4 Chronometric Studies of the Rotation of Mental Images

Originally authored by Lynn A. Cooper and Roger N. Shepard

The experimental paradigm introduced in this chapter differs from that used in the original experiment on mental rotation (described in the preceding Chapter 3). Instead of determining how the time that subjects need to compare two simultaneously presented objects depends on the angular difference between their two orientations, we determine how the time that subjects take either to prepare for or to respond to a single object (alphanumeric character) depends on its angular departure from the orientation in which that object is expected —whether its conventionally defined upright orientation or some other, prespecified orientation. The findings indicate (a) that discrimination between standard and reflected versions of a rotated character requires a compensating mental rotation, (b) that subjects who are given advanced information as to the orientation of a prespecified test stimulus can carry out the required mental rotation before that stimulus is actually presented, and (c) that during such a preparatory process, the orientation in which the test stimulus would (if presented) be most rapidly discriminated, is actually rotating with respect to external space.

Accumulating evidence indicates that to be more prepared for a stimulus is to have, in advance, a more appropriate internal representation of that stimulus. In a particularly relevant line of work, Posner and his associates have developed successful paradigms for determining the form of the internal representation remaining from a previously presented stimulus by measuring the time subjects take to respond discriminatively to an ensuing, related stimulus (e.g., Posner, Boies, Eichelman, & Taylor, 1969). They have shown that when subjects are instructed to indicate whether the second of two successively presented letters has the same name as the first, their response "same" is approximately 80 to 100 msec faster

Reprinted by permission from Chapter 3 of the volume *Visual information processing,* edited by W. G. Chase and published by Academic Press, 1973.[1]

when the two letters are physically identical ("R" and "R") than when they are identical only in name ("R" and "r"). The notion here is that a subject whose internal representation in short-term memory is of the most appropriate form (e.g., a "visual code" of the same internal structure as the ensuing visual stimulus) can respond very rapidly by matching this internal representation against that ensuing stimulus by some relatively direct, template-like process. When, however, the memory representation is of a less appropriate form (e.g., a visual code of a different structure—lower as opposed to upper case—or an auditory-articulatory code of the name of that letter), additional time is needed to access the name of the ensuing stimulus and then to test for a match between the two derived (but case-invariant) names.

Additional evidence that a visual representation mediates physical-identity matching lies in the disappearance of the superiority of the physical-identity match when the interval between the two letters reaches about two seconds. Presumably the visual representation of the first letter has faded during intervals of this length. Posner et al. (1969) also report that when subjects are motivated to attend specifically to the visual aspects of the first stimulus (i.e., when subjects always know what the case—upper or lower—of the second stimulus will be), then the speed of the physical-identity match, relative to the name-identity-only match, is maintained over longer interstimulus intervals. Finally, Posner and his associates have presented evidence that visual codes can be generated in the absence of an external visual stimulus. If subjects are given the name of the first letter in auditory form only, some 750 msec prior to the onset of the second stimulus, and if the case of this second letter is known in advance, then reaction times are as fast as those obtained for visual-visual matches of physical identity. The approximately .75-sec lead time that seems to be required for this is, presumably, the time it takes a subject to construct an internal visual representation of the named letter.

Mental Transformations

Reaction-time experiments of the sort reported by Posner and his associates appear to furnish rather strong evidence concerning the nature of particular internal representations—specifically whether they are principally visual or verbal in form. However, the question still remains whether these particular internal representations or "codes" are what we ordinarily refer to as mental images. The implied contrast, here, is with the possibility that the "visual code" postulated by Posner consists (as he himself has suggested) solely in the priming of certain relevant feature detectors in the sensory receptor system. The resulting state of heightened readiness of the receptor system for certain specific patterns could account for the demonstrated reduction in reaction times to just those patterns. However, this selective priming of lower-order feature detectors would not in itself constitute what we ordinarily refer to as a mental image, for, by hypothesis, the state of readiness would not have any cognitive consequences for the subject in the absence of the subsequent presentation of a related external stimulus. (Consider, for example, that one's perceptual system is more tuned to register the appearance of a familiar than an unfamiliar word without one's in any sense having a prior mental image of the more familiar word.) Presumably, to have a mental image, then, is to activate an internal representation that—in addition to preparing one for a specific external stimulus—can be used as a basis for further information processing even if the relevant stimulus is never actually presented. Such further information processing could include, for example, the generation of a verbal description of the mental image.[2]

The experiments that we wish to describe here follow Posner in the use of the selective reduction of reaction times to an ensuing visual stimulus for purposes of demonstrating a structural correspondence between (a) the internal representation with which the subject attempts to prepare for an upcoming stimulus and (b) the external stimulus itself. In addition, we introduce the new requirement that, in order to be fully prepared for the anticipated stimulus, the subject must first perform a transformation on the internal representation—specifically, a transformation that corresponds to a rigid rotation of the stimulus in space. This addition serves two purposes. First, by demonstrating that the subject can perform such a transformation on an internal representation, we establish that this representation is accessible to that subject for further cognitive processing. The representation then satisfies the important condition just set forth for its classification as a mental image. Second, by requiring a transformation that corresponds specifically to a spatial rotation, we provide further support for the claim that the representation or "image" is primarily visual or at least spatial in form.

Evidence that these internal representations are spatial in nature comes, also, from the postexperimental introspective reports. Our subjects typically claim that in preparing for the anticipated presentation of a rotated stimulus, they did in fact (a) form a mental picture of the anticipated stimulus and then (b) carry out a mental rotation of that picture into its anticipated orientation. Their tendency to generate such a verbal report is consistent with the supposition that the internal representation was accessible to introspection—as we should require of a mental image.

Of course, the verbalization of introspections need not be confined to reporting merely the existence and principal modality of a mental image. Reports dependent upon specific structural features of internal representations are of greater evidential value. To illustrate, in an experiment by Shepard and Feng (see Shepard, 1975) times were measured for subjects to report the identity of the letter that results when a specified spatial transformation is applied to a letter that is designated only by name. The subjects could readily report, for example, that the letter "N" turns into the letter "Z" when rotated 90°. Evidently, the internal representation they were manipulating had a definite internal structure that was analogous to the structure of the corresponding physical letter and that was internally available for further processing—including spatial transformation, visual analysis, and verbal report. Moreover, reaction times were consistently longer for more extensive transformations (e.g., longer for 180° than for 90° rotations), providing additional support for the notion that the images and operations were of a basically spatial character.

In the experiments that we shall be describing, mental transformations and the selective reduction of reaction times are used, jointly, to establish that the internal representations and mental operations upon these representations are to some degree analogous or structurally isomorphic to corresponding objects and spatial transformations in the external world. In all of these experiments, each transformation consists simply of single rigid rotation of a visual object about a fixed axis. In order to make the discrimination more demanding, hence to force subjects to carry out a mental rotation, we adopted the technique (introduced by Shepard and Metzler, 1971) of requiring the subject to discriminate between a stimulus and its mirror image—not merely between one stimulus and an entirely different stimulus.

EXPERIMENT I. DETERMINATION OF THE TIMES REQUIRED TO PREPARE FOR AND RESPOND TO A ROTATED STIMULUS

In the experiment we report now, we controlled the time during which advance information about orientation was available to the subject prior to the presentation of the rotated test stimulus itself. If subjects do indeed carry out some sort of mental rotation in the process of preparing for a tilted stimulus, this process should require more time for its completion, since the orientation indicated in the advance information departs by larger angles from the standard upright orientation. Moreover, failure to complete this process of preparation prior to the onset of the ensuing stimulus should result in an increase in the reaction time to that stimulus, since in this case some further mental rotation will have to be carried out *after* the onset of the tilted stimulus itself. Thus by determining how reaction time depends both upon the angle of the tilted stimulus and upon the duration of the advance information as to that angle, we hoped (a) to obtain somewhat more direct evidence that the generation and rotation of a mental image is in fact a part of the process of preparation for a rotated stimulus, and (b) to determine something about the time required to carry out this preparatory mental rotation.

Method

Subjects. Eight subjects—seven Stanford students and one of the authors (RNS)—were run under all experimental conditions. The first four subjects were run in the complete factorial design, which required about seven hours of participation from each subject. The second four were run in a half-replicate design and served for three to four hours each. (Although five male and three female subjects were included, no consistent differences were observed in the performances of the two sexes.)

Stimuli. The stimuli were all asymmetrical alphanumeric characters, specifically the three upper-case letters (R, J, G) and the three arabic numerals (2, 5, 7) exhibited in Figure 4.1. Each of these

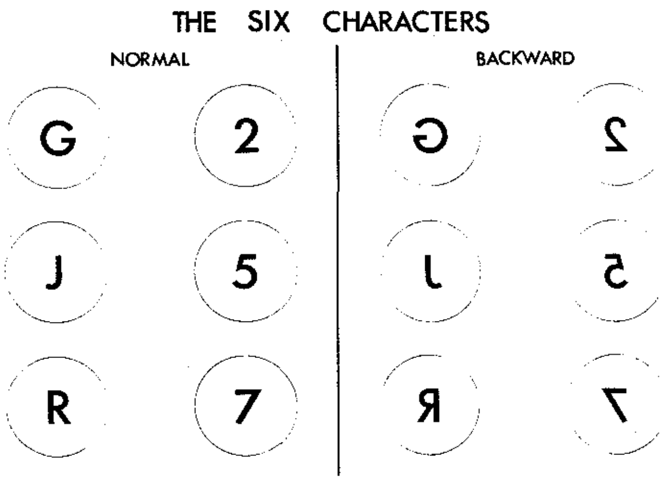

Figure 4.1 Normal and backward versions of the six alphanumeric characters used as test stimuli in Experiment I.

six characters appeared in each of six equally spaced orientations around the circle (in 60° steps starting from the standard upright position, 0°) as illustrated for the letter "R" in Figure 4.2. Since subjects were familiarized with both the set of six characters and

the set of six orientations, and since each occurred the same number of times in the test stimuli, the informational uncertainties concerning identity and orientation were equivalent in the absence of advance information. The subjects' task was simply to discriminate the normal versions of the characters (left-hand panels in the figures) from the reflected or backward versions of those same characters (right-hand panels) regardless of their orientations within the picture plane (Figure 4.2).

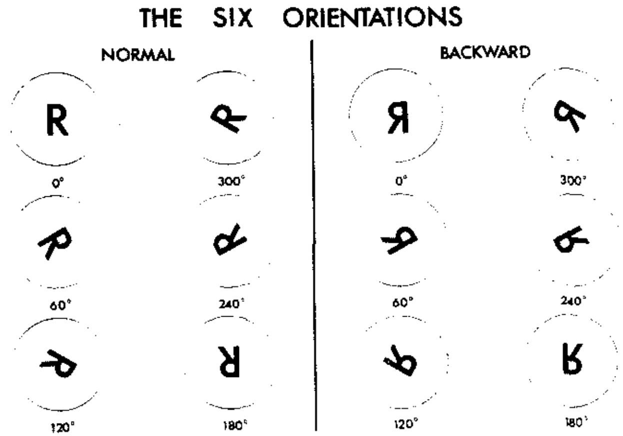

Figure 4.2 Normal and backward versions of one of the six characters, illustrating the six orientations in which it might appear as a test stimulus.

Following Shepard and Metzler (1971), we hoped that by requiring subjects to discriminate between mirror images of the same objects, we would prevent them from responding merely on the basis of some simple distinctive feature (such as the presence of an enclosed region in the case of the letter "R"), and thereby force them to carry out a "mental rotation" in order to compare a tilted character with the normal upright representation preserved in long-term memory. Notice, for example, that any one of the characters displayed in Figure 4.2 can almost immediately be identified as *some* version of the letter "R." In the cases in which that character is markedly tipped, however, it seems to take some additional time to determine whether that letter is normal or backward. Typically, subjects report that they do in fact imagine a markedly tilted character rotated back into its upright orientation to determine whether it is normal or backward, but that this is unnecessary for merely determining its identity. Indeed we suggest that subjects may have to identify a character before they can determine which is the top of the character and thereby know how the character must be rotated to bring it into its upright orientation.

The advance information cues, when presented, appeared centered within the same circular aperture as the subsequently ensuing test stimulus. The identity cue was displayed in the form of an outline drawing of the normal, upright version of the upcoming test stimulus, and the orientation cue appeared as an arrow passing through the center of the circular field and pointing in the direction at which the top of the test stimulus would appear. Figure 4.3, which shows the sequence of visual displays that would appear within the circular aperture on an illustrative trial, provides a more concrete idea of the appearance of the identity and orientation cues.

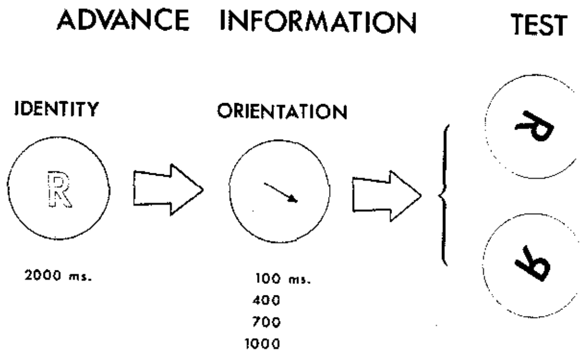

Figure 4.3 Sequence of visual displays appearing within the circular aperture on a trial of type B, in which *both* identity and orientation information were presented in advance of the test stimulus (illustrated here at 120 degrees).

The alphanumeric characters, which appeared both as test stimuli and as advanced information cues, subtended a visual angle of about 1½°. The visual angle of the circular aperture within which these characters appeared was 4°, and the luminance levels of the two or three fields that succeeded one another within this aperture (depending upon the condition) were all approximately 20 foot-Lamberts.

Structure of individual trials. The subject sat in a dimly illuminated room with head pressed against the shaped rubber light-shield surrounding the viewing window of the tachistoscope. This permitted binocular viewing of all stimuli, but prevented physical rotation of the head. Following a warning signal at the beginning of each trial, the subject fixated the circular field where the test stimulus and the advance information cues (if any) were about to appear, with left and right thumbs positioned on the two response buttons

located on a hand-held box. The subject always used the preferred thumb (i.e., the right thumb except in the case of our one left-handed subject) to register a decision that the test stimulus was normal, and the nonpreferred thumb to signal that the stimulus was backward. The two cases, normal and backward, occurred equally often according to a random sequence. The test stimulus always remained on until after the subject's response.

Each subject ran in eight different conditions of type and duration of advance information. Of central concern are the four variable time conditions (labeled "B" in Figures 4.3 and 4.4) in which *both* identity and orientation information were supplied. On these trials, the identity cue was displayed for 2000 msec, immediately followed by the orientation cue, which persisted for 100, 400, 700, or 1000 msec (depending upon which of the four conditions of type B was in effect). The orientation cue was then immediately replaced by the actual test stimulus. As indicated in Figure 4.3, even after having been provided with advance information about both the identity and orientation of the ensuing test stimulus, the subject still had to await the actual presentation of that stimulus in order to determine whether it was the normal or the backward version of that character at that orientation.

Figure 4.4 schematically illustrates the other four conditions, along with the conditions of type B (described above), for the case in which the test stimulus was to appear at 120°. The remaining four conditions were as follows: N, in which *no* advance information was provided (but only a 2000 msec blank warning and adaptation field); I, in which only *identity* information was supplied; O, in which only *orientation* information was furnished; and finally C, in which the identity and orientation information were presented in a *combined* form followed by a 1000 msec blank field before the onset of the test stimulus. The purpose of interposing the blank field in this last condition, C, was to ensure that the response to the test stimulus was based upon comparison with a representation in memory and not upon a purely sensory discrimination of continuity or change in the outline of the external visual display (for normal or backward test stimuli, respectively). For all conditions illustrated in Figure 4.4 the large unfilled arrows signify *immediate* replacement, upon the offset of one visual display, of the display shown just to its right (always within the same circular aperture).

Conditions I and C provided two reference points with which to compare the four variable time conditions B. At one extreme, when the duration of the orientation information is made very short (as in the B-condition with only 100 msec), we should expect that the subjects' reaction times to the test stimuli would approximate

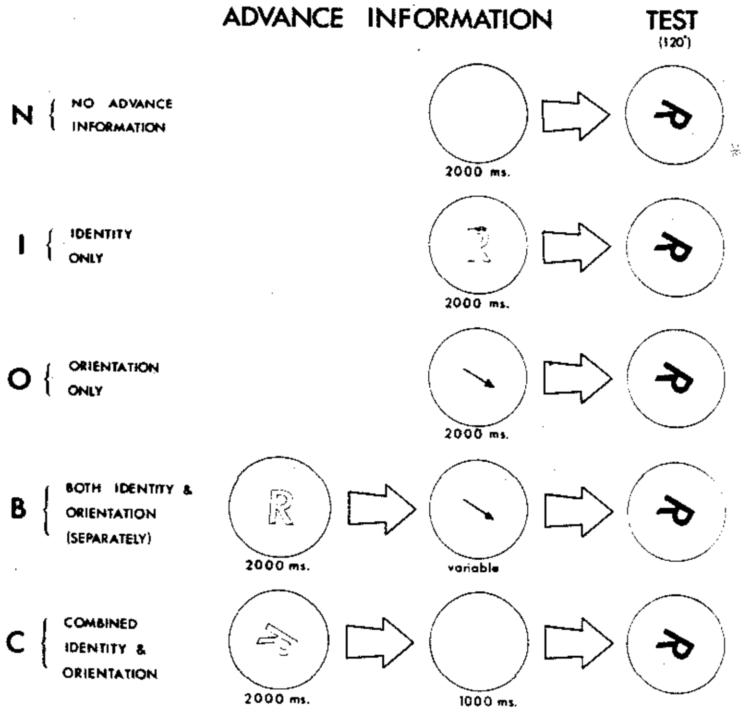

Figure 4.4 Schematic illustration of the five basically different types of conditions, N, I, O, B, and C. (Since type B subsumes four conditions, with different durations of orientation information specified in Figure 4.3, the total number of distinct conditions is eight.)

their reaction times to those same test stimuli when no advance information as to orientation has been provided (as in the I-condition). At the other extreme, when the duration of the orientation information is made sufficiently long (as in the B-condition with 1000 msec), subjects may have time to generate an appropriate mental template of the normal version of that character and to rotate it into the designated orientation. If so, their reaction times to the ensuing test stimulus should approximate those obtained when such a rotated template is supplied visually (as in the C-condition), hence does not have to be subjected to any mental rotation before comparison.

Overall experimental design. Individual trials were blocked by condition, with 12 trials to a block. At the beginning of each such block the subject was given explicit instructions concerning the nature and duration of the advance information to be provided on all trials within that block, and was then given practice trials

of that type until ready to proceed with the actual trials of the block. The order of trials within blocks was randomized subject to the constraint that each of the six orientations occurred twice within each block. Hence, although the subject knew whether there would be advance information and how long it would last, until that advance information (if any) was actually presented on a given trial, the subject did not know which of the six characters would come up next or in which of the six orientations it would appear.

The complete factorial design (used for the first four subjects) required the completion of 576 trials per subject in order to obtain one observation for each cell of the design. Each of these subjects was run for six one-hour sessions consisting of eight blocks (one for each of the eight different conditions) of 12 trials each. The order of conditions was counterbalanced over sessions. Prior to these six sessions, each subject was given an initial practice session to ensure familiarity with the stimuli, the experimental procedure, and the various conditions of advance information.

After all data had been collected from the first four subjects, we found that the mean reaction times in which we were interested were virtually unchanged when we recomputed them on the basis of only half of the observations selected from the entire factorial design by means of a "checkerboard" half-replicate design. Accordingly, the remaining four subjects were run only on the trials specified by this half-replicate design. After the initial practice session, therefore, each of these subjects completed only three one-hour sessions of 96 trials each, yielding a total of 288 observations per subject.

Subjects were instructed, for all conditions, to indicate whether the test stimulus was normal or backward (regardless of its orientation in the picture plane) as rapidly as they could, without making errors, by pressing the appropriate button on the response box. Although error rates for the different conditions and orientations were positively correlated with mean reaction times, error rates averaged over all conditions were uniformly quite low, ranging from 3.6 to 8.7 percent for individual subjects. Nevertheless, throughout the experiment all trials on which errors were made were later repeated until an errorless reaction time had been obtained from each subject for each combination of character, orientation, version (normal or backward), and condition called for by the factorial design (or its half-replicate variant).

Reaction-time Results

The effect of orientation of the test stimulus. First we consider the condition, N, in which the subject was given no advance information concerning the identity or orientation of the upcoming test stimulus. The mean reaction times for this condition (averaged over all correct responses to either the normal or the backward version of the test stimulus) are plotted as the uppermost curve in Figure 4.5. The independent variable, here, is the orientation of the test stimulus as specified in degrees of clockwise rotation from the standard upright orientation of the character. (In this and subsequent plots of this type, all points are independent except the points at 360° which merely duplicate the points at 0°.)

In the absence of any advance information concerning the upcoming stimulus, reaction time increases very markedly as the orientation of that stimulus departs from its standard upright orientation. Indeed, as we move from 0° to 180° there is a roughly twofold increase in mean reaction time, from between 500 and 600 msec at the upright orientation to nearly 1100 msec at the completely inverted orientation. From the symmetry of the curve we see that the increase in reaction time resulting from a given angle of tilt is the same for both clockwise and counterclockwise rotations. This increase is not strictly linear, however, but concave upward, with the sharpest increase occurring as we approach the completely inverted orientation of 180° from 60° away on either side (i.e., from the orientations of either 120° or 240°).

The reliability of the shape of this curve is indicated by its highly symmetric form as well as by the highly similar shapes of the two reaction-time curves plotted in Figure 4.5 just below the curve for condition N—namely, the curves for the conditions I and O, in which the subjects were given either identity information or orientation information only.

Despite the nonlinearity of these functions, we take the very marked increase in reaction time with departure of the test stimulus from its standard upright orientation to be supportive of the notion that the subject carries out some sort of a mental rotation. In particular, we suggest (a) that, in order to compare a markedly tilted character with the representation of the normal version of that character in long-term memory, the subject must first imagine the tilted character rotated into its upright orientation, and (b) that the greater this tilt, the longer it will take to complete the corrective rotation. Reasons for the nonlinearity of the increase in reaction time that are consistent with this notion of mental rotation will be presented in the theoretical discussion.

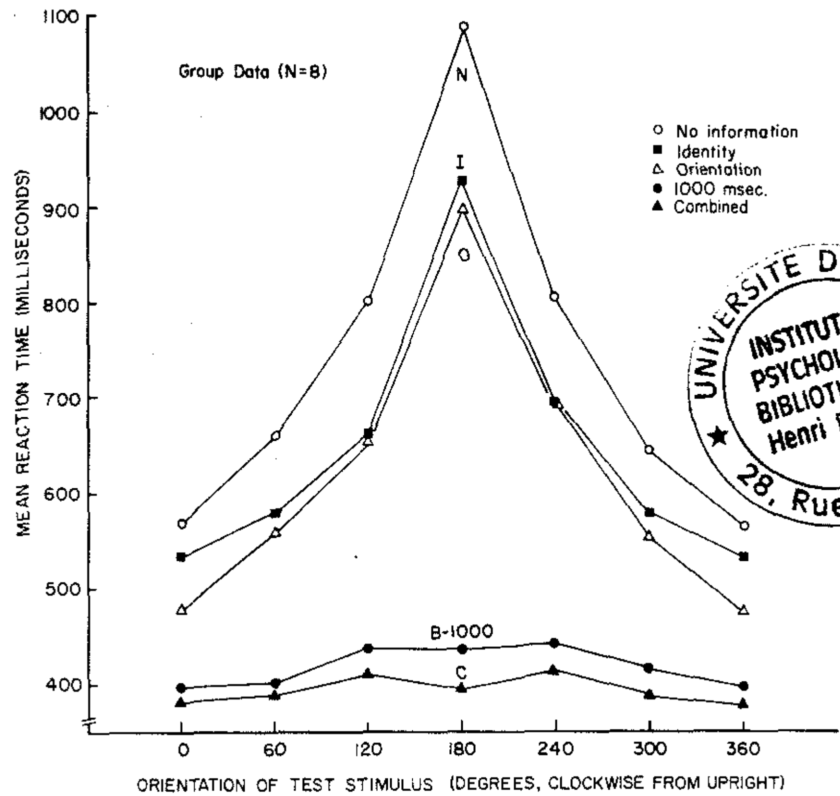

Figure 4.5 Mean reaction time as a function of orientation of the test stimulus for those conditions in which advance information, if presented at all, persisted for the maximum duration.

The effect of advance information as to identity and/or orientation. We turn, now, to a comparison among all five conditions in which the subjects were given adequate time to take full advantage of whatever advance information (if any) was provided; namely, conditions N, I, O, C, and the one B-condition in which the orientation cue persisted for the full 1000 msec. The five different curves plotted in Figure 4.5 exhibit the dependence of reaction time on orientation of the test stimulus for these five conditions.

These results show that the reaction-time curves for the two conditions (I and O) with advance information as to identity or orientation only—though somewhat lower than the corresponding curve for the condition (N) with no advance information—are nevertheless relatively close to it in height and, particularly, in overall shape. Moreover, and of central importance, they establish that the reaction-time curve for the condition (B), in which both identity and orientation information were separately presented, is dramatically lower than the other three curves and virtually flat.

Another aspect of the present results, not shown in Figure 4.5, is that the response used to signal that the test stimulus was the normal version of that character was consistently faster than the response used to signal that it was the backward version of that character—by a difference that was essentially constant over all conditions and orientations and that ranged from roughly 10 to 150 msec, depending on the particular subject. On the basis of other experiments in which the functions of the two hands have been systematically interchanged, it appears that the factor of overriding importance is the subject's choice of what to test for first (in this case, normalness or backwardness), rather than the particular hand that is then set to register a positive outcome of that test (cf. Clark & Chase, 1972; Trabasso, Rollins, & Shaughnessy, 1971).

The present results indicate three further things: First, relative to the condition in which no advance information is provided (N), the conditions in which either identity or orientation information is supplied (I or O) tend to produce reaction times that are shorter by a constant amount (roughly 100 msec), regardless of the orientation of the test stimulus. Second, the curve for the condition with both kinds of advance information presented separately (B) achieves a close approximation to a completely flat function. And third, by comparison with the new condition in which complete advance information was presented in combined form (C), we can now conclude that the internal representation that the subject constructs on the basis of separate information about identity and orientation (in condition B) is just about as efficient a mental template as a memory image of the rotated character itself (in condition C).

The effect of varying the duration of the orientation information. We turn now to a consideration of the remaining three B-conditions. In these conditions the duration of the advance information as to orientation was reduced (from the full 1000 msec) to values of only 700, 400, or 100 msec. The mean reaction times for these conditions (again averaged over all correct responses to both normal and backward stimuli) are plotted as a function of the orientation of the test stimulus in Figure 4.6. For purposes of comparison we also include the limiting reference or control conditions I and C already shown in Figure 4.5. As before these group curves are highly reliable and representative of the curves for individual subjects.

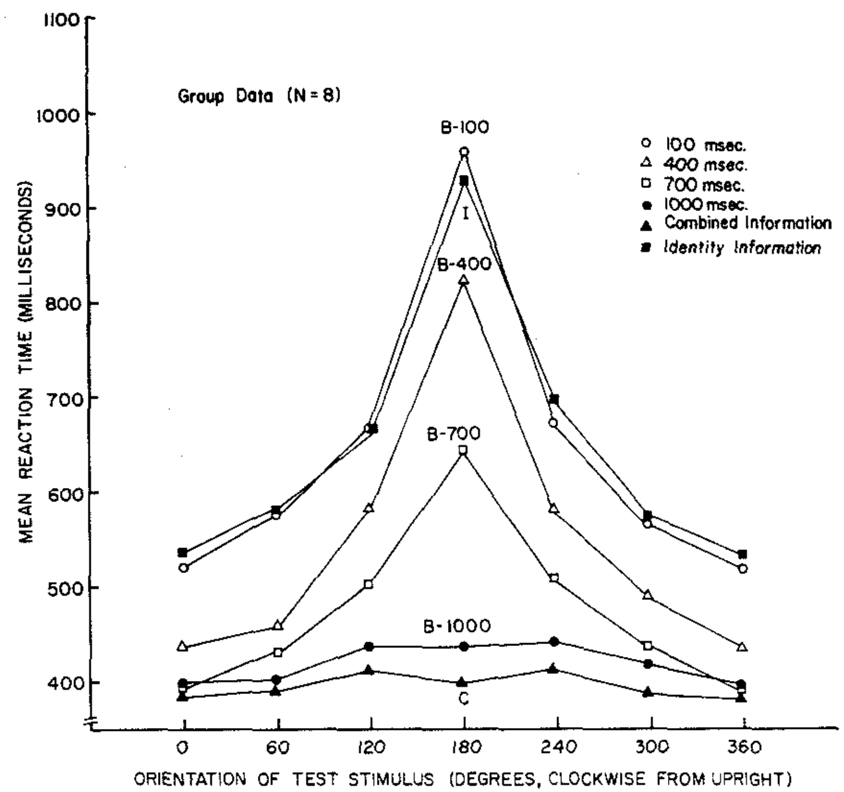

Figure 4.6 Mean reaction time as a function of orientation of the test stimulus for those conditions in which identity information was provided.

Comparisons among these curves enable us, for the first time, to make some quantitative inferences concerning the time that it takes to prepare for a stimulus that is about to appear in some rotated orientation. From the flatness of the function for the 1000-msec. B-condition, we know that this process of preparation can generally be completed within one second. However, when the duration of the orientation cue is reduced to 700 msec, a pronounced peak in the reaction-time function emerges at 180°, indicating that on the average the discriminative response to the ensuing test stimulus takes over 200 msec longer whenever that test stimulus appears in an inverted position. When the duration of the orientation cue is further shortened to 400 msec, the reaction times increase by another 200 msec at 180°, and also by some 80 msec at 60° on *either side of* 180°. *Finally, when this duration* is cut down to only 100 msec, the reaction times are essentially identical, at all orientations, to the reaction times when only identity information is provided (Condition I).

Following the experiment, the subjects themselves offered explanations for their reaction times under these B-conditions. Their explanations ran along the following lines: When the duration of the orientation cue was reduced below a second (e.g., to 700 msec), they were often unable to rotate their mental image of the anticipated stimulus around to 180° before that stimulus actually came on, although they usually were able to complete rotations of only 60° or even 120°. When the duration was further reduced (e.g., to 400 msec.), they were almost never able to get to 180° *before the onset of the test* stimulus and, now, often failed even to reach 120°. Finally, they reported that a duration of only 100 msec. was generally of no use at all, for by the time they were able to interpret the orientation cue, they had discovered that the test stimulus itself had already appeared.

Results for individual subjects. When we turn from the average reaction times for the group of eight subjects as a whole (Figures 4.5 and 4.6) to the corresponding reaction times for individual subjects, we immediately discover that there were stable and very substantial differences among subjects in their mean reaction times. However, these differences were very pronounced only for those conditions that tended to produce long reaction times; they all but disappeared for the conditions (B-1000 and C) in which complete advance information was furnished. Thus for the most difficult case in which no advance information preceded a completely inverted test stimulus (condition N at 180°), the mean reaction times varied over a more than twofold range, from just under 700 msec for the fastest subject to just over 1700 msec for the slowest. At the same time, though, the reaction times for the 1000-msec. B-condition (averaged over all orientations) ranged only from about 350 msec to a little under 500 msec for these same two subjects.

It appears that the average rate of mental rotation (which can be very roughly estimated as 180° divided by the difference between the reaction time at 0° and 180° under the N condition) varied from something like 800° per second for the fastest subject to something like 164° per second for the slowest. However, when subjects are already prepared with an appropriately oriented image of the upcoming stimulus (as in conditions B-1000 and C), these very different individual rates of mental rotation are not involved.

Consequently, in these conditions most subjects respond with approximately equal rapidity—within some 350 to 500 msec.[3]

If now we plot entire sets of reaction-time curves corresponding to those already displayed for the group of eight subjects as a whole (in Figures 4.5 and 4.6), we find that despite these enormous individual differences in average reaction time, the shapes and relational pattern of the curves are strikingly constant from subject to subject. Athough we have examined these curves for all eight subjects individually, it appears impractical to present them all here. Instead, we present complete sets of curves for just two representative subjects—namely, the one with the shortest and the one with the longest overall average reaction time. The patterns exhibited by these two appear to us to be typical of the patterns exhibited by the other, intermediate subjects.

The individual curves for these two extreme subjects are all displayed in Figure 4.7. The plots on the left, which correspond to earlier Figure 4.5, are for the conditions in which the advance information (if any) persisted for its maximum duration. The plots on the right, which correspond to the earlier Figure 4.6, are for all conditions in which identity information was provided (including those in which the advance information as to orientation was reduced in duration). The two upper plots are for the subject whose responses were, on the average, the quickest. The two lower plots are for the subject whose responses were, on the average, the slowest. Because the longest mean reaction times for this second subject were over twice as long as the longest mean reaction times for the first subject, the vertical scales in the two lower plots have been linearly compressed with respect to the vertical scales in the upper plots. Note, particularly, the extreme flatness of the curve produced by the faster subject under the B-1000 condition.

Various statistical analyses confirmed (a) that all eight subjects showed significant effects of duration, orientation, and interaction between duration and orientation, but (b) that there was no significant difference between the shapes of the flat reaction-time functions for conditions B-1000 and C or between the shapes of the peaked reaction-time functions for conditions N, I, and O.[4]

Distributions of reaction times under different conditions. So far we have been concerned with just the *means* of the distributions of reaction times for different conditions and orientations (whether for individual subjects or for the whole group). An examination of the entire distributions can provide additional information relevant to notions about what kinds of processes are going on within individual subjects. Although we have surveyed the computer-plotted distri-

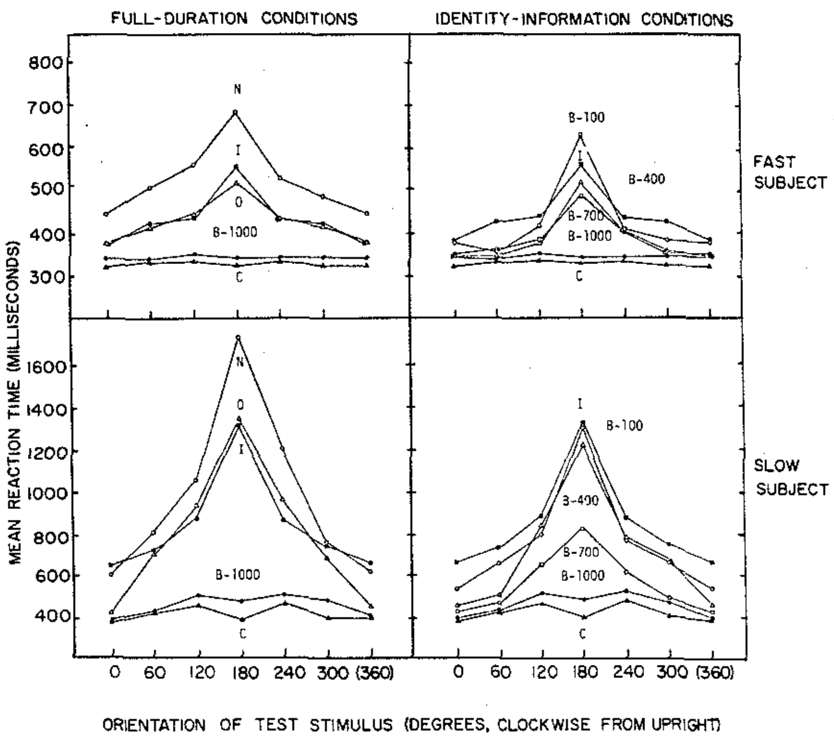

Figure 4.7 Mean reaction time as a function of orientation of the test stimulus for two individual subjects—the fastest subject (upper panels) and the slowest subject (lower panels). Left-hand panels correspond to the *group functions displayed in Figure 4.5, and right-hand panels correspond* to the group functions displayed in Figure 4.6.

butions for all eight subjects, under all eight conditions, at each of the six orientations, it is impractical to display all 384 of these individual distributions here.

Comparisons among the distributions obtained from different subjects indicated, however, that the eight subjects could be divided into a group of five subjects with relatively long reaction times and a group of three subjects with relatively short reaction times. Moreover, distributions plotted for either group as a whole then turned out to be reasonably representative of all subjects within that group. Among all 48 combinations of condition and orientation, the most informative cases appeared to be (a) those in which the test stimulus came on, essentially without advance information as to orientation, at each of the four degrees of departure from upright, 0°, 60°, 120°, and 180° (whether in the clockwise or counterclockwise direction); and (b) those in which the test stimulus came on at 180°, but following periods in which advance information as to this orientation had been presented for 100, 400, 700, or 1000 msec.

Distributions of the first sort are displayed in Figure 4.8. In order to obtain relatively stable shapes, each curve is based upon the pooled data for the two essentially equivalent conditions I and B-100 and for all (three or five) subjects within the indicated group (*fast or slow*). The distributions for the three "fast" subjects (shown by solid lines) are relatively compact and symmetrical. The distributions for the five "slow" subjects (shown by dashed lines) tend to be somewhat broader (particularly at 120°). For both groups of subjects, the distributions shift to the right and become broader as the test stimulus departs more and more from upright. This rightward shift is considerably more marked for the five slower subjects. Perhaps what most characterizes these slower subjects, then, is a slower speed of mental rotation.

The second set of distributions of interest includes those for reaction time to a completely inverted test stimulus following various durations of advance information as to the 180° orientation. These distributions are displayed in Figure 4.9, for the three fast subjects, and in Figure 4.10, for the five slow subjects. Again, these pooled distributions, though slightly broader, appeared to be quite representative of the distributions for individual subjects from each group. At the top of these figures we see that when the orientation information was available for a full second (B-1000), the reaction-time distribution was quite compact, sharply peaked and, indeed, very similar in shape to that obtained under Condition C (in which the preparatory image had already *been rotated for the* subject). At the bottom we see that when the orientation information was available for only a tenth of that time (B-100), the distribution was shifted markedly to the right, spread out, and similar in shape to that obtained under Condition I (in which there was no orientation information). As in earlier Figure 4.8, this shift to the right was much greater for the slower subjects—which we should expect if their longer reaction times were due primarily to a slower rate of mental rotation (in this case, of the preparatory image rather than of the test stimulus itself).

Here, however, the intermediate cases (B-400 and, for the five slow subjects, B-700 too) yield distributions that are more spread out than the distributions even for the extreme case B-100. We could explain this by supposing that *the rate of preparatory rotation* is somewhat variable from trial to trial, depending in part upon the particular character to be rotated. (In fact, most subjects reported

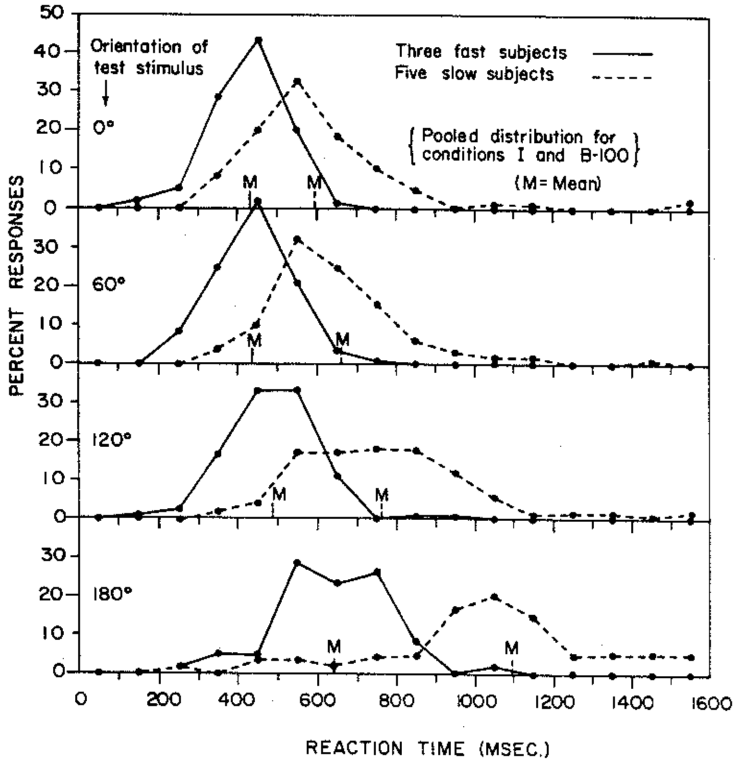

Figure 4.8 Distributions of reaction times to test stimuli presented at 0 degrees, 60 degrees, 120 degrees, and 180 degrees angular departures from the upright orientation. Distributions are pooled over conditions in which, effectively, no orientation information was provided (Conditions I and B-100). Separate distributions are plotted for the three fast subjects (solid lines) and the five slow subjects (dashed lines).

that some characters, e.g., R and 2, were generally easier than others, e.g., 7 and J.) Under the extreme conditions we should expect that on virtually all trials, the subjects either would be fully prepared (Condition B-1000) or would not be at all prepared (Condition B-100) for the inverted test stimulus. Consequently, their reaction times would be consistently short or long, respectively. Under the intermediate conditions, however, we should expect that on some proportion of the trials the subjects would be prepared and on some proportion they would not. When they were prepared, they would be able to respond rapidly (perhaps as rapidly as under Condition B-1000). When they were not prepared, though, their reaction times would be longer and more variable. In the latter case they might

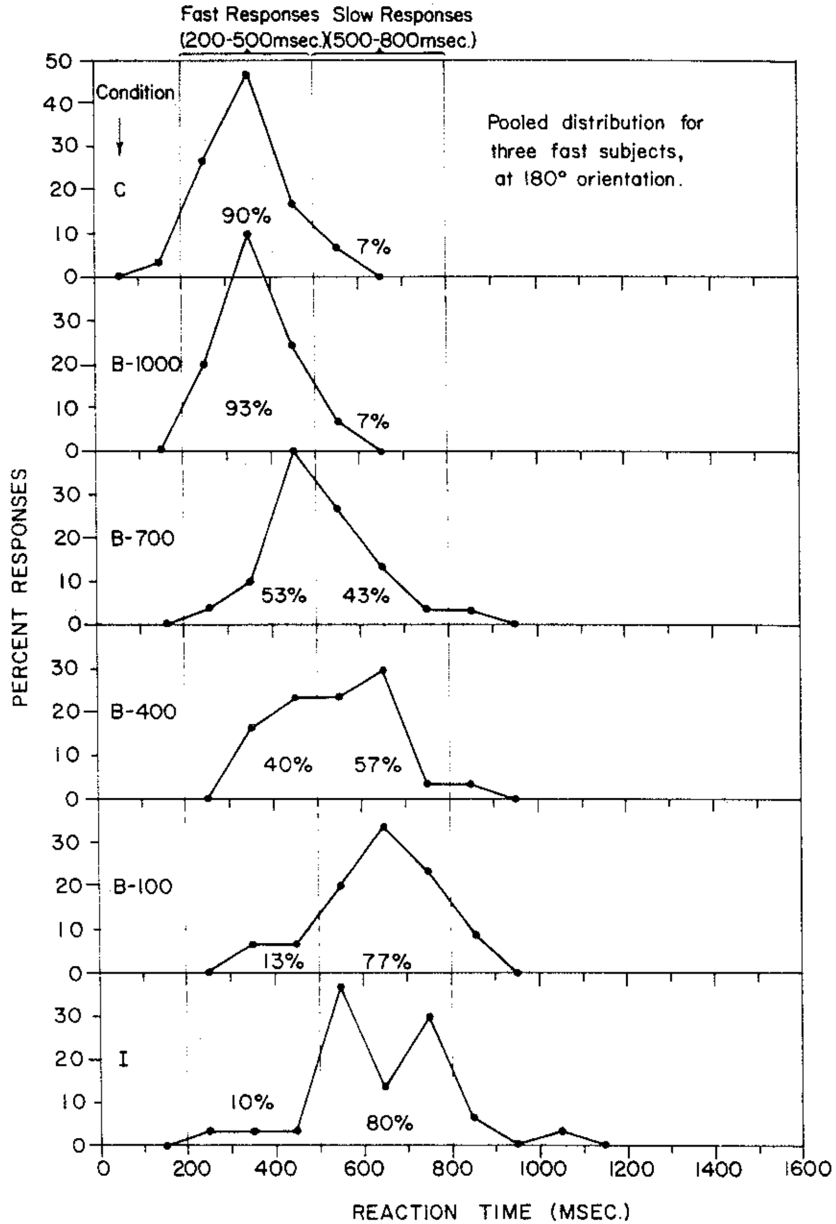

Figure 4.9 Distributions of reaction times to test stimuli presented at 180 degrees for the three fast subjects. Separate distributions are plotted for each condition in which identity information was provided.

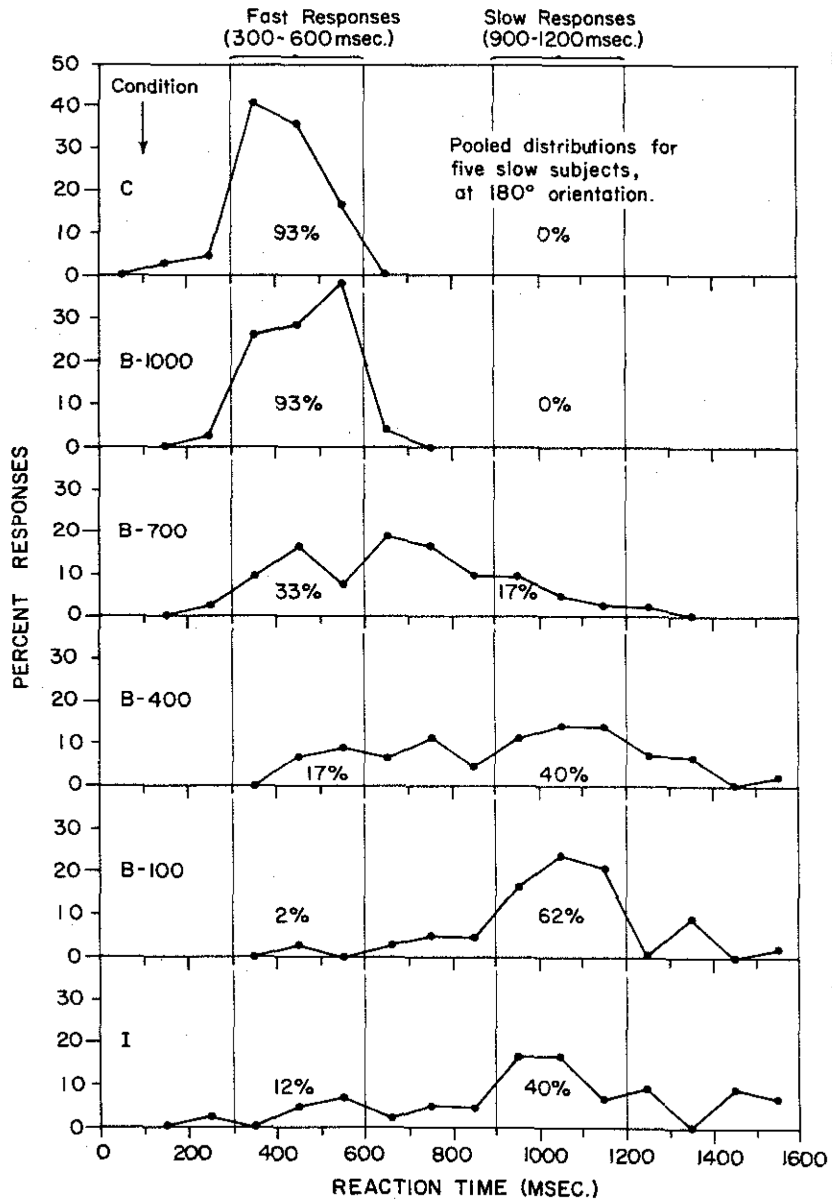

Figure 4.10 Distributions of reaction times to test stimuli presented at 180 degrees for the five slow subjects. Separate distributions are plotted for each condition in which identity information was provided.

drop their preparatory rotation and start all over with the test stimulus itself, in which case their reaction times would be comparable to their times under Condition B-100. Alternatively, they might continue their preparatory rotation into congruence with the test stimulus, in which case their reaction times could fall anywhere between their times for B-100 and for B-1000, depending on how far along they were with their preparatory rotation before the onset of the test stimulus.

It appears from Figure 4.9 that a fast subject typically responds in 200 to 500 msec when prepared for the inverted test stimulus and in 500 to 800 msec otherwise; and from Figure 4.10 that a slow subject usually responds in 300 to 600 msec when prepared and in 900 to 1200 msec when initiating rotation upon the presentation of the test stimulus itself. These two cases, labeled "fast responses" and "slow responses," are indicated in Figures 4.9 and 4.10 by the two vertical bands drawn there. Numbers have also been included to indicate, for each distribution, the percentage of its total area that falls within each of these two vertical bands. These numbers thus provide a very rough estimate of the percentage of trials of each type for which the subject was prepared for the inverted test stimulus either fully or not at all. For the five slow subjects, for example, we see that the estimate of the percentage of trials on which the subjects were fully prepared systematically declines from 93 when they had a full exposure to advance information as to orientation (in both Conditions C and B-1000), to only about 10 or less when they had little or no chance to take advantage of that advance information (in Conditions B-100 and I, respectively). The intermediate responses, which were made in 600 to 900 msec (e.g., in Condition B-700), may represent trials in which a subject, though not fully prepared at the onset of the test stimulus, was nevertheless able to continue the preparatory rotation until congruence was achieved. This would explain the reduction in overall processing time, as measured from the onset of the test stimulus.

Discussion

The nonlinear effect of orientation of the test stimulus. In the experiment by Shepard and Metzler (1971), the time subjects required to determine that two perspective pictures were of the same three-dimensional object increased in a remarkably linear manner with the angular difference in their portrayed orientations. This, together with the subjects' introspective reports, was taken to support the notion that they made each comparison by imagining one object rotated into congruence with the other and that this "mental rotation" could successfully be carried out no faster than about 60° per second in that task. If we wish to invoke a similar notion of mental rotation to account for the marked increase in reaction time to rotated alphanumeric characters, we need to explain why the increase in reaction time (shown in Figures 4.5 and 4.6) consistently departs from linearity.

At least four explanations appear consistent with the data. First, the representation or "template" for the normal, upright character in long-term memory may be so broadly tuned that this discrimination can often be made for characters tilted by 60° or even more without any need for mental rotation. Second, it may be that the rate at which an object can be mentally rotated increases with the familiarity of that object, and that letters and numbers are less familiar when viewed in more or less inverted orientations. Third, when the character was completely upside down, instead of imagining it rotated into its upright orientation within its own picture plane, subjects may sometimes have imagined it flipped rightside up about its horizontal axis. Then they either would imagine it flipped about its vertical axis (which is necessary to restore its original parity—normal or backward) or else would remember to reverse their response to the uprighted character (using the left thumb to indicate "normal" and the right thumb to indicate "backward"). An experiment by Shepard and Feng (see Shepard, 1975) has shown that it does take longer to imagine a letter flipped first about a horizontal and then a vertical axis than to imagine that same letter rotated 180° in the picture plane—even though the two ways of transforming the image result in the same final state. And fourth, the operations of determining the identity and/or the orientation of the test stimulus (which presumably must be completed before the rotation can even be started) may themselves increase nonlinearly, at least to some small extent, as the test stimulus departs from upright.

It is possible that any or all of the above explanations are to some degree correct for at least some subjects. Moreover, all four are specifically designed for the case of stimuli, such as alphanumeric characters, that have a well-defined standard or upright orientation. Hence, the fact that the reaction-time functions obtained by Shepard and Metzler were perfectly linear is compatible with these explanations (the three-dimensional nonsense shapes presented in that study had no uniquely established conventional or preferred orientation). We conclude, therefore, that the finding of a consistently nonlinear relation between reaction time and orientation of the test stimulus here does not weigh against the hypotheses that subjects typically used mental rotation to determine whether an inverted test stimulus was normal or backward.[5]

The nonadditive effects of advance information as to identity and orientation. A very central finding is that of a virtually flat reaction-time function under the 1000-msec B-condition (in which complete advance information is given, separately, for both identity and orientation). This finding offers substantial support for the notion that in preparing for the presentation of a tilted stimulus, the subject carries out a purely mental rotation of something that might be called a mental image of the anticipated stimulus. Further support comes from the subjects themselves who, after completing the experiment, typically claimed that in this B-condition—just as much as in the C-condition (in which the appropriate visual image was supplied to the subject already rotated)—they were ready with what seemed, introspectively, to be the same sort of mental image. Regardless of whether the orientation of this internal representation had been achieved by external, physical rotation (Condition C) or by purely internal, mental rotation (Condition B-1000), these subjects were equally ready to use this "rotated" internal representation as a template against which they could rapidly match the visually presented test stimulus when it then appeared in that same orientation.

What needs to be noted, however, is that such a mental rotation evidently cannot be carried out on the basis of just the orientation information alone. For, the curve for Condition O—far from being completely flat—has the same steep slope as the curves for the N, I, and 100-msec B-conditions, in which subjects were not given (or were not able to utilize) any orientation information.

This might seem surprising. To the extent that the subject has already been given information as to the orientation of a stimulus, it is natural to suppose that the redundant reappearance of the same information in the actual test stimulus should then have little effect. And this is, in fact, what has happened in the 1000-msec B-condition. There, as a result of having provided advance access to the orientation variable, that variable no longer had any effect on the latency of the subsequent response to the stimulus itself. Why is it, then, that advance access to the orientation variable did not in any way diminish the effectiveness of that same variable when it reemerged in the test stimulus in condition O?

The answer seems to be that, under the conditions of this experiment, subjects can only rotate the mental representation of a specific, concrete object or character. They evidently are not able to rotate a general, abstract frame of reference—at least when their head orientation is fixed with respect to gravitational upright (cf. Attneave & Olson, 1967; Rock & Heimer, 1957). In our experiment, when subjects were informed of both the character (e.g., the letter "R") and its orientation (e.g., at "4 o'clock"), they were able to imagine that character rotated into that orientation in advance of its actual presentation. When, however, they were informed only that the character would appear in, say, the 4 o'clock position without being told *which* character would appear in that position, they could only wait until the actual presentation of the character itself and then had to imagine it rotated into the upright orientation in order to determine whether it was normal or backward.

A consequence is that the effects of identity information and orientation information, though possibly quite additive when the stimulus appears in its upright orientation, become increasingly nonadditive as the orientation departs from upright. As is evident in Figure 4.5, each kind of information alone results in a roughly 100-msec drop in the reaction-time function as a whole, without having any appreciable effect on the shape of the function. When both kinds of information are provided, however, the reaction-time function becomes so much lower and flatter that there is a drop of at least 600 msec at 180°.

We are not saying that the advance presentation of orientation information alone has no effect on subsequent reaction time—only that it has no effect on the way in which reaction time depends upon the orientation of the ensuing test stimulus. The advance presentation of either identity or orientation information alone does have an effect—indeed, approximately the *same* effect—as is evident from the 100-msec decrease at all orientations. This is especially striking at 0° where the test stimulus appears in the standard upright position for all conditions. Consider, in particular, Conditions I and C at the 0° point. In both cases the subject sees essentially the same thing; namely, an upright outline of the appropriate character followed by the upright presentation of that same character (or its mirror image). Yet from Figure 4.5 we see that at 0°, mean reaction time was over 100 msec longer under Condition I than under Condition C; and this difference is statistically reliable.[4]

We think that the explanation for this difference does not lie in any difference in the physical displays under the two conditions. (That a one-second blank field is interposed between the advance information and the test stimulus in Condition C could only account for a longer, not a shorter, reaction time under that condition.) The explanation lies, rather, in the different interpretations that the subject has been instructed to place on the same advance information cue in the two types of trials. Under Condition C the appearance of an upright outline of the letter "R," for example, informs the subject both that the test stimulus will be the letter "R" and that it will appear in the upright orientation. Under Condition I, however, the appearance of that very same cue informs the subject only that

the test stimulus will be the letter "R"; it provides no basis for assuming anything about the orientation in which it will appear.

How, then, does advance information about either identity or orientation alone have its effect upon reaction time? We argue that the only effect of presenting advance information either about identity or about orientation alone is to cut down, by a constant amount of about 100 msec., the time that it would otherwise take to determine the identity or orientation of the test stimulus itself. When, however, both kinds of advance information are furnished, the subject is able to proceed with an entirely different preparatory process of rotating a mental image of the indicated character into the indicated orientation and, thus, becomes able to respond with uniform rapidity to the ensuing test stimulus. (See Appendix, pp. 120-1).

The relationship between prestimulus and poststimulus rotation. The present experiment, unlike previously reported experiments on mental rotation, permits us to estimate not only the time required to respond to a rotated stimulus but also the time required to *prepare* for a rotated stimulus. We assume that both times include the time needed to carry out the rotation of a mental image. However, there are differences between the two cases.

When (as in Condition I) the rotation is started only after the onset of the test stimulus, we assume that the mental image of the test stimulus is rotated from the tipped orientation in which it has been presented back into the standard upright orientation in which it can be compared with the normal representation in long-term memory. When (as in Condition B-1000) the rotation is completed prior to the onset of the test stimulus, we assume that the mental image of the normal version of the designated character is rotated from the standard upright orientation in which its outline has appeared as the identity cue into the tipped orientation in which it is about to appear as the actual test stimulus. Moreover, the character that is being mentally rotated is continuously present as a visual stimulus in the first case, but is present, if at all, only as a memory representation in the second. The question naturally arises as to whether the mental images in these two cases (which require rotation in opposite directions and in the presence of different degrees of external support) are nevertheless rotated at similar rates.

Now, whenever some portion of the required mental rotation has been carried out in preparation for the upcoming test stimulus, the subsequent reaction time should depend less strongly upon the orientation of that test stimulus. Condition B-1000 furnishes the extreme example. Here the entire rotation is completed in advance and so results in a virtually flat reaction-time function. In the case of Condition B-400, however, the reaction-time function—though lower than the functions for Conditions I and B-100—does not differ from either of them in shape (Figure 4.6). (An analysis of variance for Conditions I and B-400, in particular, showed that the condition-by-orientation interaction was nonsignificant.

Essentially the same picture emerges from the reaction-time distributions plotted in Figures 4.9 and 4.10. Although the proportion of "fast" responses does increase as we move from Condition I to Condition B-400 (particularly for the three "fast" subjects), the modal peak of the 180° distribution shifts no more than 100 msec for either the "fast" or the "slow" subjects. A full 700-msec duration of the orientation cue is necessary for the reaction-time function to flatten appreciably (Figure 4.6) and for the modal peak of the reaction-time distribution to shift markedly to the left (Figures 4.9 and 4.10). Apparently orientation information must be supplied for a 400-msec period before there is a significant reduction in the amount of post-stimulus rotation on most trials.

By analogy with the time apparently required to determine orientation on the basis of the test stimulus itself, we might have expected that the time required to determine orientation on the basis of the advance cue would also take about 100 msec. However, if 100 msec sufficed to extract this information from the inclined arrow in a fully usable form, then the curve for Condition B-100 should have been displaced downward by some 100 msec from the curve for Condition I (in which the orientation information had to be determined *after* the presentation of the test stimulus). But there is no difference between the heights of the reaction-time curves for I and B-100, whereas the curve for Condition B-400 *is* displaced downward from both of these by about 100 msec. Accordingly, we tentatively arrive at the unexpected conclusion that close to 400 msec may be needed to complete the processing of the orientation cue. In addition to the time needed merely to determine the orientation of the tipped arrow, this time may include times needed to convert this information into a form applicable to an ensuing alphanumeric character and, possibly, to overcome any backward masking of the arrrow or disruption of its interpretation caused by the sudden onset of the test stimulus.

As a rough approximation, we suggest that following the onset of the orientation cue, (a) a period of about 400 msec is usually required before the preparatory rotation is effectively started and (b) a period approaching 1000 msec is needed to ensure that nearly all of the subjects have completed a rotation of 180°. By subtraction we conclude that the time required to complete a 180° preparatory rotation for one of these stimuli is 600 msec—or less, in view of the variability in the more directly estimated poststimulus rotation times (Figure 4.8). For if 600 msec suffices to ensure completion of

a 180° rotation on almost *all* trials, then the time required to complete the preparatory rotation on an *average* trial is less than 600 msec.—perhaps closer to 400 or 500 msec. This estimate agrees well with the estimated time required to rotate 180° *after* the presentation of the test stimulus, since the differences between the mean reaction times for 0° and 180° under all of the four conditions, N, I, O, and B-100 (Figures 4.5 and 4.6), range between roughly 400 and 500 msec.

Separate consideration of the "fast" and "slow" subjects permits a further comparison of the prestimulus and poststimulus rotations. Notice in Figure 4.8 that as we move from 0° to 180°, the mean reaction time under conditions of poststimulus rotation (I and B-100) shifts about 200 msec for the three fast subjects and about 500 msec for the five slow subjects. Then notice in Figures 4.9 and 4.10 that as we move from Condition B-400 to the condition of prestimulus rotation, B-1000, the peak of the reaction-time distribution for a stimulus at 180° correspondingly shifts about 300 msec for the three fast subjects and about 500 msec for the five slow subjects. Again there is reasonably good agreement between the prestimulus and poststimulus cases (except, possibly, for the 100-msec discrepancy for the smaller group of three fast subjects).

We have been focusing here on those conditions in which a preparatory rotation either could be completed before stimulus onset (B-1000) or else, in most cases, could not even be started (I, B-100, and B-400). We turn now to a consideration of the relation between prestimulus and poststimulus rotation in the intermediate condition, B-700, in which the preparatory rotation is usually initiated but not completed prior to the onset of the test stimulus. Again it is best to discuss the groups of fast and slow subjects separately.

In the case of the group of five slow subjects, the reaction-time distribution for Condition B-700 (Figure 4.10) is widely spread out and contains the suggestion of three components—with modal peaks centered, respectively, in the 300-600 msec fast range, in the 600-900 msec intermediate range, and in the 900-1200 msec slow range (as these three ranges are depicted by vertical bands in the figure). We could interpret the (33%) fast responses as resulting from trials in which subjects were able to complete the 180° rotation before stimulus onset, and the (17%) slow responses as resulting from trials in which subjects abandoned their incomplete preparatory rotation and started over with a (reverse) rotation of the test stimulus itself. If so, it is most natural to attribute the remaining intermediate responses to trials in which the subject continued a preparatory rotation (which had been started but not completed before stimulus onset) until it was completed *after* stimulus onset.

The quantitative values of the means and relevant modal points of the reaction-time distributions are in satisfactory agreement with this account in the case of the larger group of five "slow" subjects. Notice, again, that as we move from Condition B-400 to B-700 to B-1000 in Figure 4.10, the modal point (and indeed the overall mean) shifts left by roughly 300 msec, each time, from the middle of the band of slow responses to the middle of the band of intermediate reponses to the middle of the band of fast responses. In other words, each additional 300 msec of time allowed for preparatory rotation cuts about 300 msec off the time that the subject then requires to complete that rotation after the presentation of the actual stimulus.

This trade-off between pre- and poststimulus rotation time does not emerge so clearly in the case of the smaller group of three "fast" subjects. In Figure 4.9, the leftward shift of the mode of the reaction-time distribution from Condition B-400 (or B-100) to B-700 is closer to 200 msec than to 300 msec, and the full 300-msec shift is not achieved until condition B-1000 when, according to our analysis, the subjects have had at least 600 msec in which to complete their preparatory rotation. If the three fast subjects were faster simply by virtue of faster mental rotation (as their poststimulus reaction-time distributions suggest in Figure 4.8), we should expect that they would nearly always be ready for the test stimulus in Condition B-700. The distribution for B-700 would then be essentially identical to that for B-1000, rather than displaced 100 msec to the right—as it is in Figure 4.9. (Possibly there is a connection between this 100-msec discrepancy and the other 100-msec discrepancy, noted earlier, for these same three subjects.)

Indeed, several of the differences between the results for these two internally homogeneous groups of subjects suggest that the "fast" subjects may have used a different method in addition to their postulated greater rate of mental rotation. These differences include (a) the two 100-msec. discrepancies that, as we just noted, occurred only for the group of fast subjects, (b) the more compact reaction-time distributions obtained from this group (Figures 4.8 and 4.9), and (c) the greater tendency of the individual reaction-time functions obtained for this group to bend sharply at about 60° on either side of 180° (as is illustrated for one of these subjects at the top of Figure 4.7). Because of the small number of subjects in the "fast" group, we do not feel prepared at this time to make a definite claim concerning their method of processing.

Despite some uncertainty about these three subjects, we believe that the data from all eight subjects support our general conclusion that subjects carried out mental rotations both to respond to a rotated stimulus in the absence of advance information as to its orientation, and to prepare for a rotated stimulus on the basis of such advance information. Moreover, we believe that the data from the majority of the subjects agree quite well with our more specific conclusion that subjects carried out such prestimulus and poststimulus rotations at essentially the same rates.

EXPERIMENT II. DEMONSTRATION OF A CORRESPONDENCE BETWEEN AN IMAGINED AND AN ACTUAL ROTATION

Our purpose in undertaking this second experiment was to obtain evidence bearing upon two notions that have been implicit throughout our discussion of Experiment I but that could not be directly substantiated on the basis of that experiment. One notion concerns the nonlinearity of the obtained relationship between reaction time and orientation of test stimulus. Although the original evidence for a process of mental rotation was largely based upon the finding of a strongly linear relationship (Shepard & Metzler, 1971), we have been arguing that the relationship found in the present Experiment I (as well as in some pilot experiments by Shepard and Klun[6])—despite a consistent and marked nonlinearity—is also the result of an underlying process of mental rotation. The other notion concerns the sense in which this mental process is one specifically of rotation. Since the term "rotation" is ordinarily defined only in relation to an external physical process, our use of this term in connection with an internal mental process implies that there is some sort of one-to-one correspondence or isomorphism between mental and physical processes of rotation.

It is possible that an internal process quite different from rotation—such as visual search, feature detection, verbal analysis, or some other digital computation—might enable a subject to determine that two objects (or images) are identical except for orientation. Such a process might even become more difficult and protracted as the two objects differ more widely in orientation. In order for an internal process to qualify as the kind of analog process that we would call a mental rotation, however, intermediate stages of the process must have a one-to-one correspondence to intermediate stages of an actual physical rotation of the one object into congruence with the other in the external world. Central to the concept of an analog process, then, is the idea of the path or trajectory of the process. If the internal process is one of rotation, not only should the starting point and end point of such a process correspond to the two objects compared, but also any intermediate point on this trajectory should correspond to an external object in an intermediate orientation—even though such an external object may not be physically present.

But what does it mean to say that at a given point in time an internal process has a one-to-one corresondence to a particular external object that is not physically present? It does not necessarily mean that there is any concrete structural resemblance or first-order isomorphism between the pattern of activity in the subject's physical brain at that moment and the corresponding external object (if it were present). Nor does it mean that there has been anything actually rotating within the subject's physical brain. It means only that there is a one-to-one relation between the internal representation and the corresponding external object in the specific sense that the subject is especially disposed to respond to that particular object in that particular orientation at that particular moment—if it were actually to be presented (Shepard, 1975).

To be "turning something over in one's mind," then, is to be passing through an ordered series of dispositional states. On the basis of introspective evidence and other considerations presented earlier, we are inclined to refer to these states as successive mental images of a rotating object. Although there are reasons to suppose that there may be some degree of abstract "first-order" isomorphism between such an internal process and a corresponding external rotation of the imagined object (Shepard, 1975) all we require in order to speak of "mental rotation" is that the internal process produces the necessary series of dispositional states. This may be achieved by means of a second-order isomorphism (Shepard & Chipman, 1970) according to which the internal process—whatever its neurophysiological nature—has an important part in common with the internal process that goes on when one is actually *perceiving* such an external rotation.

In this second experiment we have subjects imagine an alphanumeric character rotating at a certain (externally paced) rate within a blank circular field. At a random point during this purely mental process, we display a normal or backward version of the imagined character—either in the orientation that the subjects should be imagining at that moment in their rotation or in some other orientation chosen at random. If the subjects are actually carrying out a mental rotation, the speed with which they can discriminate whether this probe stimulus is normal or backward should be greatest when the probe appears in the orientation momentarily assumed by their mental image, for only then will they be able to make an immediate match between their internally rotating image or template and the externally presented test probe. Results of this sort should enable us to say that something is indeed rotating during this process—namely, the orientation at which the subjects are most prepared for the external presentation of the corresponding physical stimulus.

With regard, next, to the matter of nonlinearity, all four of the alternative explanations that we offered for the concave upward shape of the reaction-time functions were based on two observations. The first was that the alphanumeric characters (unlike the stimuli used by Shepard and Metzler) have a uniquely defined and well-learned upright orientation. The second was that the hypothesized rotation was always between the orientation of the presented test stimulus and this unique upright—not (as in the experiment by Shepard and Metzler) between the two orientations within all possible pairs. If any or all of the four explanations that we offered for the obtained nonlinearity are correct, we should be able to counterbalance the asymmetrical effects of the unique upright orientation and thus be able to obtain a more nearly linear function. We merely need to ensure that any mental rotations that the subject must perform *after* the onset of the test probe have starting and stopping points that are *both* evenly distributed around the 360° circle.

The theoretically critical point therefore resides not in the linearity itself but in the indication that it can give us that the underlying process is composed of parts (corresponding to rotations through smaller angles) that are necessarily performed in sequential order and for which, consequently, the performance times are additive. Thus no matter what the effective rotation times may be between particular adjacent points separated by 60° around the circle, if the time required to go from any one point to any other nonadjacent point is an additive combination of the component times to go between the intervening adjacent points, then the average time to go between the points in all pairs separated by n 60° steps should increase linearly with n. To show that the time to rotate from A to C is an additive combination of the times to rotate from A to B and from B to C is to furnish another kind of evidence that the process of rotating from A to C passes through a point, B, corresponding to an intermediate orientation. A finding of linearity would thus support further our claim that the process of mental rotation is an analog process.

In order to provide for a test of this predicted linearity, we depart from the previous procedure in which the presentation of the test stimulus was always in strict agreement with any advance information as to identity or orientation. This time the test stimulus does not always appear in the orientation for which the subject is preparing at the given moment. On half of the trials it appears in each of the five other evenly spaced orientations with equal probability. In this way we are for the first time able to determine the relationship of reaction time to departure of the test probe—not only from upright but also from the orientation in which that probe was expected. On the basis of our previous results, we might anticipate that any nonlinearity will be confined to the former relationship.

If our explanations as to the source of this nonlinearity are correct, we should expect the latter relationship to approximate the linearity found by Shepard and Metzler (1971).

Method

Subjects. Eight subjects were individually run through the complete experimental design, requiring four hours of participation from each. Of these subjects, two were female and two others (including RNS) had previously participated in Experiment I. (The results for the two females and for the two experienced subjects were all typical of those for the remaining four experimentally naive male students.)

Stimuli. This time we used only two from the set of six asymmetric alphanumeric characters employed in Experiment I—the upper-case "R" and the numeral "2." As before, each of these characters could appear as a normal or backward test stimulus in any of the six orientations, spaced in equal 60° steps around the circle. And again the subjects were instructed to determine, as rapidly as possible, whether each such test stimulus had appeared in its normal or in its backward version, regardless of its orientation within the two-dimensional test field (see Figures 4.1 and 4.2). The tachistoscopic apparatus, the visual angles of the stimuli, and the right- and left-hand pushbutton responses were all the same as described for Experiment I.

Structure of individual trials. All trials in this experiment were analogous to trials of type B in Experiment I in that cues as to both identity and orientation were available in advance of the onset of every test stimulus. There were, however, three important changes from the procedure used in Experiment I. (a) The advance cues were presented in auditory rather than in visual form. (b) While the cue as to the identity again agreed with the identity of the ensuing test stimulus on all trials, the cue as to the orientation this time agreed with the orientation of the test stimulus on half of the trials only. (c) The cue as to orientation, rather than being fixed in a single, random orientation on each trial, indicated an orientation that was progressively moving in a clockwise direction throughout the course of any one trial.

During each trial, the subject sat fixating the blank circular field in the tachistoscope. To start a new trial, the experimenter orally announced which of the two characters, "R" or "2" (with which the subject had previously been visually familiarized), was scheduled to appear as the test stimulus on that trial. Then the experimenter started playing a magnetic tape on which the verbal commands "up," "tip," "tip," "down," "tip," "tip" had previously been recorded at a controlled rate of one command per half second. On the basis of prior instructions and practice trials, the subject was to imagine the normal version of the announced character starting in its upright orientation and rotating clockwise at a rate of 60° per half second, in synchrony with the auditory commands. (Thus the initial command "up" notified the subject to begin rotating the internally generated mental image from its initial upright orientation and, three commands later, the word "down" indicated that this image should now be rotated around into its 180° orientation.) To assist the subject to keep pace with the auditory commands, there were six small tick marks visible around the border of the circular field at 60° steps (starting at the center of the top).

Quickly following a randomly preselected one of these verbal commands, the probe stimulus appeared in one of the six equally spaced orientations within the circular field. Thereupon the subject was to actuate the right- or left-hand switch as rapidly as possible to indicate whether the visual probe was the normal or backward version of the character, respectively. (As in Experiment I, however, this response assignment was reversed for the one left-handed subject.) The auditory commands terminated with the onset of the visual test stimulus, but the test stimulus remained on until after the subject made a response. The interval from the onset of the visual probe to the actuation of one of the two response buttons was recorded as the reaction time.

Overall experimental design. One half of the trials, determined according to a random sequence, were "probe-expected" trials in which the probe stimulus was presented in the orientation designated by the current auditory command. The other half of the trials were "probe-unexpected" trials in which the probe appeared, with equal probability, in any one of the five other possible orientations. (As a consequence, at each orientation there were five times more probe-expected observations than probe-unexpected observations.) Within both the probe-expected and probe-unexpected trials, one half of the probe stimuli were presented in their normal and one half in their backward versions, again according to a random sequence. This entire set of observations was collected for both test characters ("R" and "2"), yielding a total of 240 experimental trials per subject (following an initial set of 48 practice trials). Order of trials was randomized anew for each subject, and there was no blocking of trials by any experimental factor. Trials on which errors were made were retaken, if possible within the same session, and error rates were low (ranging from 10% to 2% for individual subjects).

Reaction-time Results

The effect of absolute orientation of an expected test stimulus. "Probe-expected" trials, in which the visual probe appeared in the orientation corresponding to the current auditory command, could be considered analogous to trials of Type B-1000 in Experiment I. For if the subjects were in fact rotating a mental image of the designated character in time with the auditory commands, they should be able to make a direct match of their current mental image against the visual probe when it suddenly appears. Thus they should be able to determine with great rapidity whether that probe is normal or backward—whatever the currently designated orientation happens to be. Just as in the earlier Condition B-1000, the function relating reaction time to absolute orientation of the test stimulus should be relatively flat (cf. the earlier Figures 4.5 and 4.6).

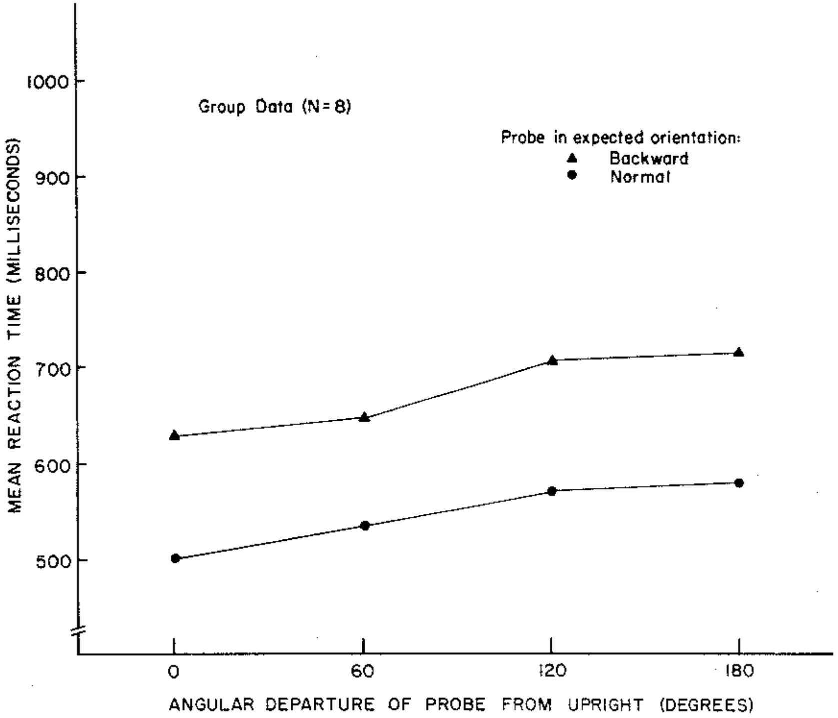

Figure 4.11 Mean reaction time as a function of angular departure of the probe stimulus from the upright orientation for those trials on which the probe stimulus appeared in the expected orientation. Separate curves are plotted for reaction times to "normal" and "backward" test stimuli.

Mean reaction times for those trials in the present experiment in which the probe stimulus appeared in the expected orientation are plotted in Figure 4.11 as a function of the angular departure of the

(expected) probe from its standard upright orientation. These times are averaged over correct responses to both test characters for all eight subjects. However, they are plotted separately for trials in which the probe was normal or backward, as indicated in the figure.

In agreement with Experiment I, the responses to the backward probes were consistently longer than the responses to the normal probes—by an amount that was independent of departure from upright (and also independent of departure from expected orientation—as we shall soon see). In the present experiment, this difference in reaction time was on the order of 100 to 150 msec. The average difference is therefore some 50 msec longer than the average difference between reaction times to normal and backward stimuli in Experiment I. We shall argue that this is the result of an additional operation that must be carried out in the present experiment. Since the test stimulus was always presented in the orientation indicated by any advance orientation cue in Experiment I, a mismatch with an already rotated preparatory image automatically ensured that the stimulus was backward. In the present case, however, the subject must make a further determination as to whether the mismatch was the result of the probe's being presented in its backward version *or* in an unexpected orientation.

For simplicity of presentation in this and the following figures, we have averaged the reaction times for the symmetrically related orientations of 60° and 300° and of 120° and 240°. Thus the independent variable is now departure from upright in either direction, rather than absolute orientation in a specifically clockwise direction. As a consequence, the points plotted for 60° and 120° are based upon twice as many observations as the points plotted for 0° and 180°. However, if these reaction-time functions are "unfolded," so that all six orientations of the test stimulus are plotted separately, the functions become symmetrical about 180°—as in the earlier Figures 4.5 and 4.6.

The results displayed in Figure 4.11 are in good agreement with our expectations. The slight humping up near 180° in the previously obtained reaction-time functions for Conditions C and B-1000 reappears here as the small but consistent rise in reaction time from 0° to 180°. As we suggested before, subjects may simply require a little more time to compare two images when those images are both in a less familiar orientation. In any case, the average increase from 0° to 180° is only about 80 to 90 msec. If we contrast this with the 400 to 500 msec increase in reaction time at 180° found (in the previous conditions N or I) when no orientation information was provided, we see that the effect of the absolute orientation of an expected test stimulus is again very small.

That subjects in the present experiment were able to classify a test stimulus as normal (as opposed to backward) in only 500 to 600 msec, as long as the orientation of that stimulus coincided with the rotating orientation expected, supports our claim that they arrived at their classification by matching the presented probe against a "rotating" mental image. As further support, however, we also need to show that subjects required appreciably more time whenever the probe appeared in an unexpected orientation.

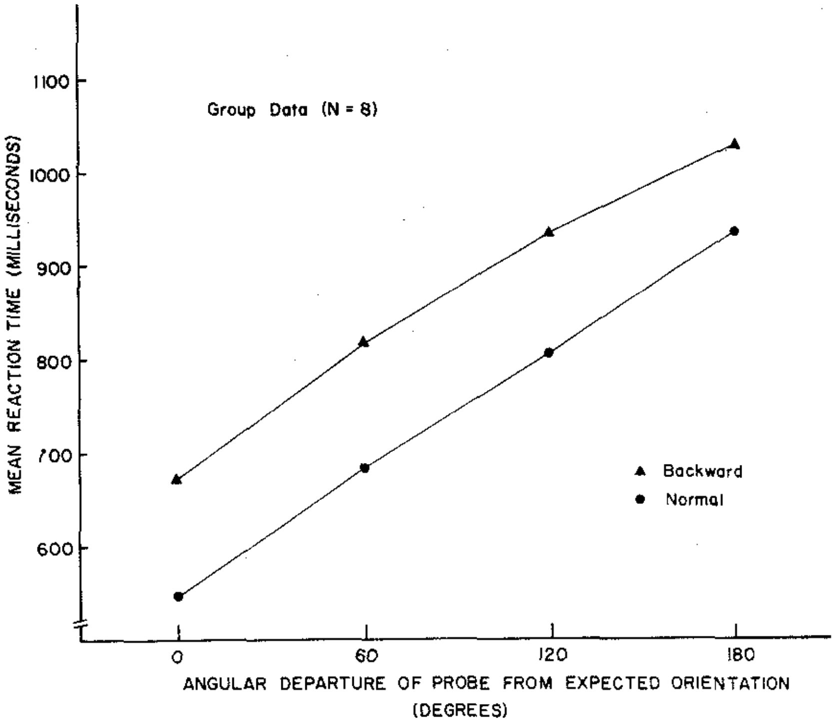

Figure 4.12 Mean reaction time as a function of angular departure of the probe stimulus from the expected orientation. Separate curves are plotted for reaction times to "normal" and "backward" test stimuli.

The effect of angular departure from expected orientation. In this experiment we can observe for the first time what happens when the test stimulus is presented in an orientation that departs from what the subject has been led to expect. Figure 4.12 illustrates the main result. Mean reaction time is plotted as a function of the angular difference between the orientation of the visual probe and the orientation that the subject should have been expecting at that point in the sequence of auditory commands. As in the preceding figure, separate curves are presented for trials in which the probe was normal or backward and, again, responses to probes of the latter type were some 100 to 150 msec slower. Also as before, the reaction times have been averaged over corresponding clockwise and counterclockwise departures at both 60° and 120° over both test characters and over all eight subjects. In addition, for this figure the reaction times have also been averaged over all angular departures of the probe from upright.

As predicted, when the probe appeared in some orientation other than the (rotating) orientation expected, reaction time increased markedly with the difference between the expected and the actually presented orientations. Indeed the overall increase, from 0° to 180°, is close to 400 msec. This increase is some five times greater than the 80 to 90 msec increase shown in the preceding Figure 4.11. At the same time, it is comparable to the 400 to 500 msec increase observed in Experiment I for departure from the orientation naturally expected when there was no advance information about orientation—namely, the standard upright orientation.

However, although the functions exhibited in Figure 4.12 thus agree with the functions displayed (in Figures 4.5 and 4.6) for the earlier conditions N, I, O, and B-100 with respect to overall slope, they differ markedly from those earlier functions with respect to shape. For whereas the earlier functions were uniformly concave upward, the functions shown in Figure 4.12 are both strikingly linear. But the linearity of these new functions is just what we have predicted on the basis of two assumptions. The first is that when the orientation of the probe fails to agree with the imagined orientation, the subject must undertake an additional poststimulus rotation in order to achieve a match between that probe and the internal representation of the corresponding normal character. The second assumption is that when the starting and ending points of the required poststimulus rotations are evenly distributed around the circle, the biasing effects of the special upright orientation should cancel out in such a way as to reveal the linear and, hence, the underlying sequential-additive nature of mental rotation.

The joint effects of departures from expected and upright orientations. We have argued that any nonlinearity in the present data should emerge when reaction time is plotted against departure from upright (as in Experiment I) rather than against departure from expected orientation (as in the preceding Figure 4.12). Evidence supporting this argument is illustrated in Figure 4.13. Here the dependence of reaction time upon angular departure of the probe from upright is separately plotted for each value of angular departure of the probe from its expected orientation. The plotted means are averaged over subjects, characters, and, this time, both "normal" and "backward" responses as well. The lowest curve (for 0°) is thus the average of the two curves displayed earlier in Figure 4.11. Again, the flatness of this bottommost function shows that when the probe appears in the expected orientation, the absolute angle of this orientation has relatively little effect on reaction time. In this case we cannot expect to find any nonlinearity.

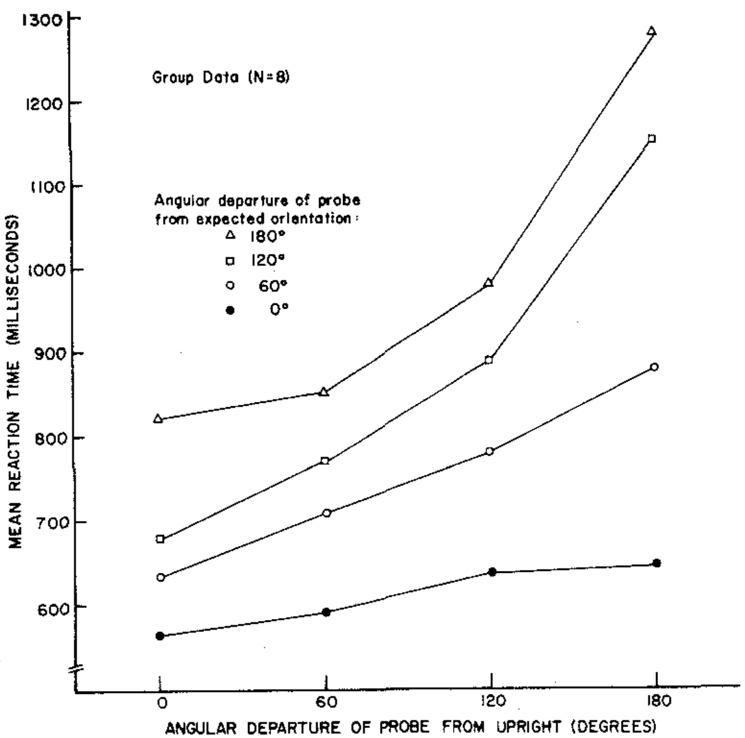

Figure 4.13 Mean reaction time as a function of angular departure of the probe stimulus from the upright orientation, plotted separately for each of the four angular departures of the probe stimulus from the expected orientation. (Reaction times to "normal" and "backward" test stimuli are averaged together.)

As the visual probe departs more and more from its expected orientation, however, three changes take place in the average data. The function relating reaction time to departure from upright (a) rises in overall height, (b) increases in positive slope, and (c) becomes increasingly concave upward. The first change—the increasing height of successive functions in Figure 4.13—is just the effect illustrated in Figure 4.12. As we saw there, this increase is essentially linear with departure from expected orientation. The second and third changes—the increasing steepness and curvature of the successive functions—are to some extent analogous to the changes in the family of curves for the B-conditions in Experiment I (Figure 4.6). In both cases the relation of reaction time to departure from upright becomes stronger and more nonlinear as the advance information as to orientation is decreased in effectiveness—whether by a reduction in duration (Figure 4.6) or in accuracy (Figure 4.13).

The analogy between the family of curves in the present Figure 4.13 and the B-family curves in the earlier Figure 4.6 should be advanced with caution in view of the smaller number of observations contributing to each of the points in the three upper "probe-unexpected" curves of Figure 4.13 (particularly at 0° and 180°). Moreover, an additional consideration, which does not arise in connection with Experiment I, may come into play in the present experiment. Consider trials represented by the leftmost datum point of the highest curve in Figure 4.13. These are trials on which the subject has presumably rotated a mental image of the designated character into its upside-down position but on which the test probe unexpectedly appears in its upright position. Some subjects reported that when this happened, they identified the upright test probe as normal or backward without rotation—presumably by comparing it against the normal upright representation in long-term memory, rather than by continuing to rotate their short-term image (which was then at 180°) all the way back to 0° for comparison with the upright probe. We would expect subjects who operated in this way to yield relatively short reaction times to probes that unexpectedly appear near 0°. Such relatively rapid reaction times may contribute considerably to the slopes of the upper curves in Figure 4.13.

Perhaps the most striking comparison illustrated in Figure 4.13 is between the rightmost point of the lowest curve and the leftmost point of the highest curve. This comparison reveals that on the average subjects responded nearly 200 msec faster to an expected test stimulus at 180° than to an unexpected test stimulus in the standard upright orientation. As a consequence of the relatively small number of observations contributing to the upper point, the numerical value of this difference is somewhat unreliable; however, the direction of this difference is the same for all eight subjects. This comparison therefore furnishes strong evidence that the internal representation that the subjects were "mentally rotating" in time with the auditory commands was more available for matching against the externally presented test stimulus than was the permanent representation of the normal upright character that the subjects presumably retain in long-term memory.

Analysis of variance indicated that the effects of angular departure of the probe from upright, angular departure of the probe from the expected orientation, and the interaction between these two were all statistically reliable.[7]

Results for individual subjects. In Figure 4.14 we follow our earlier convention of presenting the curves for the two subjects that were most extreme with respect to overall reaction time. The two panels on the left correspond to the earlier Figure 4.11, while the two panels on the right correspond to the earlier Figure 4.12. There is a considerable difference in overall height of the curves obtained for the "fast" subject (top panels) and the "slow" subject (bottom panels). (Note, incidentally, that the vertical scales for the "slow" subject have been linearly compressed with respect to the vertical scales for the "fast" subject.) Nevertheless both subjects yielded functions that were essentially flat when the probe appeared in the expected orientation and increased monotonically and approximately linearly with departure of the probe from its expected orientation.

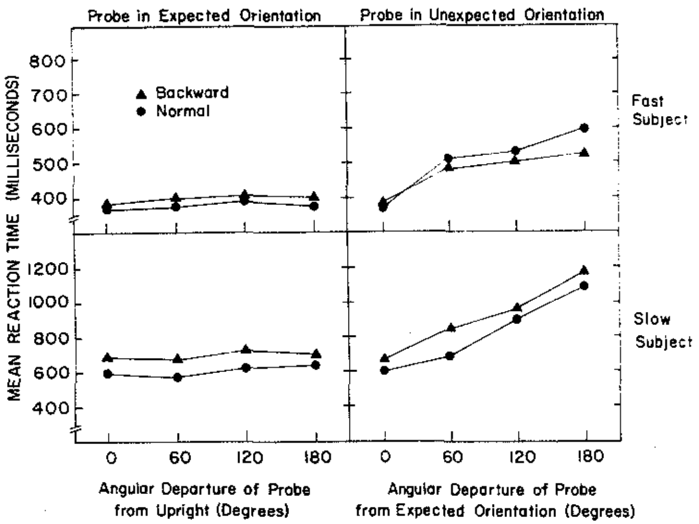

Figure 4.14 Mean reaction-time functions for two individual subjects—*the fastest subject (upper panels) and the slowest subject (lower* panels). Left-hand panels correspond to the group functions displayed in Figure 4.11, and right-hand panels correspond to the group functions displayed in Figure 4.12.

Curves for all of the other subjects were similar in shape to these curves (and for some of the subjects the curves corresponding to those plotted on the right in Figure 4.14 were even more linear).

Discussion

The introspective reports that the subjects gave us following their participation in Experiment II are consonant with our interpretations of their reaction times. The subjects claimed that they were indeed able to imagine the normal version of the designated character rotating clockwise in time with the auditory commands. The subjects also indicated that they used this rotating mental image as a sort of template against which to compare the test stimulus when it suddenly appeared within the circular field. If a match was achieved, they immediately executed the readied (e.g., right-hand) response. If not, they then had to determine whether the mismatch resulted from the fact that the stimulus had been presented in its backward version or in an unexpected orientation.

If the subjects determined that the visual probe was backward (perhaps by performing an additional mental operation of reflection), they then switched control to the other hand before executing their response. The average additional time for a mismatch response in Experiment II was some 50 msec longer than in Experiment I. This difference may be the time required for the additional operation (possibly of reflection) that was required only in Experiment II (in which, for the first time, the probe could differ from what was expected by a reflection *or* a rotation).

Concerning trials on which the visual probe differed from the subject's internal representation by a rotation, subjects reported two different strategies: When the probe departed markedly from upright, they tended to carry out a further (poststimulus) rotation in order to bring their rotating mental image into congruence with the external probe and then proceeded as described above for the case in which no such further rotation was necessary. When the probe was close to upright while their mental image was far from upright, they tended to abandon their rotating mental image and to determine whether the probe was normal or backward directly—after imagining the probe itself rotated back to upright, if necessary.

With respect to the reaction-time results, we regard as especially significant (a) that reaction times were short and relatively independent of orientation whenever the probe appeared in the (rotating) orientation expected and (b) that these times increased as the probe *departed from this expected orientation—even when this very departure* brought the probe closer to its standard upright orientation.

This conclusively demonstrates that during the alleged process of mental rotation, something was indeed rotating—namely, the orientation in the external world at which the subject was most prepared for the presentation of the test stimulus. It also demonstrates that this process of mental rotation was an analog process, at least to the extent that it went through intermediate states that had a one-to-one relation to intermediate orientations in the external world. As the lowest curve in Figure 4.11 shows, the subjects required an average of only 500 to 600 msec to perceive the probe, to match it against the "mentally rotating" internal representation, and to make a correct response. This suggests that the representation was in a form that was particularly suitable for comparison with the visual stimulus. Although for several subjects there undoubtedly were kinesthetic concomitants, it is tempting to refer to this "rotating" internal representation as a visual image.

CONCLUSIONS

Empirical Findings

The stimuli used in both experiments were two-dimensional visual patterns (certain upper-case letters and numerals) that (a) are highly familiar, (b) are characteristically seen in a well-learned and uniquely defined upright orientation, and (c) are not symmetrical about any axis. On each trial one such stimulus or its mirror image was presented in some orientation within its two-dimensional plane. The subjects were instructed to actuate a right-hand or left-hand switch, as rapidly as possible after the onset of the stimulus, to indicate whether that stimulus was presented in its normal or backward version, respectively—regardless of its orientation within the picture plane. The principal findings concerning the measured reaction times can be summarized as follows:

1. Throughout, reaction times were consistently shorter to the normal stimuli than to their mirror images. The difference averaged about 100 msec or somewhat less and was independent of orientation and several other experimental variables (Experiment I, and Figures 4.11, 4.12). This difference varied somewhat from subject to subject, however, and it averaged some 50 msec longer when the stimulus could depart from what was expected by a rotation as well as by a reflection (Figures 4.11, 4.12).

2. When the stimulus appeared in an orientation for which the subject was not specifically prepared, reaction time increased monotonically with the departure of the stimulus from the natural upright orientation (Experiment I, Conditions N and I), or from the rotated orientation that the subject had been set to expect

(Experiment II). For the maximum possible departure of 180°, the average increase in reaction time was on the order of 400 to 500 msec (but, again, varied from subject to subject).

3. When the subject was not set to expect the stimulus in a particular (rotated) orientation, the increase in reaction time with departure from upright consistently conformed to a concave upward function (Figure 4.5, Conditions N and I).

4. When the subject was set to expect a particular stimulus in a particular orientation, the increase in reaction time with departure from that expected orientation (when averaged over all such orientations) conformed to a remarkably linear function (Figure 4.12)—reminiscent of the functions reported by Shepard and Metzler (1971) for rotated three-dimensional objects.

5. When the subject was given advance information only about the identity of the stimulus or only about its orientation, reaction time was reduced by a constant amount of about 100 msec in either case, regardless of the (nonconflicting) orientation of the ensuing stimulus (Figure 4.5, Conditions I and O versus Condition N).

6. When the subject was given valid advance information about both identity and orientation, reaction time to an upright stimulus was reduced by the sum of the reductions attributable to the two kinds of advance information provided separately—i.e., by a total of some 200 msec (Figures 4.5 and 4.6, Condition B-1000 at 0° or Condition B-400, as compared with Condition N).

7. When the subject was thus set for a specific stimulus in *any* specific orientation, reaction time to such a stimulus was consistently short (about 500 msec) and remarkably independent of the angle of that expected orientation (Figure 4.5, Condition B-1000, and Figures 4.11 and 4.14). Consequently, when the expected orientation was other than upright, the reduction in reaction time resulting from the two kinds of advance information exceeded the sum of the reductions attributable to either kind alone. (For a stimulus at 180° from upright, the reduction averaged about 600 msec rather than the 200 msec prescribed by simple additivity.)

8. When a particular stimulus was expected in a particular rotated orientation, the reaction time to such a (rotated) stimulus was consistently shorter than the reaction time to that stimulus presented in the standard upright orientation—in which it was then no longer expected (Figure 4.13, second, third, and fourth points on the lowest curve versus the corresponding leftmost points on the second, third, and fourth curves, above).

9. The time required to be fully prepared for presentation of a particular stimulus at a particular orientation increased as the (validly) indicated orientation departed from the standard upright orientation.

For an orientation of 0°, minimum reaction times (of about 500 msec) were achieved if the orientation information had been available 400 msec prior to stimulus onset. But for an orientation of 180°, reaction times approaching this minimum (500 msec) level were not achieved until the orientation information had persisted for about 1000 msec (Figure 4.6, Conditions B-400 and B-1000).

10. Except, possibly, in the case of a few "fast" subjects, there was an approximate trade-off between preparation time and reaction time such that each additional 300 msec of orientation information (beyond the 400 msec needed for optimum performance at 0°) reduced the subsequent reaction time to a stimulus at 180° by another 300 msec (Figures 4.6, Conditions B-400, B-700, and B-1000 at 180°; and Figure 4.10).

11. Reaction times to an expected stimulus had approximately the same short (500 msec) values whether the advance information was presented (a) in the form of a visual outline of the stimulus already rotated to the orientation in which it was to appear (Figure 4.5, Condition C), (b) in the form of a visual outline of that stimulus in its upright orientation followed by a rotated arrow to indicate the orientation in which it was to appear (Figure 4.5, Condition B-1000), or (c) in a purely auditory form (Figure 4.11).

12. Subjects were able to maintain an optimal level of readiness for the appearance of a visual stimulus in an orientation that was rotating clockwise (in accordance with concurrent auditory cues) at a rate of 120° per second (Figure 4.11). (Higher rates may also be feasible for some subjects but have not yet been tried.)

Theoretical Interpretations

Nowhere in the preceding list of "empirical findings" did we speak of any such things as a mental image or a mental rotation. Our intention in preparing that list was to confine ourselves to objective and, we hope, reproducible patterns evident in the "hard data" of the recorded reaction times. We even avoided any mention of the subjects' postexperimental introspective reports. The list was thus put forward as a set of facts standing in need of a theory. It may be that someone will be able to formulate a theory that satisfyingly accounts for this particular set of facts without invoking any such concepts as mental imagery or mental rotation. Until this happens, however, we are inclined to favor a theory (a) that *has* now been formulated (at least in outline) and (b) that *is* consonant with the subjects' own introspective reports that they did indeed rotate mental images. Accordingly, the theoretical interpretations that we have been proposing to account for the empirical findings can be summarized as follows:

1. Subjects determine whether a familiar asymmetrical stimulus has been presented in its normal or in its backward (mirror-image) version by comparing the stimulus against an internal representation or mental image. Unless instructed otherwise (or unless presented in advance with the other version to be used as a model), subjects use the internal representation of the normal version for this purpose. (See Findings 1, 2, and 7, listed above.)

2. If the stimulus fails to match the image, subjects perform an additional operation (requiring 100 msec or somewhat less) in order to switch control to the alternative response (Finding 1). If there is the possibility (as in Experiment II) that the mismatch may result from a rotation rather than just from a reflection, subjects may perform still another operation (requiring perhaps 50 msec more) in order to verify that the mismatch was the result specifically of a reflection (Finding 1).

3. If the stimulus appears in its familiar upright orientation for subjects who are not specifically expecting it in that orientation, the internal representation used for comparison against the stimulus is the permanent representation of the normal upright version of that stimulus in long-term memory (Findings, 2, 6, 8).

4. If the stimulus appears in some other orientation that departs markedly from any orientation that the subject may have been specifically expecting, a mental image of the presented stimulus is mentally rotated from the orientation in which the stimulus appeared back into its standard upright orientation for comparison with the upright representation in long-term memory (Findings 2, 8).

5. Before subjects can compare a stimulus against an internal representation or can begin to rotate a mental image of it into a particular orientation, they must know both the identity and the orientation of that stimulus. Either or both of these two pieces of information that they do not have in advance must be independently extracted from the stimulus itself at the cost of about 100 msec for each piece—hence of about 200 msec for both (Findings 5, 6).

6. If information about both the identity and the orientation of the ensuing stimulus is provided sufficiently in advance, a mental image of the normal version of the designated stimulus is formed in its standard upright orientation (from the identity cue, in Experiment I, or from long-term memory, in Experiment II) and is mentally rotated into the designated orientation for rapid comparison with the stimulus when it appears (Findings 4, 7, 8, 10, 11).

7. If the stimulus appears in a (rotated) orientation different from the orientation of the preparatory mental image, the preparatory

image is then rotated into the orientation of the stimulus for purposes of comparison. This is true whether the difference in orientation arises because the preparatory rotation has not been completed (Findings 9, 10) or because the orientation of the stimulus is different from what the subject has been set to expect (Findings 2, 4). *In the former case, however, the preparatory rotation will not* even be started unless the advance orientation information (when presented in the form of a rotated arrow) has been available for some 400 msec (Findings 9, 10). When the duration of the orientation cue is too short, preparatory rotation is abandoned and an image of the stimulus itself is rotated back into upright, as described in Point 4, above.

8. Although the rate of mental rotation depends upon the individual subject (and, perhaps, upon the complexity and familiarity of the stimulus), on the average subjects can rotate a mental image of an alphanumeric character through 180° in some 500 msec (Findings 2, 9, 12). Moreover the rate is about the same whether the rotation is carried out in the presence of the external stimulus, after it has appeared (Finding 2) or in the absence of that stimulus, in preparation for it (Finding 9).

9. Subjects are able to carry out only the mental rotation of an internal representation of a particular concrete object, not a general abstract frame of reference. Thus even though subjects are given adequate advance information about the orientation of an upcoming *stimulus, they are unable to perform any preparatory rotation unless* they are also given advance information about the identity of that stimulus (Findings 5, 7).

10. In the case of stimuli (such as alphanumeric characters) that have characteristically been seen in (or close to) one particular orientation, that familiar upright orientation plays a special role in processes of rotation and/or matching. Such processes are relatively slow when the mental image being rotated or matched is in an unfamiliar orientation, and especially so when the image is close to 180° from upright (Finding 3).

11. Mental rotation is an analog process with a serial structure bearing a one-to-one relationship to the corresponding physical rotation. The time required (mentally) to rotate from an orientation A to an orientation C is just the sum of the times required to rotate from A to some intermediate orientation B, and to rotate from B to C (Finding 4). Moreover, in mentally rotating an object between any two widely separated orientations, A and C, the internal process passes through the mental image corresponding to that same object in some intermediate orientation, B (Finding 12). Consequently, *the orientation at which the subject is most prepared for the appearance of that object at each moment is actually rotating* with respect to the external world (Findings 7, 8, 12).

12. Even though the mental image may be internally generated in the absence of the corresponding visual stimulus (Finding 11), it nevertheless possesses an internal structure that is at least abstractly *isomorphic* to the structure of that external visual stimulus, since the image can be transformed in a way that corresponds specifically to a rigid rotation of the external stimulus (Finding 12); and, even after such a transformation, the image can be matched against the subsequent presentation of the rotated visual stimulus with essentially the same speed and precision as *if it were a straight memory image* remaining from an immediately preceding presentation of the identical rotated stimulus (Findings 7, 11).

Final Remarks

We are not claiming that the many familiar objects that we all encounter in various positions in our every-day lives have to be mentally rotated into some canonical orientation before they can be recognized. In the case of most objects and symbols with which we have to deal, there are sufficiently numerous, redundant, or orientationally invariant distinctive features that we may well achieve recognition directly—without need of a preliminary mental transformation. However, we *are* claiming that a demonstration that human subjects are capable of mentally rotating spatially structured objects is of considerable importance, even if a sharp demonstration requires special circumstances.

If rotation is mentally performed in an analogical manner, then the imagining, understanding, and planning of many other kinds of operations in the physical world may also be accomplished in an analogical manner. Hence, the mere demonstration that humans solve some problems (however contrived) by means of a basically analog process has important implications for the nature of the human mind. At the very least, it raises a question about the advisability of formulating theories of human behavior solely in terms of discrete processes of verbal mediation of symbol manipulation—as has been characteristic in experimental psychology and in computer simulation, respectively. Moreover, the fact that internal representations can be operated upon by such analog processes tells us something important about the nature of those representations. Clearly, such an internal representation cannot adequately be regarded either as an undifferentiated neural event (such as the activation of a particular neuron or population of mutually interchangeable neurons) at the physiological level, or simply as an unanalyzable symbol at the information-processing level. Instead, such a representation must have an internal structure that is itself to some extent analogically related to the structure of its corresponding external object. For during the process of rotation, the parts and the relationships among the parts must be transformed in very constrained ways in order to enable the kind of rapid, template-like match against an ensuing visual stimulus that we have demonstrated here.

Appendix: A Tentative Information-Processing Model

Figure 4.15 schematically portrays our first approximation to what may take place when a subject without advance information attempts to determine whether an alphanumeric character presented in some arbitrary orientation is normal or backward, as in a trial of type N in Experiment I. Following the onset of the blank field, the subject fixates that circular field and readies the hand used to signal that the ensuing character is normal (Box 1). Then, upon the appearance of the test character in that field, the subject determines the identity of the character (Box 2) and its orientation (Box 3). Evidently, each of these operations takes about a tenth of a second: advance presentation of either piece of information alone lowers the entire reaction-time function by about 100 msec (Conditions I and O), and advance presentation of both pieces of information lowers the reaction time for an upright test character by about 200 msec (Condition B). Possibly the times to determine identity and orientation increase slightly with departure from upright (e.g., see Kolers & Perkins, 1969). However, our results provide no indication of this. Moreover, in a subsidiary experiment, one of us (LAC) found that the mean reaction time for orally reporting the identity of each of the six characters (about 550 msec) was virtually the same for all six orientations. In any case, once the identity and the orientation of the presented character have been established, the subject imagines the character rotated into its upright orientation—requiring roughly 100, 250, or 500 additional msecs for stimuli tipped by 60, 120, or 180 degrees, respectively (Box 4). The mentally righted image of the presented character is then compared with the representation of that character in long-term memory (Box 5). If there is a match, the response that has previously been readied is immediately initiated (Box 7); otherwise, control is first switched to the other hand (Box 6)—consuming (for a typical subject) another tenth of a second.

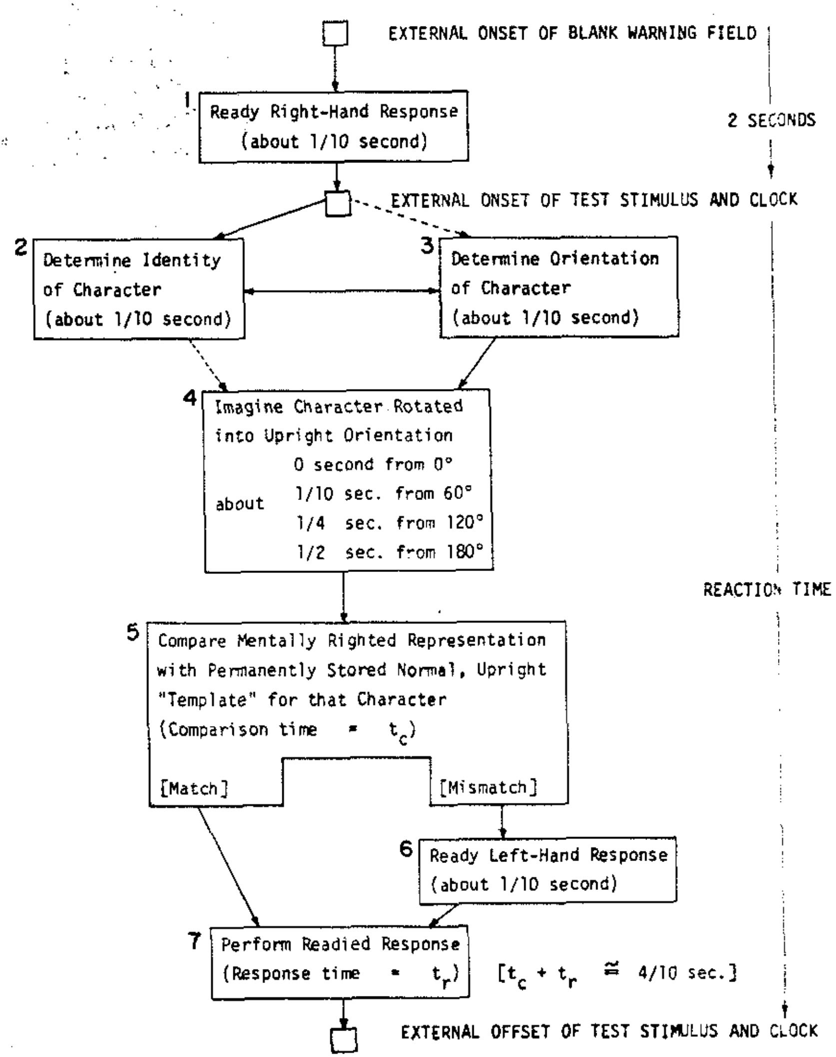

Figure 4.15 An information-processing model for determining whether a rotated alphanumeric character is normal or backward.