the slopes of the post-stimulus RT functions from Experiment I were computed. For the group data, the mean difference between the slopes of the Experiment I RT function and the preparation RT function failed to achieve traditional levels of statistical significance ( $t(7) = 2.20$ ;  $.05 < p < .07$ ).<sup>3</sup> For five of the eight *Ss*, similar *t*-tests revealed that the functions were not significantly different in slope, and for the three other *Ss* the difference was marginally significant ( $p < .05$ ).

RT functions corresponding to the group functions shown in Fig. 5 are plotted for each of the individual *Ss* in Fig. 6. (Note that in Fig. 6 the scale for  $RT_2$  has been displaced downward by 200 msec for JK and SR.) Although the eight *Ss* differ considerably in overall RT (particularly in rotation rate), the data for each of the *Ss* capture the principal features of the group data. In addition, each *S* shows cross-experiment consistency. That is, there is a significant positive correlation, among the individual *Ss*, between rank order of the slopes of the post-stimulus RT functions and rank order of the slopes of the preparation RT functions ( $r_s = .74$ ,  $N = 8$ ,  $p < .05$ ). On the other hand, a similar Spearman rank order correlation between slopes of the  $RT_1$  functions and slopes of the  $RT_2$  functions is nonsignificant ( $r_s = .36$ ,  $N = 8$ ,  $p > .05$ ). The right-hand panel of Table 1 presents additional individual *S* data, and Cooper (1973) provides complete data for each of the eight *Ss*.

Five-way analyses of variance (*Ss*  $\times$  sessions  $\times$  directions of rotation  $\times$  versions  $\times$  orientations) were performed on the group data for both preparation and discriminative RT's. Since the assignment of the four random forms was not consistent across *Ss*, the scores used in these analyses were RT's averaged over forms. The results of the analysis of variance on the  $RT_1$  data may be summarized as follows: (a) The only two main effects to reach statistical significance were "*Ss*" and "orientations" (for both,  $p < .001$ ). (b) Two two-way interactions—"Ss  $\times$  sessions" and "Ss  $\times$  orientations"—and two three-way interactions—"Ss  $\times$  sessions  $\times$  orientations" and "Ss  $\times$  sessions  $\times$  versions"—were significant ( $p$ -values ranged from .01 to .001). (c) None of the other interactions achieved statistical significance.

In order to test more sensitively for the possibility of a practice effect on the slope of the  $RT_1$  function, a second analysis of variance, which

<sup>3</sup> The mean difference between the slopes and a confidence interval of one SD was  $.505 \pm .229$ .

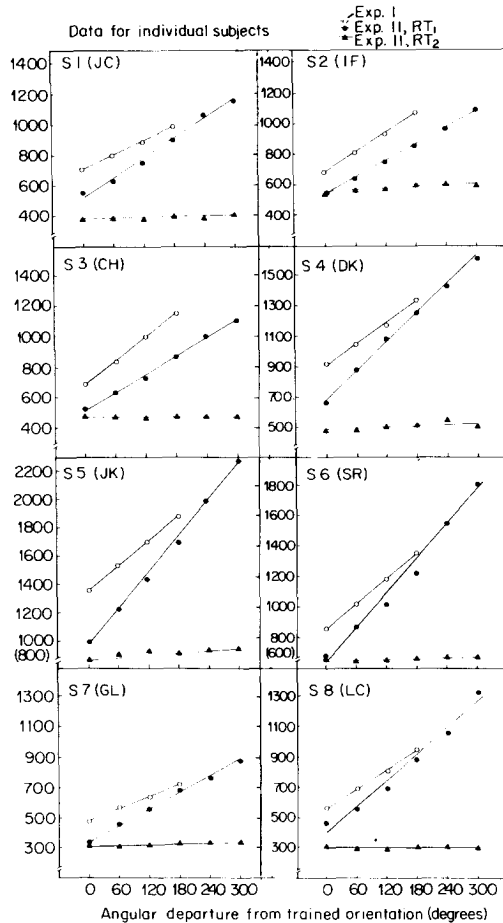


FIG. 6. Mean RT as a function of angular departure of the test form from the trained orientation, plotted separately for each of the eight individual Ss. The functions for each S correspond to the group functions illustrated in Fig. 5.

used just the linear component of the orientation factor, was performed. When tested against the "Ss  $\times$  sessions  $\times$  linear orientations" interaction, the "sessions  $\times$  linear orientations" interaction still proved nonsignificant ( $p > .10$ ), as did all the other interactions which failed to achieve statistical significance in the previous analysis.

The analysis of variance on the RT<sub>2</sub> data is summarized below: (a) The main effects of Ss, standard-vs-reflected versions, and orientations were all statistically significant ( $p$ -values ranged from .05 to .001). (The significance of the main effect of orientation for the group data is due, primarily, to the data of one of the eight Ss, cf., Cooper, 1973.) (b) All of the two-way interactions with the factor "Ss" were significant ( $p$ -values

ranged from .05 to .001), and three of the three-way interactions with this same factor were significant (for " $Ss \times sessions \times rotations$ ",  $p < .01$ ; for " $Ss \times rotations \times versions$ ,"  $p < .01$ ; for " $Ss \times versions \times orientations$ ,"  $p < .05$ ). (c) None of the other interactions or main effects were statistically reliable.

Polynomial regressions for both  $RT_1$  and  $RT_2$  functions were computed for the group data, for individual  $Ss$ , and for forms within  $Ss$ . For  $RT_1$ , all of these regressions revealed a highly significant linear trend ( $p < .001$ ), and no quadratic or higher-order effects were significant ( $p > .05$ ). For  $RT_2$ , several of the regressions yielded marginally significant linear components ( $p < .05$ ).

Error rates were somewhat higher than in Experiment I. For individual  $Ss$ , error rates ranged from 1% to 13.3% with an average of 7.9% (cf., Fig. 5).

### Discussion

The linearity of the  $RT_1$  function supports the notion that  $Ss$  prepare for an upcoming, rotated test form by imagining the standard version of the form indicated by the identity cue rotated into the orientation designated in advance. The internally-generated and pre-rotated mental image can then be compared directly with the ensuing test form. If the two representations match, the  $S$  can respond "standard" immediately. If the two representations are incongruent, extra time is needed to respond "reflected."

The flatness and overall lowness of the  $RT_2$  function indicates that  $Ss$  are fully prepared for the onset of the test form when they execute the preparation response. This flatness further indicates that comparisons between the transformed preparatory image and the visually-presented test form can be made equally rapidly at any angular departure from the trained position *for which the  $S$  is prepared*. Thus we can conclude that to be prepared for the onset of an external stimulus is to have, in advance, an appropriate internal representation of that external stimulus.

The extension of the linear preparation-time function beyond  $180^\circ$  has theoretical significance. One might suppose that the process of preparing for a rotated stimulus consists merely in the selective priming of feature detectors appropriate to the upcoming stimulus in the designated orientation, and, further, that it is more difficult (and hence takes more time) to activate detectors at greater angular departures from a familiar orientation. If this account were correct, then we should expect the  $RT_1$  function to peak at  $180^\circ$  and to be symmetrical about the inverted position. This symmetrical shape would result from the differential difficulty of priming feature detectors at  $0^\circ$ ,  $60^\circ$ ,  $120^\circ$  and  $180^\circ$  angular departures from the trained orientation. Instructions for clockwise or coun-

terclockwise "rotations" should not affect the  $RT_1$  function, for, presumably, the process of activating feature detectors at a certain orientation does not have a directional component.

Such a feature-priming theory has in itself no mechanism to account for the sensitivity of the  $RT_1$  function to direction-of-rotation instructions. A process of mental rotation, however, predicts exactly the relationship illustrated in Fig. 5. The extension of the  $RT_1$  function out to a  $300^\circ$  angular departure provides strong support for the claim that *Ss* carry out a mental analog of an external physical rotation in preparing for an upcoming test stimulus. For, the linear increase in preparation time up to and, particularly, beyond  $180^\circ$  indicates that the mental process is passing through a trajectory in a specific direction, and that the time required to complete the process is a function of the length of the trajectory, rather than the physical departure of the end point (in either direction) from the trained position.

Like Experiment I, the present experiment is not sensitive enough to specify the degree of structural correspondence between the preparatory image and its perceptual counterpart. However, the flatness of the  $RT_2$  function and the reasonably low error rates indicate that the preparatory image must have embodied a sufficient degree of the spatial structure of the corresponding external stimulus to be used for rapid and accurate matching against that stimulus, when it physically appeared.

The primary objective of the present experiment was to examine the relationship between pre-stimulus and post-stimulus mental rotation. While both processes consist of the transformation of a mental image, they differ with respect to the degree of external support. In the case of preparatory rotation, the transformed representation is an internally maintained image of a visual object which is no longer externally present. In the case of post-stimulus rotation, the internal representation is extracted from a visual stimulus which remains externally present during the entirety of the subsequent rotational process.

For the group data and for six of the individual *Ss*, the estimated rate of preparatory rotation was somewhat slower than the rate of post-stimulus rotation. Nevertheless, the difference, even when present, was generally quite small for both the group data (cf., Fig. 5) and for the individual *Ss* (cf., Fig. 6). Indeed, the group data and over half of the *Ss* showed no statistically significant differences between pre-stimulus and post-stimulus rotation rates. These results thus provide more direct confirmation of the earlier, indirectly-derived suggestion of Cooper and Shepard (1973b) that the external (but nonrotating) presence of the object that is being mentally rotated has little effect on the rate of that mental rotation.

The data plotted in Fig. 5 reveal a difference of 325 msec between the

intercept of the post-stimulus rotation RT function (Experiment I) and the intercept of the discriminative function (Experiment II). This may seem surprising at first, for in both cases the test stimulus—a form presented in the trained orientation—and the required response—determination of the version of the upright form—are identical. However, this intercept difference may relate to an information-processing model which proposes that, before post-stimulus rotation can be initiated, Ss must first determine the identity and the orientation of the tilted test stimulus (cf., Cooper & Shepard, 1973b).

In the present experiment, Ss were provided with advance information as to *both* the identity and the orientation of the upcoming test form prior to the choice response. However, in Experiment I, *neither* sort of information was provided in advance. It is tempting to suppose that the 325-msec difference between the intercept in these two situations reflects the time needed to determine the identity and the orientation of the test form when given no advance information (Experiment I).

### SUMMARY AND CONCLUDING REMARKS

In this final section, the central findings from the experiments reported above are reviewed, and general implications of these findings are discussed.

(a) RT for determining whether a rotated test form is standard or reflected in version increases linearly with the angular departure of that form from a previously trained orientation (Experiment I). This linearity provides evidence that Ss mentally rotate an internal representation of the test form into the trained orientation in order to determine whether the presented version is standard or reflected.

(b) "Standard" responses are faster than "reflected" responses for all test-stimulus orientations (Experiments I and II). This constant RT difference suggests that Ss initially compare the mentally transformed representation of the test form with a memory image of the *standard* version of that form at the trained orientation.

(c) The intercepts and slopes of the RT functions for each of the eight random forms are approximately equal (Experiment I). This unexpected finding suggests that, regardless of the perceptual complexity of the test form, mental rotation is carried out at a constant rate of some 460°/sec (for these particular Ss).

(d) When Ss are given advance information concerning the identity and the orientation of an upcoming test form, the time required to prepare for the onset of that test form is a linearly increasing function of the angular departure of the predesignated orientation from the previously learned position (Experiment II). Furthermore, this linear function extends from 0° to 300° of clockwise or counterclockwise departure,

depending upon which direction was specified in advance. This finding indicates that *Ss* prepare for the upcoming test form by mentally rotating an image of that form, in the specified direction, into the predesignated orientation even though that orientation sometimes indicates a preparatory rotation of more than 180°.

(e) Following a preparatory mental rotation, RT for determining whether a visually-presented test form is a standard or a reflected version is rapid and constant, regardless of the orientation of the test form (Experiment II). The flat RT function demonstrates that *Ss* are prepared for the test-form onset, in the sense of having an appropriate internal representation pre-rotated into the designated position, when they so indicate.

(f) The slope of the post-stimulus RT function (Experiment I) is somewhat smaller than the slope of the pre-stimulus RT function (Experiment II); however, they do not differ significantly. This similarity in slope indicates that both sorts of mental rotation are carried out at comparable rates.

In the experiments reported above, discussions of the processes involved in preparing for and responding to disoriented test stimuli have made reference to the "rotation of a mental image." The central assumption has been that the internal processes and representations underlying performance in these tasks are structurally analogous to the external operations and objects to which they correspond. Recently, a fundamentally different view of the nature of these mental operations and representations has emerged (cf., Pylyshyn, 1973). These alternative accounts claim that the internal representations and processes involved in tasks such as those studied here bear no structural resemblance to their corresponding external objects and operations.

An example of one such model for the representations and processes underlying "mental rotation" tasks has recently been proposed by Levin (1973). In this model, visual objects are represented propositionally in the form of networks, and orientation is represented in the network in the form of a reference point and position predicates which relate the orientation of subparts of the representation to the reference point. Rotation is accomplished by successively changing these orientation predicates. This model (and other models of the same variety) makes a clear and interesting prediction: The amount of time required "to rotate mentally" an internal representation of a visual object should increase not only with the amount of disorientation of the object, but also with its visual complexity.

The results of Experiment I above indicate that rotation rate does *not* depend upon complexity for random forms within the range of six to 24 points. Unfortunately, this experiment does not rule out the possibility

that the internal representations of the test forms were reduced or highly schematic in nature. One way to test between these alternatives is to require Ss to distinguish between standard versions of random forms and test probes which differ from the standard in fairly subtle aspects (rather than merely reflection). Such research might provide an indication of the degree of structural detail embodied in a transformed internal representation, as well as resolve the question of the relationship between complexity and rotation rate.

## REFERENCES

- ATTNEAVE, F. Physical determinants of the judged complexity of shapes. *Journal of Experimental Psychology*, 1957, **53**, 221-227.
- ATTNEAVE, F. & ARNOULT, M. D. The quantitative study of shape and pattern perception. *Psychological Bulletin*, 1956, **53**, 452-471.
- BOWER, G. Mental imagery and associative learning. In L. Gregg (Ed.), *Cognition in Learning and Memory*, New York: John Wiley & Sons, 1972.
- BROOKS, L. R. Spatial and verbal components of the act of recall. *Canadian Journal of Psychology*, 1968, **22**, 349-368.
- CLARK, H. H. The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 1973, **12**, 335-359.
- CLARK, H. H. & CHASE, W. G. On the process of comparing sentences against pictures. *Cognitive Psychology*, 1972, **3**, 472-517.
- COOPER, L. A. Internal representation and transformation of random shapes: A chronometric analysis. Doctoral dissertation, Stanford University, 1973.
- COOPER, L. A. & SHEPARD, R. N. The time required to prepare for a rotated stimulus. *Memory and Cognition*, 1973a, **1**, 246-250.
- 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, 1973b.
- EGETH, H. & BLECKER, D. Differential effects of familiarity on judgments of sameness and difference. *Perception and Psychophysics*, 1971, **9**, 321-326.
- LEVIN, J. A. Network representation and rotation of letters. Unpublished manuscript, University of California, San Diego, September, 1973.
- METZLER, J. & SHEPARD, R. N. Transformational studies of the internal representation of three-dimensional objects. In R. L. Solso (Ed.), *Theories of Cognitive Psychology: The Loyola Symposium*. Potomac, MD: Lawrence Erlbaum Associates, 1974.
- PAIVIO, A. *Imagery and Verbal Processes*. New York: Holt, Rinehart & Winston, Inc., 1971.
- POSNER, M. I. Abstraction and the process of recognition. In G. Bower & J. T. Spence (Eds.), *The Psychology of Learning and Motivation*. Vol. 5 New York: Academic Press, 1969. Pp. 44-100.
- POSNER, M. I., BOIES, S. J., EICHELMAN, W. H., & TAYLOR, R. L. Retention of visual and name codes of single letters. *Journal of Experimental Psychology Monograph*, 1969, **79**, 1-16.
- PYLYSHYN, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 1973, **80**, 1-24.
- SEGAL, S. J. (Ed.) *Imagery*. New York: Academic Press, 1971.
- SEGAL, S. J. & FUSELLA, V. Influence of imaged pictures and sounds on detection of

- visual and auditory signals. *Journal of Experimental Psychology*, 1970, **83**, 458-464.
- SEGAL, S. J., & GORDON, P. E. The Perky effect revisited: Paradoxical thresholds or signal detection error? *Perceptual and Motor Skills*, 1969, **28**, 791-797.
- SHEPARD, R. N. Form, formation, and transformation of internal representations. In R. L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1975.
- SHEPARD, R. N. & FENG, C. A chronometric study of mental paper folding. *Cognitive Psychology*, 1972, **3**, 228-243.
- SHEPARD, R. N. & METZLER, J. Mental rotation of three-dimensional objects. *Science*, 1971, 701-703.
- TRABASSO, T., ROLLINS, H., & SHAUGHNESSY, E. Storage and verification stages in processing concepts. *Cognitive Psychology*, 1971, **2**, 239-289.
- VANDERPLAS, J. M. & GARVIN, E. A. The association value of random shapes. *Journal of Experimental Psychology*, 1959, **57**, 147-154.

(Accepted August 17, 1974)