Memory & Cognition
1979, Vol. 7 (1), 19-28

The rate of "mental rotation" of images: A test of a holistic analogue hypothesis

ZENON W. PYLYSHYN
Departments of Psychology and Computer Science, University of Western Ontario
London, Ontario, Canada

This paper discusses the analogue-propositional distinction and argues that, given an appropriate understanding of this issue, the question of whether a particular cognitive function is analogue or not is an empirical one. As an example of how the question can be empirically investigated, the proposed analogue operation for mental rotation of images is considered. It is argued that the view that images are rotated in a holistic analogue manner should predict that rotation rate is independent of such factors as the conceptual complexity of the stimulus or of the comparison task. Two experiments are described that investigated the effects of several stimulus and task variables on the apparent rate of "mental rotation" of images in a Shepard-type task. Instead of comparing a stimulus and misoriented probe figure to determine whether they are identical (except for orientation) or mirror images, as was the case in most of previous studies, the present experiments required subjects to judge whether the misoriented probe was a subfigure of the target stimulus. The results showed that the "rotation rate" (i.e., the slope of the RT vs. angle of misorientation function) was influenced by practice, stimulus attributes, and the nature of the comparison task. In particular, when the probe was a "good" subfigure of the reference stimulus, apparent rotation rate was greater. These results are interpreted as indicating that the linear RT vs. angle relation is not due to a holistic analogue rotation of images, as had been supposed, but arises from a more articulated and piecemeal process in which analysis of the stimulus figure interacts with the comparison task.

Shepard and Metzler (1971) showed that the amount of time it took subjects to compare two figures (which were either identical or mirror images of each other) was a linear function of the angle between them. Their interpretation of this finding was that, in making the comparison, what the subjects did was "image one object as rotating into congruence with the other," subject to the restriction "that they could only do this at a certain rate without losing the essential structure of the rotated image" (Shepard & Metzler, 1971, p. 703).

Since that initial study, there have been a large number of investigations of various aspects of this so-called "mental rotation" effect. It has been widely cited as the strongest possible evidence for "analogue" processes. For example, Attneave (1974) claims, "this and other studies from Shepard's laboratory . . . show beyond any reasonable doubt that when one rotates a mental image from one aspect to another, the representation of the object is in fact going through all of the intermediate aspects in a continuous manner. I have no idea how anybody could possibly account for these results without postulating an analogue representational medium" (p. 498).

Elsewhere (Pylyshyn, 1978), I have argued that such a conclusion is at the very least premature since (1) there is no difficulty in conjuring up some articulated or propositional model that would produce the same reaction time (RT) results, and (2) simply appealing to an "analogue representational medium" does little to explain the underlying cognitive process.

This is not to say that the analogue position is not predictive. By claiming that mental rotation is in some ways like real physical rotation, we inherit some predictive potential from our knowledge of the latter. But such predictive power is highly ephemeral, inasmuch as the "rigid physical object" metaphor is a loose one at best. It is not clear exactly what aspects of the metaphorical object are supposed to transfer to the image rotation case. Some people (e.g., Cooper & Shepard, 1973) have claimed that between imaging a figure in one orientation and imaging it in another orientation, the image–like the corresponding physical object–passes through all intermediate orientations (at least in the situations that they investigated). But no one, to my knowledge, has suggested that the image must accelerate and decelerate or that the relation among torque, angular momentum, and angular velocity has an analogue in the mental rotation case. Of course, it may turn out that it takes subjects longer to rotate an object that they imagine to be heavier, thus increasing the predictive value of the metaphor. But in that case, it seems clearer that, even if it was predictive, the metaphor would not be explanatory (surely, no one

The work supported here was sponsored by the National Research Council of Canada, operating Grant A4092.

0090-502X/79/010019-10$01.25/0

believes that some images are heavier than others and that heavier ones accelerate more slowly). The reason that the latter case seems not to be explanatory is that it is clear that in this case the explanation must appeal to the subject's knowledge of the behavior of heavier objects rather than to an intrinsic property of images. It is our contention that such an appeal to tacit knowledge will be required in most cases of cognitive activity. This is the primary reason for preferring a propositional to an analogue account of mental processes. We return to this point below.

Analogue and Propositional

A number of writers (notably, Anderson, 1978; Palmer, 1978) have concluded that it may not be possible to decide between the analogue and propositional positions by appeal to behavioral data alone. I believe this conclusion is mistaken and rests on a misconception of both the primary purpose of the analogue-propositional distinction and the requirements on explanatory, as opposed to merely locally predictive, theories. An adequate defense of this belief is clearly beyond the scope of this report and has been treated elsewhere (Pylyshyn, Note 1). Since, however, the present study is presented as an example of how the distinction might be approached empirically, a brief digression on this issue is in order.

Part of the problem in distinguishing between forms of representation is one to which I alluded some time ago (Pylyshyn, Note 2) and which Anderson (1978) has emphasized more recently, namely, that in general we can only observe the behavior of a representational system, consisting of a representation plus some process that accesses it. In principle, one can modify either one to accommodate any changes in the other. Thus, in principle, one can pair either an analogue or some descriptive representation with an appropriate process, so the two behave identically. But although the two systems may make the same predictions in some particular limited context, they would very likely not be equally explanatory (recall the example of rotating imagined heavier objects cited above). One might, for example, require an interpreting process that is itself in need of further cognitive analysis, but whose behavior we can predict from common sense experience (as is often the case with the "mind's eye"). Moreover, the representation-process tradeoff cannot be carried out freely without loss of generality.[1] Mechanisms posited to account for one experimental finding must be adaptable to explain other findings, or else they are justifiably dismissed as being ad hoc. Furthermore, this kind of potential tradeoff of properties is true in every area of science. For example, we could have a very different picture of the nature of matter if we posited a different set of laws governing interactions at a distance. Concepts like mass and force also trade off, and we could make correct predictions in a limited empirical domain by adjusting hypothesized properties of one to compensate for changes in the other. But the price we would pay is measured in terms of loss of generality and explanatory power. Theoretical indeterminacy in the case of mental representation is thus no worse than it is in any area of science where inference is always to the best available hypothesis in the light of a combination of intrinsic and extrinsic criteria.

The analogical-propositional debate has, in my view, been further defocused by a lack of clear conception of what this distinction is really about. In the absence of a shared understanding of these notions, the arguments have frequently been over whether processes are continuous or discrete, or even whether mental events are "analogous" to real world events (a quite different use of the word "analogue"). Below, I present a brief sketch of what I take to be the essential point of the distinction—especially as it applies to cognitive operations.

In my view, the analogue-propositional distinction has to do primarily with whether or not one believes that the appropriate explanation of a certain behavioral phenomenon must appeal to some explicit representation, such as factual or procedural knowledge, or whether the behavior can be explained by appeal to intrinsic properties (e.g., physical, biological) of the system itself. This in turn depends on the particular description under which one wishes to understand the system's behavior. For example, an explanation of what happens when I speak might appeal to movements of my jaw, tongue, and vocal cords, as well as to intrinsic elastic properties of air. Thus, one could have an analogue model of this process by choosing a modeling medium in which a set of physical properties intrinsically behaved in a manner isomorphic to the physical properties being modeled. But an explanation of the same event under the description that I was, say, "lecturing" would have to appeal to explicit (propositional) representations of such things as grammatical rules, knowledge, beliefs, goals, and so on.

This is forced on us by two factors. One is that what we want is not simply to explain the occurrence and structure of some event. To address the phenomenon of interest, the explanation must deal with the event described as lecturing, not as lip moving or sound making, although, clearly, it was that too. The second factor is the empirical fact that there are crucial regularities and generalizations relevant to the event thus described that can only be captured by referring to such abstractions as knowledge, beliefs, goals, and the like. A further discussion of why this difference is crucial can be found in Pylyshyn (Note 1). For the time being, we can summarize the analogue-propositional distinction as follows.

When we believe that an explanation requires an appeal to explicit symbolized representations, we generally refer to it as propositional (or, equivalently, as dependent upon such things as rules, knowledge, or strategies). On the other hand, when we wish, in our explanation, to attribute certain aspects of the behavior to intrinsic properties of the biological mechanism—especially when this involves a reasonably complex set of relationships—we say that the behavior in question is produced by an analogue process whose properties are fixed by the biological medium.[2] The issue, then, is whether certain determinants of the behavior we are interested in are to be explained by appealing to properties of the biological system (to be further explicated by a biological rather than a cognitive account), or to internally represented rules and knowledge. Although both may be true, it will generally be the case that only one explains the phenomena *when the latter are described in certain terms* (e.g., cognitive ones).

I have argued elsewhere (Pylyshyn, Note 1) that many of the phenomena that appear to be reasonable candidates for an analogue explanation show what I called "cognitive penetration"; that is, the phenomena can be critically influenced by cognitive factors such as the subjects' beliefs and interpretations. Whenever that is the case, we conclude that the whole phenomenon cannot be explained by a single analogue mechanism (although analogue subprocesses may still be involved if they resist cognitive penetration).

As might be expected, this is but the tip of a philosophical iceberg. However, in specific cases, a theorist is often faced with rather clear theoretical options that can be empirically distinguished. The argument in this paper, as well as those I have made elsewhere, is not directed at the question of whether or not mental processes can be analogue; there is no argument in principle against this. The concern is always with particular models that propose certain specific analogue components.

The above analysis can be focused on the case of mental operations such as rotation by considering under what conditions a particular operation is best described as an analogue operation. One way to answer this is to say that an operation is analogue if it itself does not involve any cognitive or computational processes, but rather, is carried out primitively by the biological system; that is, its explanation (if anyone wished to pursue the matter further) would be given in terms of intrinsic physiological properties and would not need to appeal to rules, knowledge, beliefs, or other cognitive representations. An analogue operation, therefore, is one that does not require that we posit, for its explanation, any internal cognitive states or subprocesses.

Like the wired-in functions in a computer, an analogue process is computationally opaque. This, in turn, entails that there be no behavioral phenomenon whose explanation requires that we appeal to how that operation is carried out. If there were, then we would need a model of what goes on inside the analogue process, thus destroying the assumed analogicity of that process and replacing it by a more articulated one. While the new process might still contain analogue components, they would not be the same as those originally hypothesized. That is the crucial criterion and the one we shall refer to in examining the mental rotation proposal.

The claim that mental rotation of an entire image is carried out by an analogue operation (which is holistic and computationally opaque) entails the following prediction.[3] The rotation carried out in this (physical) analogue cannot depend on how complex the figure is nor on any other property of how the figure was interpreted except insofar as these are represented as initial conditions in the physical analogue (e.g., rotation could be slower for more complex figures only if increased complexity made the analogue object heavier, or the medium more viscous, or some such physical alteration). While such models of how rotation rate could be affected by the nature of the stimulus are possible, they are not plausible, since they assume that some global index is first computed from a conceptual analysis of the stimulus, and this index (although not the analysis) is used to fix an initial condition in the analogue. A more plausible view is surely that the apparent rotation rate is a consequence of the way in which the figure is processed. In any case, this type of model could not be tailored to explain such things as why, for example, some parts of a figure might appear to rotate faster than other parts, or why the rotation rate might depend on the nature of the post-rotation comparison task. This class of model would therefore predict that the apparent rate of mental rotation would not be affected by the nature of the comparison task and would be independent of properties of the stimulus (although the latter prediction perhaps does not have the same decisiveness as the former).

In contrast to the analogue approach, a view that assumes the representation to be an articulated symbol structure makes it possible for the rotation and comparison processes to be sensitive to any type of component feature, to overall conceptual complexity, and so on, as well as to the nature of the comparison task, since "rotation rate" becomes a side effect of a sequence of processing steps. In fact, it is hard to see how a model based on such a representation could avoid going through more operations for each unit of angular displacement in "rotating" a conceptually more complex figure (unless shortcut methods exist that require processing only certain landmark features), or for more difficult comparison tasks (e.g., ones in which conceptually more complex figural information enters into the discrimination). Thus the articulated representation view would naturally predict that a number of figure-specific and task-specific factors would affect the relation between RT and the angle between figures being compared, that is, that the linear RT vs. angle relation would have a different slope or "apparent rotation rate" under different experimental conditions.

To examine whether the apparent rate of rotation (i.e., slope of RT vs. angle function) varies with task, the following experiments were designed. Instead of the probe against which subjects were to compare the mentally rotated stimulus figure being either identical to the figure or its mirror image (as in experiments by Cooper & Shepard, 1973; Shepard & Metzler, 1971), the probes in these studies were potential subfigures (embedded figures) of the reference stimuli. If the stimulus figures are rotated in some holistic analogue fashion into the reference orientation in making the judgment (as to whether the probe was a part of that stimulus), the amount of time it takes to make the response should be the same function of angle for each probe and for each stimulus. In other words, although some probes may be more difficult than others to judge as being subfigures, the extent of the difference should be independent of angle. This follows from the view that after rotation, an image of the stimulus at the relevant angle is available for any subsequent comparison task. A secondary prediction from a holistic image rotation view is that so long as a figure is not too complex to be imaged, the apparent rate at which it is rotated should be independent of such stimulus attributes as complexity. Although there are few agreed-upon measures of relevant stimulus attributes, one can at least make the prediction that whatever determines the difficulty of the comparison task will remain constant over angle; that is, the apparent rotation rate should be constant over different figures. However, because it is unknown which figural attributes are relevant to the task, failure to find a difference in rate might simply be a consequence of not sampling from the relevant attribute dimensions.

The experiments reported below represent two studies from a series we have been conducting to explore the factors that contribute to apparent mental rotation rates as measured by the technique of compairing pairs of figures in different orientations. Experiment 1 extends the method to a variety of stimuli and to "embedded figure" probes. Experiment 2 focuses on one specific dimension: goodness of subfigure in relation to the stimulus.

EXPERIMENT 1

Method

Materials. The stimulus figures were designed so as to be as simple as possible and yet to lend themselves to various decompositions into subparts. The complete set of stimuli and probes is shown in Figure 1. These stimuli and probes were chosen in an attempt to sample from a variety of levels of complexity and task difficulty. A large number of additional probes and figures, which would have provided a greater range of figure and probe types, was discarded after a pilot investigation revealed they were too difficult. The four stimulus figures represent two basic figure types (triangle and quadrilateral) with a range of possible embedded subfigures. There are a total of 6, 24, 24, and 120 possible subfigures (including disjoint subfigures) for Stimuli 1, 2, 3, and 4, respectively. Probe A in each case was an outline of the stimulus figure, a condition most closely approximating the figure and mirror image probes used by previous investigators.

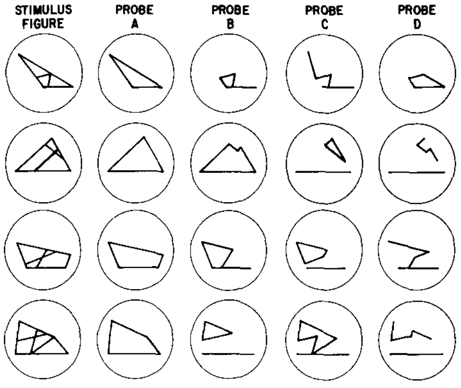

Figure 1. Stimulus figures and "true" probes used in Experiment 1. Distractor trials used either mirror images of the probes shown or the above probes paired with mirror image stimuli. Figures are numbered top down (1-4).

Each of the probes includes the baseline of the original figure in order to provide an orientation reference. Each of the 20 figures appearing in Figure 1 was used in both the form shown and its mirror image (i.e., rotated about a vertical axis). Each of the eight reference stimulus figures was paired with four "true" subfigure probes and four "false" mirror image probes. Each of these pairs of figures in turn was arranged so that the probe was at an angle of 0, 35, 70, and 105 deg clockwise to the orientation of the reference figure. This resulted in 256 different pairs for one complete replication. The stimulus pairs were drawn on cards (circle diameters were 9 cm, and centers were 9.5 cm apart) and shown in a Gerbrands tachistoscope.

Subjects. Sixteen paid subjects responded to an advertisement and were run through a practice session of 128 trials. Two of the subjects had an overall error rate of over 30% and were eliminated. One resigned because he found the task "too difficult and frustrating." The remaining 13 provided the data reported.

Procedure. After the 128-trial screening/practice session, subjects were scheduled for four additional sessions of 128 trials, each taking about 1 h on separate days. This provided 512 post-practice trials, or two complete replications. Each session consisted of four blocks of 32 trials with a brief break between blocks. In the first block of each session, an additional seven pairs were added containing "distractor" false probes that were not mirror images of true probes, in order to discourage the early induction of a strategy peculiarly suited to identifying mirror image probes.

Subjects were instructed to rotate the reference figure appearing on the left until it matched the orienting baseline of the probe figure on the right and then to indicate whether the probe figure was a true subfigure of the resulting superimposed image (which would then be both in the correct orientation and in the correct location). Instructions emphasized that, although there might sometimes appear to be a simpler way to obtain the answer, the subject was to continue using the prescribed method or to inform the experimenter if he could not do so. Instructions stressed both accuracy and speed.

There were three keys in front of the subject. When he was ready for a new trial (and when he knew, by auditory cues, that the experimenter had changed the stimulus card), the subject depressed the middle key briefly. This turned on the display and started the Hunter timer. The display stayed on until the subject indicated his choice by pressing either the right key (for the true subfigure response) or the left key (for the false response). After the last session, the subject was asked about any strategies he might have used and was also asked to indicate which stimuli and probes he had found easiest and which most difficult.

Results

The average error rate was 4%, ranging from 3.3% to 9.8% for individual subjects. Error rate in the second complete replication was one-third of that of the first replication. Error rate for true responses was 30% higher than for false responses and was correlated with angle of probe. The highest error rate was, surprisingly, for Stimulus 1, largely due to the nearly 20% error rate on Probe D. For this reason, that particular probe condition was excluded from subsequent analysis. In the case of the remaining errors, the RTs for erroneous responses were replaced by mean times for the correct responses in that cell (there were four data points in each cell in the analysis to be described below).

Only data from true pairs were used in the analysis. Graphical inspection of the false cases revealed no systematic trends and high variability. Because strategies for deciding the false cases are unknown (and likely variable), these data were not further examined.

Because different probes appeared with different reference figures, a separate analysis was carried out for each of the four reference figures shown in Figure 1 together with their mirror images. The exception to this was in the case of Probe A. Because this probe was always the outline of the reference figure, it was possible to consider all the Probe A data in one analysis. This provides a test for the relation between stimulus figure type and apparent rotation rate. The results are shown in Figure 2. Because an analysis of variance revealed highly significant stimulus, angle, practice, and Stimulus by Angle interaction effects (all ps < .005), it was decided to analyze the data directly for the effects of primary interest, that is, the relation between these factors and the slope of the RT vs. angle curves appearing in Figure 2. This was done using an orthogonal components trend analysis, analyzing for linear and quadratic trends across angle.

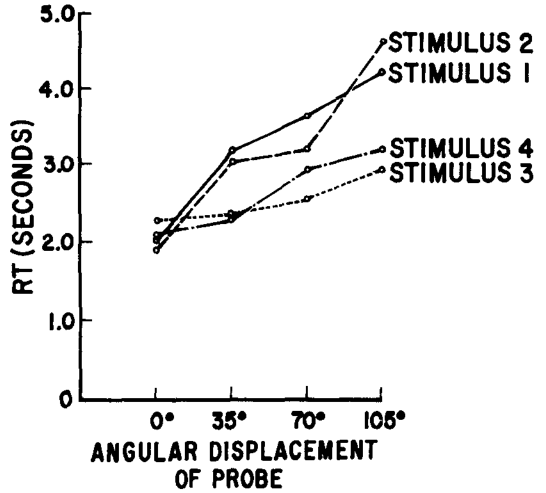

Figure 2. Mean reaction time functions for the four different stimulus figures (plus their mirror images) using the outline probes (Probe A) only.

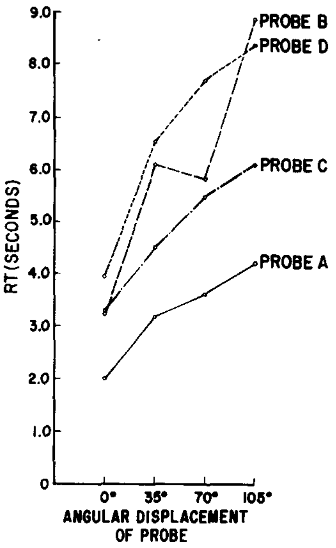

Figure 3. Mean reaction time functions for Stimulus 1 with the four probe types shown in Figure 1.

Trend analyses confirmed the statistical reliability of the linear effects displayed in Figure 2. There was a highly significant overall linear trend [F(1,12) = 96, p < .001], as well as a significant interaction between reference figure and linear trend [F(3,36) = 5.6, p < .005]. No interactions involving practice (first vs. second half of the experiment) with linear trend and no quadratic trends approached significance. Interactions with the linear trend component of the angle factor represent differences in slopes of the best-fitting straight line through the RT vs. angle function. The analysis thus reveals that the four figures produced different apparent rotation rates when assessed against their outline probes. Actual rotation rates (reciprocal of slope) were estimated as 42, 57, 1,250, and 106 deg/sec for Stimulus Figures 1, 2, 3, and 4, respectively. It appears that in assessing whether the probe was a correct or mirror image outline of the reference figure, the type of figure (particularly whether triangle or quadrilateral) affected the apparent rate of rotation.

The relation between slopes and individual probes was assessed separately for each reference stimulus figure (taken together with its mirror image). Results are displayed in Figures 3-6. Four separate analyses of linear trends were carried out. (All trend results reported as significant also showed a significant interaction with angle on an analysis of variance, a condition considered to be a prerequisite for the use of trend analysis.) All four stimuli showed highly significant overall linear trends (df = 1,12, p < .001), and all but Reference

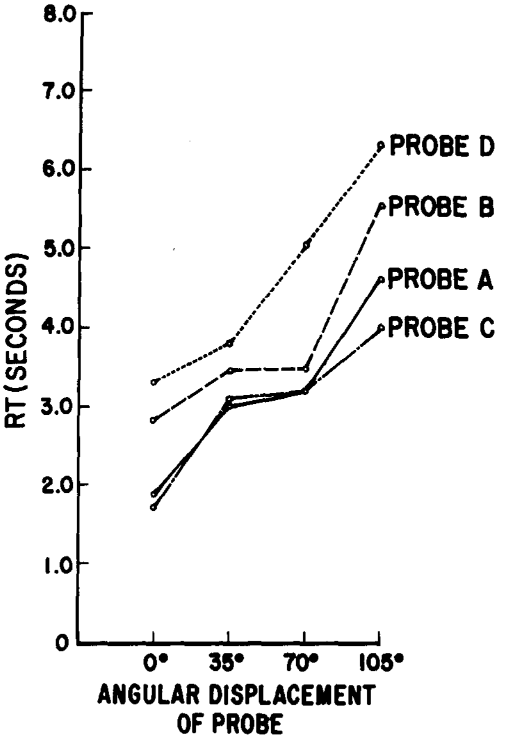

Figure 4. Mean reaction time functions for Stimulus 2 with the four probe types shown in Figure 1.

Stimulus 2 showed a significant probe effect on linear trend [Stimulus 1, $F(2,24) = 6.95$, $p < .005$; Stimulus 3, $F(3,36) = 3.35$, $p < .03$; Stimulus 4, $F(3,36) = 3.35$, $p < .03$]. All but Stimulus 1 also showed a significant effect of practice on linear trend [Stimulus 2, $F(1,12) = 9.02$, $p < .01$; Stimulus 3, $F(1,12) = 7.38$, $p < .02$; Stimulus 4, $F(1,12) = 9.51$, $p < .01$]. In each case, the slope decreased with practice (i.e., the apparent rate of rotation increased). With one exception, again no quadratic component of trend (including the overall effect over angle) approached significance. The one exception was the overall angle effect for Reference Stimulus 2 [$F(1,12) = 8.3$, $p < .05$].

Discussion

Two main findings emerge from these analyses: The apparent rate of rotation varies with attributes of the figure being rotated (as measured by the Probe A conditions), and, for three out of the four reference figures, the apparent rate depends on the particular subfigure used as probe and increases with increasing practice. More specifically, it appears that the apparent rotation rate is higher for the quadrilateral figures (3 and 4) than for the triangular ones (1 and 2) and is highest for the (easier) outline probe (Probe A) and lowest for the more difficult open line-segment probes (Probes 1-C, 3-D, and 4-D), with other probes yielding intermediate rates.

In other respects, these results raise additional questions. For one thing, the variability of the data was rather high. Not only did one of the reference figures fail to yield the same results as the other three, but between-subjects variability was such that graphs for individual subjects revealed very little in the way of a consistent pattern (although, clearly, the variability was not enough to mask the statistical reliability of the group means).

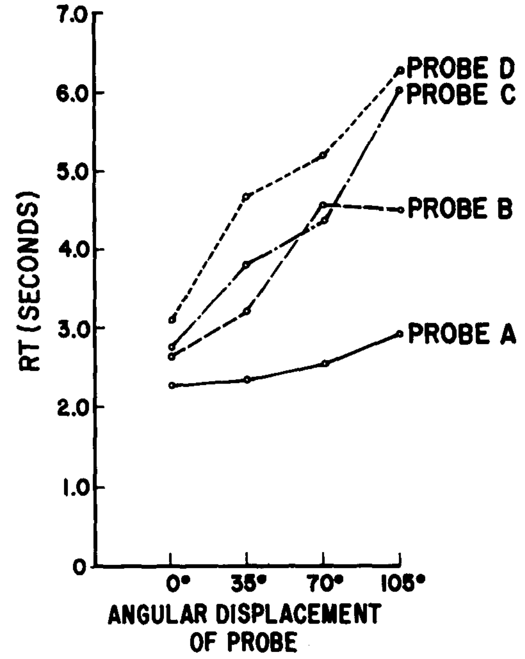

Figure 5. Mean reaction time functions for Stimulus 3 with the four probe types shown in Figure 1.

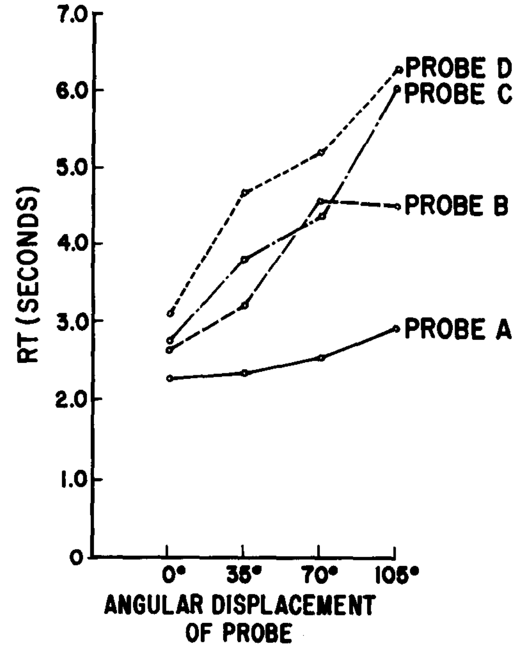

Figure 6. Mean reaction time functions for Stimulus 4 with the four probe types shown in Figure 1.

Such variability might have been to some extent anticipated given the nature of the task. Figures were sufficiently complex in the way they could be "parsed" or restructured that one might expect a variety of different strategies for recognizing an embedded subfigure. In a similar task, although without the additional requirement of rotating the image, Reed (1974) also reported considerable variability. The additional requirement of rotating an image of the reference stimulus was also difficult for subjects to adhere to uniformly. In the debriefing, one subject admitted frequent failure to rotate, another said he sometimes rotated the probe, and five others spontaneously suggested they occasionally used some iterative process of attending to fragments of the reference figure, and falling back on rotations, and rescanning the reference figure when the initial glance at the probe failed to reveal the answer. There is reason to suspect that this sort of thing happens in most, if not all, mental rotation studies. For example, the eye movement records collected by Just and Carpenter (1976) and by Metzler and Shepard (1974) on the original Shepard and Metzler (1971) three-dimensional figures (which produce the cleanest linear time-angle relations) suggest that even there, subjects used a serial processing strategy. Thus, it seems likely that there was nothing special about our experimental set-up that led to this. Such a heterogeneous set of reported strategies might be expected in our case, in spite of the explicit enjoinder to rotate, if one takes the view that the encoding of the reference figure is some sort of structural description rather than a holistic image or analogue. In such a case, there is much room for variability in the way the figure is conceptually structured or parsed and, thus, considerable variability in the subsequent rotation and matching process. If the probe happens to be a subfigure of the reference stimulus that received a unitary interpretation in the original encoding, the comparison task might be relatively easy and quick. In contrast, if lines making up the probe subfigure had been encoded originally in relation to other lines not in the probe, the comparison might require going back to the stimulus figure to construct a new parse or structural description.

Such considerations suggest that what is needed is a more systematic and independently motivated way of constructing figures and subfigure probes. Such a method might also help to get away from some of the remaining puzzling aspects of the present results, namely, the question of why certain of the figures and probes appear to be more difficult than others. The initial figures were designed to sample from a range of figural complexity and comparison difficulty. However, some of the initial expectations as to difficulty were not borne out. For example, 10 out of 13 subjects rated Stimulus 1 (expected to be one of the simplest figures) as the most difficult, and this was confirmed by error data. In fact, the two triangular figures that were thought to be the simplest yielded the longest RTs and the slowest rotation rates. Clearly, we were misled in our intuitions of what constitutes the relevant aspect of complexity for this particular task and using the particular probes we had selected.[4] A more clearly interpretable set of findings might be hoped for if a more systematic process of constructing figures and probes could be found.

Experiment 2 was directed at solidifying the general findings reported above by using reference stimuli for which subfigures with empirically established "goodness" measures are available. These were selected from a set of figures developed by Palmer (1978) for studying mental manipulation of figural information.

EXPERIMENT 2

Method

Materials. Palmer (1978) developed and validated in several different ways a measure of the "goodness" of a subfigure in relation to a reference figure in which it is embedded. The measure takes into account how well integrated the elements of the subfigure are in relation to each other (using a number of empirically weighted "Gestalt" criteria), in contrast to how closely the same elements are related to other reference figure elements not in the subfigure. Thus, the index is a relative one: The same subfigure can yield a different measure, depending on the reference figure in which it is embedded. Palmer found that the measure predicted not only subjective ratings but also the ease with which figures can be mentally synthesized from given subfigures. In the present study, we adopted four of Palmer's figures together with two subfigures that can be classed as "good subfigures" and two that can be classed as relatively "bad subfigures" of each of the four reference figures. The complete set of stimuli (excluding false probes) is shown in Figure 7. Reference figures were selected from a set consisting of six randomly arranged line segments drawn between points in a square 3 by 3 dot matrix. The three-line subfigure probes were constructed to maximize the goodness measure or to give a relatively low goodness measure by Palmer's principles. False probes were identical to the true ones shown in Figure 7 but

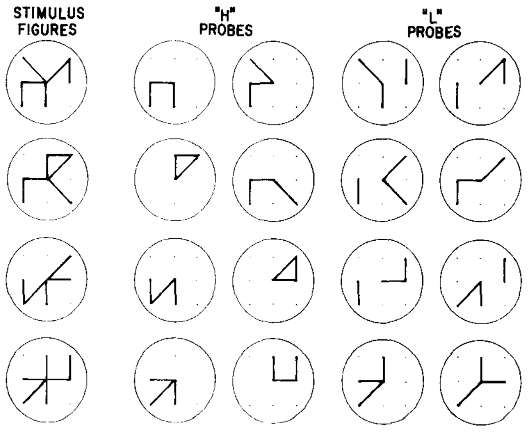

Figure 7. Stimulus figures and "true" probes used in Experiment 2. The "H" probes are "good" subfigures of the reference stimulus and the "L" probes are poor subfigures, using Palmer's (1977) criteria.

were paired with inappropriate reference figures. To keep the false probes sufficiently difficult, they were chosen so that two of the three lines matched reference figure segments. There were equal numbers of false probes that were connected (and thus resembled "H" or "high-goodness" probes) and partly connected or not as compact (and thus generally resembled the "L" or "low-goodness" probes).

The figures were drawn the same size as those of Experiment 1. Instead of the base of the figure appearing in the probe as an orientation reference, the nine-dot matrix aided by a small arrow at the top of the probe served to unambiguously indicate the orientation and relative location of the probe subfigure.

Subjects. A pilot study showed that highly stable data could be obtained using these stimuli, so only four subjects were used. These subjects had all taken part in Experiment 1.

Procedure. As in Experiment 1, four 1-h sessions with 128 trials in each were run. Unlike Experiment 1, additional false distractor trials and mirror images of stimuli appearing in Figure 7 were not used. Thus, the total 512 trials of the experiment represented four complete replications, rather than two as in the previous experiment.

Apart from the differences noted above, the procedure was identical to that of Experiment 1.

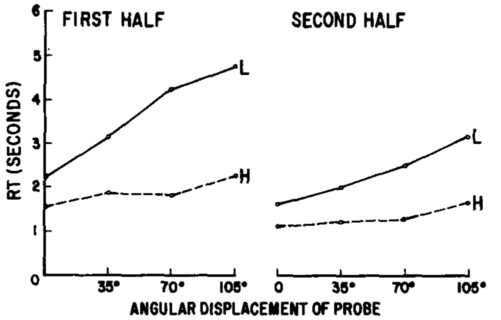

Figure 8. Mean reaction time functions for "good" (H) and "poor" (L) subfigure probes obtained in Experiment 2. Curves on the left are for the first two blocks of trials and curves on the right are for the last two blocks of trials of the experiment.

Results

The error rate was similar to that of Experiment 1: The mean was 5.1% and the range for individual subjects was from 2.3% to 8.8%. This rate decreased from 8.4% in the first session to just over 3% in the fourth session. Nearly 9.5 times as many errors were made in the L-probe condition as in the H-probe condition, there was a 33% higher error rate for false probes than for true ones, and a low positive correlation between errors and angle (showing that RT differences did not arise from a speed-accuracy tradeoff).

The treatment and analysis of the RT data was the same as for Experiment 1, except that this time it was possible to include all the data in one analysis since we had a common basis for grouping probes across all stimuli (i.e., H vs. L probes). The results are shown in Figure 8 separately for the first two and last two sessions of the experiment. A linear trend analysis confirmed the statistical reliability of the differences apparent in the figure: Rotation rate was affected by probe goodness [$F(1,3) = 102.7$, $p < .002$] and by practice [$F(3,9) = 131.9$, $p < .002$], although the interaction of these two factors did not significantly affect the linear trend over angle. Apparent rotation rate was faster for the more easily extracted H subfigures (121 deg/sec) than the L subfigures (54 deg/sec). Results for individual subjects followed the same overall pattern.

GENERAL DISCUSSION AND CONCLUSIONS

These two experiments provide clear evidence that the apparent "rate of mental rotation" (i.e., the slope of the RT vs. angular displacement curve) is a function of (1) practice, (2) intrinsic properties of the stimulus, and (3) the nature of the comparison task carried out on the rotated image. The influence of practice on rotation rate is found routinely in studies such as these (Shepard, Note 3), although it has not generally been reported in the literature, since published results are invariably obtained from highly practiced subjects using overlearned stimuli.

The influence of stimulus attributes might have been expected since there is precedence for this finding in the literature. For example, while the rotation rate for drawings of three-dimensional block figures was estimated to be around 60 deg/sec (Shepard & Metzler, 1971), the rate for rotating familiar shapes such as block letters appears to vary anywhere from 164 to 800 deg/sec (Cooper & Shepard, 1973, p. 109). Furthermore, Hochberg and Gellman (1977) showed that the apparent rotation rate depends on such figural attributes as the presence of what they refer to as "landmark features." The results of Experiment 1 (shown in Figure 2) confirm the view that the rate does depend upon properties of the figure. But these results also caution against any simple interpretation of which figural attributes contribute to this dependence. In particular, higher values on such indices as the number of vertices in the stimulus figure or the number of ways in which the figure could be decomposed into subfigures, which we had initially thought of as measures of stimulus complexity, produced higher, rather than lower, rotation rates and lower "subjective complexity" ratings. Clearly, such indices can be rendered impotent as predictors in this task by other properties of the task situation. What appears to be relevant here is the complexity of the task as a whole rather than the complexity of the stimulus alone. For example, what may have made Stimulus 1 the one rated most difficult and the one producing the slowest rotation rate is that it was the most nearly symmetrical in outline and, hence, the most similar to its mirror image distractor.

Observations such as these may explain why Cooper and Podgorny (1976) failed to find an effect of complexity on rotation rate. In their studies, Cooper and Podgorny used random polygons and found that the number of points in the polygon (which could be 6, 8, 12, 16, or 24) did not affect the rotation rate.

But, as we found in Experiment 1, the number of vertices is not as important as the complexity of the comparison task itself. Thus, the nature of the comparison, and, consequently, of the distractors used, is of central importance. Now, Cooper and Podgorny used small random perturbations of the original polygons as distractors. However, they selected distractors for the various stimulus figures that were rated by subjects as being equally similar to the target stimulus. But, if these similarity ratings are correlated with comparison difficulty (as is generally the case in comparison experiments), then the selection of distractors was in fact made in such a way as to equalize the complexity of the comparison task over the five levels of stimulus complexity. On the basis of our present discussion, we would be led to predict no difference in task difficulty, and, consequently, no difference in rotation rate–just as Cooper and Podgorny indeed found.

The finding, which was especially clear in Experiment 2, that the nature of the comparison task affects rotation rate has important implications for the nature of the processing in this situation. In particular, it provides strong evidence that the process is not one in which a stage of holistic analogue rotation of the image is followed by an independent stage of comparison, as suggested, for example, by Cooper and Shepard's (1973) model. If that were the case, then the nature of the probe might be expected to affect the intercept but not the slope of the RT vs. angle curve. Results such as those depicted in Figure 8 make it clear that if there is anything that might be called "rotation" in this situation, the whole figure is not carried along rigidly. Rather, there must at least be some analysis of the original stimulus and some piecemeal "rotate and compare" subprocesses. Such a view receives independent corroboration from Hochberg and Gellman's (1977) "landmark feature" effect, as well as from Just and Carpenter's (1976) more fine-grained analysis of eye movements involved in such rotation experiments.

As pointed out earlier (Footnote 3), such evidence as this does not exclude the possibility that some analogue process is involved in the manipulation of component parts of the image. It only argues against the holistic rotation view. Nevertheless, it is precisely this holistic version that is so intuitively appealing and that leads writers like Attneave (1974) and Kosslyn and Pomerantz (1977) to decry the "unnaturalness" of articulated or propositional views. The more carefully we examine phenomena, such as the mental rotation findings, the more we find that the informally appealing holistic image-manipulation views must be replaced by finer grained piecemeal procedures that operate upon an analyzed and structured stimulus using largely serial, resource-limited mechanisms. By repeated application of experimental methods such as those used in the present studies, the building blocks of the process explanation of imagistic phenomena can be progressively refined. By the time the reduction of global phenomena such as "mental rotation of images" to a sufficiently primitive process form has been achieved, the resulting model may contain few, if any, components worthy of the name "analogue." In any case, it seems clear that the appeal of such informal accounts as those given by subjects who claim, for example, to "rotate an image" will have been seriously eroded in a more adequate explanation of the phenomena.

REFERENCE NOTES

1. Pylyshyn, Z. W. *The explanatory adequacy of cognitive process models.* Paper presented at the M.I.T. workshop on Cognitive Representation, Massachusetts Institute of Technology, January 1978 (mimeo).
2. Pylyshyn, Z. W. The symbolic nature of mental representations. Paper presented at a symposium on *Objectives and methodologies in artificial intelligence*, Australian National University, Canberra, May 1974 (mimeo).
3. Shepard, R. N. Personal communication, 1978.
4. Pomerantz, J. Personal communication, 1978.

REFERENCES

Anderson, J. R. Arguments concerning representations for mental imagery. *Psychological Review*, 1978, **85**, 249-277.

Attneave, F. How do you know? *American Psychologist*, 1974, **29**, 493-499.

Cooper, L. A., & Podgorny, P. Mental transformations and visual comparison processes: Effects of complexity and similarity. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, **2**, 503-514.

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

Hochberg, J., & Gellman, L. The effect of landmark features on "mental rotation" times. *Memory & Cognition*, 1977, **5**, 23-26.

Just, M. A., & Carpenter, P. A. Eye fixations and cognitive processes. *Cognitive Psychology*, 1976, **8**, 441-480.

Kosslyn, S. M., & Pomerantz, J. R. Imagery, propositions, and the form of internal representations. *Cognitive Psychology*, 1977, **9**, 52-76.

Kosslyn, S. M., & Shwartz, S. P. A simulation of visual imagery. *Cognitive Science*, 1977, **1**, 265-295.

Metzler, J., & Shepard, R. N. Transformational studies of the internal representation of three-dimensional objects. In R. L. Solso (Ed.), *Theories of cognitive psychology: The Loyola symposium*. Hillsdale, N.J: Erlbaum, 1974.

Palmer, S. E. Hierarchical structure in perceptual representation. *Cognitive Psychology*, 1977, **9**, 441-474.

Palmer, S. E. Fundamental aspects of cognitive representation. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. Hillsdale, N.J: Erlbaum, 1978.

Pylyshyn, Z. W. What the mind's eye tells the mind's brain: A critique of mental imagery. *Psychological Bulletin*, 1973, **80**, 1-24.

Pylyshyn, Z. W. Imagery and artificial intelligence. In W. Savage (Ed.), *Perception and cognition: Issues in the foundation of psychology. Volume IX of the Minnesota studies in the philosophy of science*. Minneapolis: University of Minnesota Press, 1978.

Reed, S. K. Structural description and the limitations of visual images. *Memory & Cognition*, 1974, **2**, 329-336.

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

NOTES

1. Anderson (1978) claims that a system can always be constructed that not only mimics a given system with a different representation but also goes through the same computational steps in doing so. But his proof rests on the untenable assumption that one is free to designate any arbitrary subprocess in the mimicking system as a single computational step. Such a subprocess could (and in his proof it does) require the computation of a translation between the two representations in the course of mimicking each step of the original process. (Anderson's "existence proof" relies on the possibility of mimicking one transforming operation T by the operation T*, constructed by translating the representation in the mimicking system to that in the mimicked system, applying T, and then translating back again.) But, if the translation itself has a computational complexity that varies with its input (as would be the case even with such simple examples of representations as different number systems), the individual steps of the original process cannot be taken as being mimicked by individual steps in the mimicking process (and because the complexity is variable, no uniform speed up of the computation will resolve this difficulty). Now, because in the mimicking process complexity will vary in ways not attributable to the number of steps taken (each step corresponding to a variable length computation), such a process will no longer be consonant with various behavioral indices of complexity (such as RT), which invalidates Anderson's claim. The issue of when we are entitled to count a function as a single computational step is addressed theoretically in Pylyshyn (Note 1) and empirically in the present report.

2. This is a slight simplification of the position presented in Pylyshyn (Note 1). What we have characterized here is the notion of a computational primitive. It is only when such a primitive behaves in a way that encapsulates a whole system of relationships, such as those formalized in euclidean or metrical axioms or Newtonian laws, that we usually speak of it as an analogue process. Thus, an operation for comparing two symbols for identity, even if it were primitive, would not ordinarily be referred to as an analogue operation, although there is no difference in principle between it and, say, the proposed operation of image rotation–only a difference in formal complexity.

3. Note that the argument here is not that the Shepard phenomenon involves no analogue processes of any kind. The concern here is with a particular analogue process, one that rotates the entire image. Other hybrid analogue-propositional proposals, such as those of Anderson (1978) and Kosslyn and Shwartz (1977), are not being tested in the studies about to be described. On the other hand, the great attraction of the mental rotation idea arises from the simple holistic account. To deal with the hybrid models, one would have to apply an experimental procedure such as the one to be discussed below to what the model assumes are the component parts in order to see whether they, in turn, seem to be rotated holistically. What makes the hybrid proposals seem unpromising in the long run is that, in order to apply rotation to elements, the latter must be geometrical pieces. But the arguments presented by Pylyshyn (1973, 1978) suggest that representations of objects are of nonuniform grain and factor out perceptual qualities like color, size, identity, particular relations such as "above" or "left of," and general abstract features such as "elongated." It is difficult to see how parts of such a representation can be analogically rotated. Nonetheless, the present investigation is limited to providing evidence only against the pure analogue view.

4. Pomerantz (Note 4) has suggested that Stimulus Figures 3 and 4 might have ended up being easier on the average because each of the probes for these stimuli preserves a salient stimulus landmark, namely, the prominent upper left vertex of the quadrilateral. This may explain the discrepancy between our a priori complexity judgments and objective performance on the task.

(Received for publication June 30, 1978;
revision accepted December 12, 1978.)