CHAPTER 37

VISUAL FUNCTIONS OF MENTAL IMAGERY

RONALD A. FINKE
Department of Psychology, State University of New York at Stony Brook, Stony Brook, New York

ROGER N. SHEPARD
Department of Psychology, Stanford University, Stanford, California

CONTENTS

1. **Visual Imagery in Comparison, Identification, and Resolution** 37-3
    1.1. Comparisons Among Objects When They Are Imagined or Actually Observed, 37-3
        1.1.1. Outline Shapes, 37-3
        1.1.2. Familiar Faces, 37-5
        1.1.3. Numerical Symbols, 37-5
        1.1.4. Colors, 37-5
        1.1.5. Sizes and Lengths, 37-7
        1.1.6. Influences of Nonstimulus Factors on Comparisons, 37-8
    1.2. Interpretive and Structural Influences on Visual Representations, 37-8
        1.2.1. Influences Revealed in Judged Relations Among Objects, 37-8
        1.2.2. Influences Revealed in the Imagining of Individual Objects, 37-9
    1.3. Imagining Objects and Visually Identifying Them, 37-9
        1.3.1. Imagining an Object Formed of Separately Presented Parts, 37-9
        1.3.2. Imagining an Object from a Description of Its Parts, 37-10
        1.3.3. The Time Required to Imagine an Object or Scene, 37-13
    1.4. Visual Resolution for Imagined and Observed Objects, 37-15
        1.4.1. Verification Time for Features of Objects Imagined at Different Sizes or Distances, 37-15
        1.4.2. Acuity for Objects Imagined or Observed in the Visual Periphery, 37-15
    1.5. Visual Aftereffects of Imagination and Perception, 37-17
        1.5.1. McCollough Effect Obtained When Bar Gratings Are Only Imagined, 37-17
        1.5.2. Some Unresolved Issues Concerning the Imagery Form of the McCollough Effect, 37-18
    1.6. General Issues Concerning Information Representation in Imagery, 37-18
    1.7. Summary, 37-18
2. **Visual Imagery in Determining Distance, Location, and Visual–Motor Control** 37-19

Preparation of this chapter was in part supported by National Science Foundation Grant No. BNS 80-05517 to Roger Shepard and by a Sloan Foundation Postdoctoral Fellowship awarded to Ronald Finke. We thank Michael Kubovy, Lloyd Kaufman, and Kenneth Boff for their extensive and helpful editorial suggestions on the manuscript, Teresa Putnam for her typing and retyping of numerous revisions, and our many colleagues who provided useful information and reports.

that an object's location is remembered descriptively or symbolically than by assuming a strict functional equivalence between perceived and imagined location. It is still possible, though, that even when the locations of objects are distorted in memory they may be imagined and judged as if they were actually being observed in those altered locations.

2.2. Mental Scanning of Imagined and Observed Arrays

2.2.1. Scanning an Imagined Picture or Array. Memory for the relative locations of objects and for details on pictures of objects is facilitated by scanning over the objects and their features (e.g., Loftus, 1972; Nelson & Loftus, 1980). In addition, information in a fading "iconic" image is retrieved by scanning the image along particular directions (e.g., Neisser, 1967; Sperling, 1960). Such findings raise the question of whether an imagined object or array, too, can be scanned, to retrieve information about the relative locations of objects (or features of objects) no longer physically present.

Kosslyn (1973) had subjects inspect drawings of objects and then verify from memory whether the objects contained a specific part. When the subjects were also instructed to form mental images of the object, and to start by mentally "focusing" on one end of the imagined object, verification time increased in direct proportion to the distance along the object between the part and the point of focus. The result was interpreted as evidence that imagined objects can be "scanned" in much the same way as physically presented objects, with more time required to scan across greater distances.

Lea (1975), however, pointed out that distance between features and number of other intervening features had been confounded in the Kosslyn (1973) study. Accordingly, Kosslyn, Ball, and Reiser (1978) then varied the distance between objects independently of the number of intervening objects. Subjects learned the locations of three letters on a straight line. After mentally focusing on one end of the line, a letter was named, and the subjects scanned along the line to where they had imagined the letter and indicated whether that letter was upper- or lowercase. Reaction time increased both with the distance and with the number of intervening letters along the path of the scan. See Figure 37.13(a). Likewise, in an earlier experiment, Weber and Harnish (1974) found that when subjects were to indicate the location in a lowercase printed word of a letter extending above the others (e.g., letters such as *b*, *d*, *f*, etc., as opposed to letters such as *a*, *c*, *e*, etc.) their reaction times increased with the length of the word and with the position of the taller letter in that word. Moreover, the same sort of increase was obtained when the subjects imagined the word as when it was actually presented.

In the second experiment of the Kosslyn et al. (1978) study, subjects learned the locations of objects on a map, arranged so that the interobject scan paths would never include other objects along the same path [Figure 37.13(b)]. Reaction time was almost perfectly correlated with distance separating the object [Figure 37.13(c)]. Kosslyn et al. concluded from these experiments that it takes longer to scan greater distances between imagined objects and, because additional time is needed to "inspect" nontarget objects encountered along the scan path, it also takes longer to scan over greater numbers of imagined objects.

These findings suggest an equivalence between the scanning of imagined and visually perceived objects, but they are vulnerable to the objection that experimental subjects could intentionally delay their responses by greater amounts when asked to scan between objects remembered to be farther apart (Richman, Mitchell, & Reznick, 1979). Subjects in an experiment by Mitchell and Richman (1980), for example, were given the same map and instructions used by Kosslyn and were then asked to guess the time required to scan between all possible pairs of remembered objects. Their estimated scanning times were highly correlated with interobject distance (although the estimated rate of scanning was much slower than that found by Kosslyn et al., 1978). Clearly, various control experiments are required to decide between alternative hypotheses. In particular, tacit knowledge and eye movement hypotheses could also have predicted many of these results. Subjects may have attempted to simulate perceptual scanning, drawing on tacit knowledge about how long it takes to scan between certain distances, or may simply have moved their eyes between the remembered locations of the imagined objects, thereby increasing their response time as the interobject distance increased. Nevertheless, subjects' expectations do not seem to provide the most likely explanation for some of the scanning results (see Kosslyn, 1981; Kosslyn, Pinker, Smith, & Shwartz, 1979). The hypothesis that the results arose from experimentally introduced bias is less likely in light of recent findings that the linear dependence of scanning time on distance is still present even when the experimenter is led to have contrary expectations (see Jolicoeur & Kosslyn, Note 5).

2.2.2. Scanning an Imagined Three-Dimensional Scene. The findings for the mental scanning of imagined pictures and maps have been extended to the mental scanning of objects imagined at various positions in three-dimensional space. Pinker and Kosslyn (1978) had subjects learn the locations of objects suspended in a box, and, as in the previous picture or map scanning studies, instructed them to imagine this scene and to imagine themselves visually scanning it. Their scanning times were highly correlated with the three-dimensional distances between the objects in the scene. Moreover, when the subjects were first instructed to imagine that selected objects were moved up or down by specified amounts in the remembered configuration, their reaction times were linearly related to the interobject distances in these altered configurations. This suggests that the rate of mental scanning of the imagined scene was constant, with resulting scanning times proportional to distances in three-dimensional space, even after the subjects imagined rearranging the objects.

2.2.3. Scanning an Imagined Two-Dimensional Projection of a Three-Dimensional Scene. In an investigation of the perspective properties of an imagined three-dimensional scene, Pinker (1980a) instructed subjects to imagine themselves scanning these scenes through the sight of a rifle as it moved its aim from the projection of one object to the projection of the other on a plane perpendicular to the original viewing direction. When they imagined scanning the scene in this manner, reaction time was proportional to the distances between the objects in the projection plane. In a variant of the task of *mental rotation* (considered in Section 3.1), the subjects were instructed to imagine that they were looking at the three-dimensional configuration of objects from the side or top of the transparent box. This time, when they imagined scanning the scene with the rifle sight, reaction times were proportional to the distances between the objects in the new projection planes. Pinker's findings showed that it is possible to imagine oneself or a point moving between objects in a three-dimensional scene or to

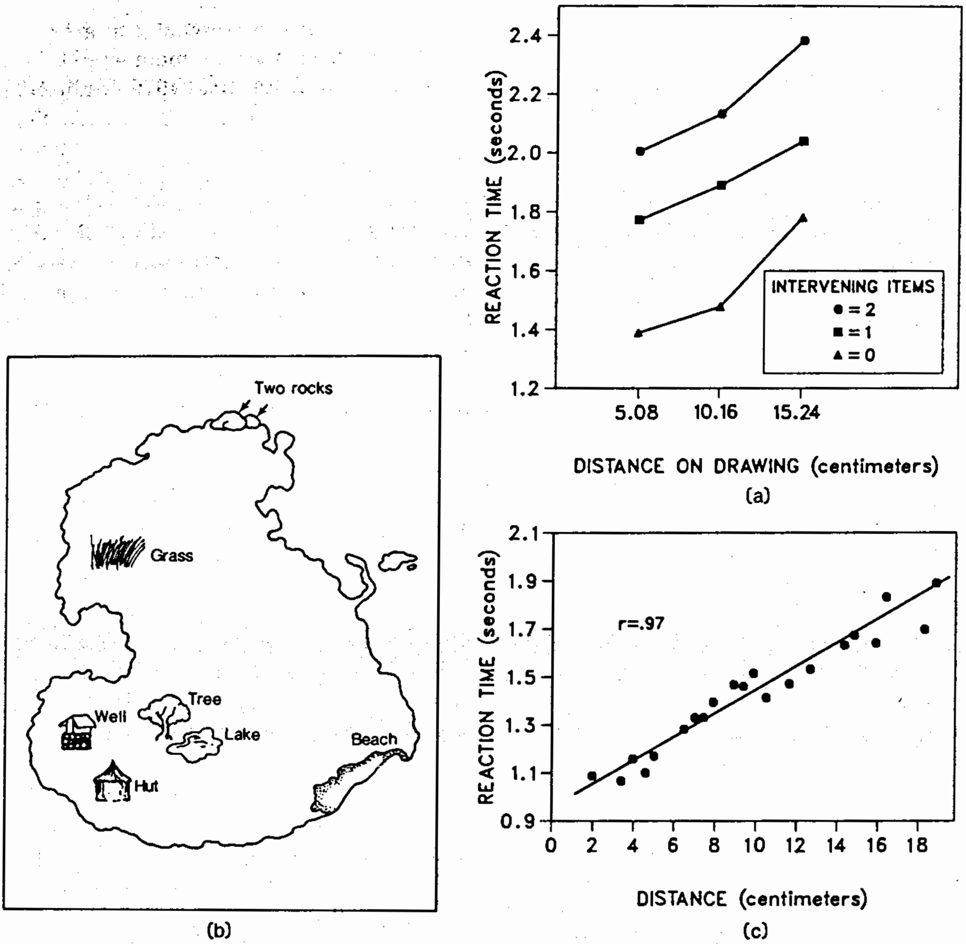


Figure 37.13. Dependence of time to scan between objects mentally in an imagined spatial layout on the distance between the objects in that layout. (a) Time to scan between imagined letters on a line containing zero, one, or two intervening letters, plotted as a function of distance between the terminal letters. (b) Fictitious map learned and then imagined by subjects. (c) Mean time to scan mentally between designated landmarks on the imagined map, plotted as a function of distance between those objects. The linear increase of scanning time with distance suggests an equivalence between the imagined scanning of objects and the scanning of objects that are actually observed. (From S. M. Kosslyn, T. M. Ball, & B. J. Reiser, Visual images preserve metric spatial information: Evidence from studies of image scanning, *Journal of Experimental Psychology: Human Perception and Performance, 4*. Copyright 1978 by American Psychological Association. Reprinted with permission.)

imagine oneself or the point moving between the projections of those objects on various two-dimensional planes, including those planes corresponding to points of view not previously adopted.

The Pinker (1980a) study also included perception conditions in which subjects actually observed the array of objects as they were scanning it. When the subjects were asked to imagine themselves scanning between the presented objects in the three-dimensional array, response times were proportional not only to the objects' distances of separation in that three-dimensional array, but also to those distances as projected onto a plane perpendicular to the line of sight. When subjects were asked to imagine themselves scanning the presented objects through an imaginary rifle sight, response times were very similar to those obtained in the corresponding imagery condition, proportional only to the distances in the plane. To consider the possible contribution of eye movements to these measured times, subjects in a control experiment were told simply to move their eyes from one object to the other. Since the eye movement times were much shorter than the times previously taken to scan either the observed or the imagined array, and were proportional to the two-dimensional but not the three-dimensional distances, it is unlikely that eye movements could have been responsible for all the reported findings.

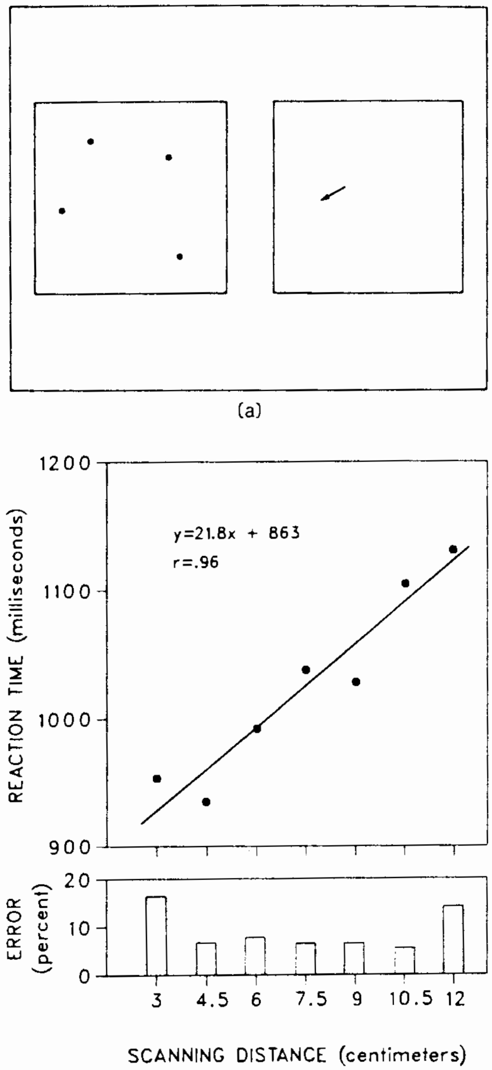
2.2.4. Spontaneous Scanning to Determine Position, Direction, or Distance. In Section 2.2.1 we consider evidence that the linear reaction time functions obtained in imagery scanning experiments might be due to subjects' performance expectations. Pylyshyn (1981) has advanced the related argument that tacit knowledge about relations between distance, time, and velocity or about changes in the appearance of objects with changes in distance and vantage point might also explain these findings. Such possibilities raise the question of whether the mental scanning of an imagined scene or picture is ever used spontaneously as a method of retrieving information from memory, or of judging spatial relationships.

One possible spontaneous use of mental scanning might be to verify that objects that can no longer be seen lie along particular directions. For example, several experiments suggest that people can use imagery to tell when someone is pointing to objects that are out of view (see Attneave & Farrar, 1977; Attneave & Pierce, 1978). Other experiments, however, have shown that, when subjects learn the locations of objects on a map and then judge from memory whether one of the objects lies in a certain direction with respect to the other, reaction times are largely independent of the relative distances between the objects (see Bannon, 1981; Wilton, 1979), which suggests that such judgments are not made by imagining that one is scanning a scene.

Finke and Pinker (1982, 1983) tested the possibility that the mental scanning of a scene might prove useful in judging relative directions to the previously learned locations of objects from new locations designated unexpectedly. Subjects were shown simple dot patterns, and after a 2-sec delay were presented an arrow in an unexpected location and orientation as shown in Figure 37.14(a). Their task was to decide whether the arrow pointed at any of the dots they had just seen. Although no mention was made of mental imagery or of scanning, reaction time increased linearly with distance between the arrow and the dot to which it pointed, at a rate similar to that obtained when instructions are given to imagine scanning a picture [Figure 37.13(c)]. Since the subjects in this experiment were never told to use imagery, and, more important, would not have had sufficient time to memorize the arrow–dot distances before having to make their judgments, these results are less likely to have been due to subjects' expectations or to tacit knowledge about relations between scanning times and distance (see also Pinker, Choate, & Finke, Note 6, for evidence that these findings may be extended to longer retention intervals).

When advance information was provided about the location of the arrows, however, reaction time in this task was not significantly correlated with distance (Finke & Pinker, 1983). Whereas in the former case most subjects reported having made their judgments by imagining they were scanning the patterns, in this case they reported having been able to determine the "correct" directions from the arrow locations to the dots before the arrows were presented. In agreement with earlier reports by Wilton (1979) and Bannon (1981), these findings show that subjects need not imagine that they are scanning a scene when interobject directions are encoded at an earlier time.

Another possible use of the imagined scanning of a scene has been suggested by Thorndyke (1981). Subjects estimated distances between points on a map from memory or while actually viewing the map. In both conditions, distance estimates increased linearly with increasing interpoint distance, and, independently, with increasing number of intervening points (note, incidentally, the agreement with the finding of Hartley, 1977, 1981, described in Section 1.1.5). The correspondence of these results with those of Kosslyn, Ball, and Reiser (1978) led Thorndyke to propose that people make estimates of distance by imagining themselves or an object moving between locations on the visible or imagined map and by noting the scan duration.



(a)

(b)

Figure 37.14. Dependence of time to verify that an arrow points at a previously presented dot, on the distance of the arrow from that dot. (a) An example of a dot pattern and subsequently presented test arrow. (b) Mean reaction time to verify that the arrow was pointing at one of the previously presented dots, plotted as a function of the distance between the arrow and the dot. The similarity between this function and that typically found when subjects are specifically instructed to form and to scan mental images (see Figure 37.13) suggests that these patterns had also been imagined and then mentally scanned. (From R. A. Finke & S. Pinker, Directional scanning of remembered visual patterns, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9. Copyright 1983 by American Psychological Association. Reprinted with permission.)

2.3. Changes in Visual–Motor Coordination Resulting When Errors of Movement Are Imagined or Observed

2.3.1. Effects of Imagined Practice. It has often been proposed that imagery can serve as a kind of cognitive "workspace," to test new actions mentally and thus to avoid the risk and effort involved in actually performing them (e.g., see Attneave, 1974; Baylor, 1972; Metzler & Shepard, 1974; Pinker & Kosslyn, 1978; Shepard, 1978b). Experiments on "mental practice," moreover, have shown that, when people mentally rehearse a skill, their performance on the skill can improve (e.g., see L. V. Clark, 1960; Rawlings, Rawlings, Chen, & Yilk, 1972; A. Richardson, 1967; Smyth, 1975). It is less clear whether visual imagery, as opposed to proprioceptive imagery or some less specific movement-planning process, is responsible for the beneficial effects of mental practice (e.g., Start & Richardson, 1964).

Greenwald (1970) has proposed that visual imagery permits one to anticipate visual *feedback* (i.e., the visual consequences of movement), so that errors of movement can be detected (see also G. A. Miller, Galanter, & Pribram, 1960). This possibility suggests that imagery might help control movements made in the absence of continuous visual feedback. In a study by Thomson (1983), subjects inspected an array of objects located at various distances along the ground and then walked to designated target objects with their eyes closed. The subjects were able to do this accurately up to 8 sec after closing their eyes; after that, they reported that they could no longer visualize how close they were coming to the targets. Additional findings suggested that they had imagined the visual consequences of their actions, rather than merely counting their steps or using proprioceptive cues. For example, when they were told to stop at unexpected points along their walk and to throw a block at the target, they could do so accurately for, again, up to 8 sec after closing their eyes.

2.3.2. Effects of Imagined Errors of Movement. The experiments reviewed next explore the related question of whether imagined errors of movement affect visual–motor coordination in the same way as perceived errors of movement. When people attempt to point at objects while looking through prisms that displace the apparent locations of the objects, they point erroneously at first, but if permitted to see their movements they quickly adapt to the distortion and become more accurate over time (although actual movement is not a necessary condition for prism adaptation; see Welch, 1978; and Welch, Chapter 24). Finke (1979b) had one group of subjects look through laterally displacing prisms and point repeatedly at a target. They were allowed to observe their errors at the moment they completed each movement. A second group of subjects also looked through the prisms and were asked to point repeatedly at the target, but they were *not* allowed to observe their errors. Instead, they were instructed to imagine that they saw themselves making the same sequence of pointing errors that subjects had made in the perception condition, denoted by markers placed to one side of the target. A third group of subjects participated in a control condition, identical to the imagery condition except that no imagery instructions were given.

The size and direction of pointing aftereffects in the adapted and unadapted hands were determined by subtracting the postadaptation errors from the preadaptation errors. While pointing aftereffects in the imagery condition were half as large as those in the perception condition (which revealed an adaptation of approximately 40% of the original prism displacement), they were in the predicted direction and exhibited roughly the same amount of intermanual transfer from the adapted to the unadapted hand (Figure 37.15), as is characteristic of prism adaptation under terminal as opposed to continuous feedback conditions (see M. M. Cohen, 1967; C. S. Harris, 1965; Wilkinson, 1971). In contrast, subjects in the control condition did not show significant pointing aftereffects. The correspondence between the intermanual transfer of pointing aftereffects in the first two conditions suggests that imagined errors of movement are functionally equivalent to those that are observed, possibly as a result of activation, during imagination, of visual mechanisms concerned with the detection of such errors.

These findings are not easily explained in terms of subjects' expectations, because, even when subjects were led to expect that their actual pointing errors would be contrary to the errors they were asked to imagine, the aftereffects were still determined by the imagined errors (Finke, 1979b, Experiment 3). Nor can the results be explained in terms of tacit knowledge, because the experimental subjects had never previously adapted to distorting prisms.

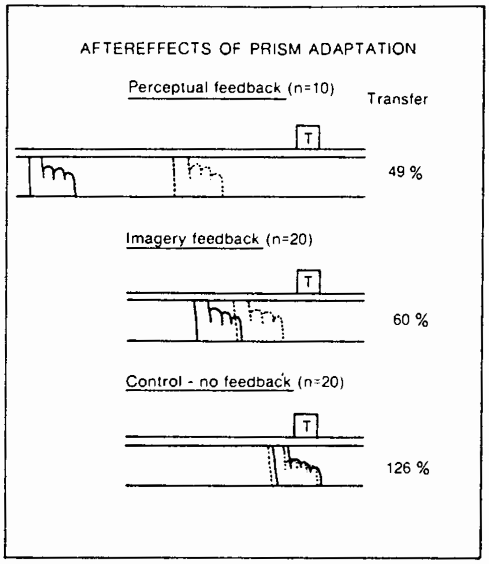


Figure 37.15. Mean relative pointing aftereffects when errors of movement occurring during adaptation to visual displacement prisms were observed (top), imagined (middle), and neither observed nor imagined (bottom). Schematic hands drawn in solid and dashed lines represent the relative sizes of the aftereffects for the adapted and unadapted hands, respectively. Intermanual transfer, which is shown as a percentage, suggests a functional equivalence between errors of pointing that subjects actually observed or only imagined. (Data from R. A. Finke, Levels of equivalence in imagery and perception, *Psychological Review, 87*. Copyright 1980 by American Psychological Association. Reprinted with permission.)

These results could be explained, however, at least in part by assuming that the aftereffects of prism adaptation are caused by potentiation of the subjects' eye muscles while they look through the distorting prisms at their real or imagined errors (e.g., Paap & Ebenholtz, 1976). Evidence against this alternative comes from an extension (by Finke, 1979a) of the above experiments, in which subjects imagined making the same errors of movement without looking through prisms or moving their eyes. Significant pointing aftereffects were still obtained, although in this case they did not transfer intermanually.

As one example of a possible application of these findings, people could "fine-tune" their visual-motor coordination by imagining movements and their consequences. However, these findings also indicate that, in order to do this appropriately, people would already need to have some knowledge of the actual consequences of their movements.

2.4. Summary

The findings in this section suggest that mental imagery can have useful visual functions when (1) comparing distances among objects not physically present, (2) comparing visual angles between objects as they would appear when viewed from a different vantage point, (3) verifying that an object at a newly specified location lies along a particular direction with respect to objects previously observed, and (4) practicing a skill mentally to refine the anticipation of the visual consequences of movement.

With regard to the functional equivalence claim, the strongest evidence among these studies come from those demonstrating (1) correspondences between changes in visual-motor coordination following the imagination and observation of errors of movement (Finke, 1979a, 1979b) and (2) correspondences between the effects of distance on the time required to "scan" a configuration of objects that is remembered or imagined (e.g., Finke & Pinker, 1982, 1983; Kosslyn et al., 1978; Pinker, 1980a). The weakest evidence comes from studies on remembering the relative locations of objects or places, the inaccuracies of which can be better explained by assuming that the information is recoded into some kind of symbolic representation than by assuming that mental imagery is used (e.g., Baird et al., 1982; Moar & Bower, 1983; A. Stevens & Coupe, 1978; Tversky, 1981).

The tacit knowledge alternative, which holds that subjects have, on the basis of previous experience, acquired enough abstract information about the relevant physical or perceptual processes to give responses imitating what would be expected if they were actually experiencing such processes, could account for many of the findings on the mental scanning of imagined objects (Mitchell & Richman, 1980; Richman et al., 1979). It is least successful in explaining the characteristics of visual-motor aftereffects following the imagined adaptation to distorting prisms, where the required tacit knowledge was not available (again, Finke, 1979a, 1979b).

The eye movement alternative, in contrast, can account for some of the results of imagined prism adaptation but cannot explain the findings suggesting that the mental scanning of objects occurs in three as well as in two dimensions (Pinker, 1980a; Pinker & Kosslyn, 1978).

The other alternative, experimentally introduced bias, offers the least satisfactory explanation for these findings, since few opportunities existed for introducing such biases into most of these studies. Moreover, even when such opportunities did exist (as in some of the image scanning studies), the results appear to be unaffected by the expectations of the experimenter (e.g., Kosslyn & Jolicoeur, Note 5).

3. VISUAL IMAGERY IN REPRESENTING SPATIAL TRANSFORMATIONS

This section reviews studies investigating: (1) the imagined rotation of objects in two and three dimensions; (2) the effects of object complexity, methods of presentation, and frames of reference on imagined rotation; (3) the imagined transformation of the size, color, and shape of objects; and (4) the illusions of *apparent motion* between objects alternately displayed in different orientations or sizes. The studies of imagined rotation, particularly, bear on fundamental issues concerning the extent to which mental transformations can be continuous and/or holistic. These issues are specifically taken up (after describing the studies themselves) in Section 3.1.9.

3.1. Mental Rotation

3.1.1. Imagined Rotation of Objects That Have a Standard Upright Orientation. In a study by Cooper and Shepard (1973a, 1973b), subjects were asked to discriminate normal from mirror-reversed letters and numerals when these were presented at various orientations in the picture plane. In the absence of advance information as to the identity or orientation of the test character, discrimination time increased markedly with departure of the character from its normal, upright orientation.

In another condition, the subjects were first shown an outline drawing of a normal character in the upright orientation and were then shown an arrow designating the orientation at which that character or its mirror image was to be presented. See Figure 37.16(a). Following a delay of 100–1000 msec, the test pattern was presented, and subjects indicated whether it was the normal or mirror-reversed version of the target character they had just seen. Figure 37.16(b) shows that discrimination time increased monotonically with increasing rotation of the test patterns (up to 180°), and that this monotonic increase diminished with increasing preparation time. Moreover, when subjects were given a full second to prepare, their times were essentially the same for all test orientations, and they matched the times obtained when the normal characters were actually shown in advance at those orientations. The proposed interpretation of these results—that people imagine a pattern rotated into a particular orientation to make the required discrimination concerning that pattern—was supported by the subjects' introspections.

Several control conditions for this experiment were reported in Cooper and Shepard (1973a). Subjects were shown the target character in advance but were given no information about orientation, or were given advance information about orientation but not about the identity of the character, as seen in Figure 37.17(a). As shown in Figure 37.17(b), verification times for both of these control conditions increased markedly with increasing rotation of the test patterns. It is noteworthy that verification times obtained when orientation information only had been provided, though somewhat faster for all presented orientations, showed the same strong dependence on orientation as those obtained without any advance information. The constant difference in height of these two reaction time functions presumably reflects a savings in the time to identify what is the top of the test character, following the advance information as to orientation (compare Rock, 1973). The similarity in the shapes of the two functions, however, suggests that the subjects were not able to prepare in advance for a particular orientation in which a stimulus was to be discriminated unless they also knew which stimulus was to be discriminated in that orientation.

Similar conclusions follow from the results of a later study by Cooper and Shepard (1975), who found that subjects took less time to identify schematic drawings of right and left hands presented at various orientations when instructed to visualize, in advance, the appearance of each hand at its proper orientation. When the imagined and observed hands did not match, verification time increased monotonically with increasing angular departure from the upright position (i.e., with fingers pointing straight up). Corresponding results were obtained when the experimenters presented no advance information, or information only about orientation. However, when the imagined hand matched the observed hand, verification time was about equally fast across all orientations. These findings further suggest that people make, or prepare themselves to make, refined discriminative responses to a disorientated pattern, by imagining the pattern transformed into the orientation that will most facilitate the required discrimination.

One problem with the mental rotation account is that it leads one to expect that verification time should increase linearly with increasing angular departure of the test patterns from the canonical upright, on the assumption that the rotations are imagined at a constant rate, whereas in these studies this time increased more rapidly with increasing angular departure. (See again Figure 37.16 and Figure 37.17.) Possibly, subjects did not imagine the patterns rotated completely to the upright position, because the patterns could be discriminated even when imagined as somewhat tilted away from the upright (Cooper & Shepard, 1973a). In support of this hypothesis, Hock and Tromley (1978) found that reaction time increased linearly with degree of disorientation for letters that do not appear to be upright when slightly tilted; the wider the range of orientations in

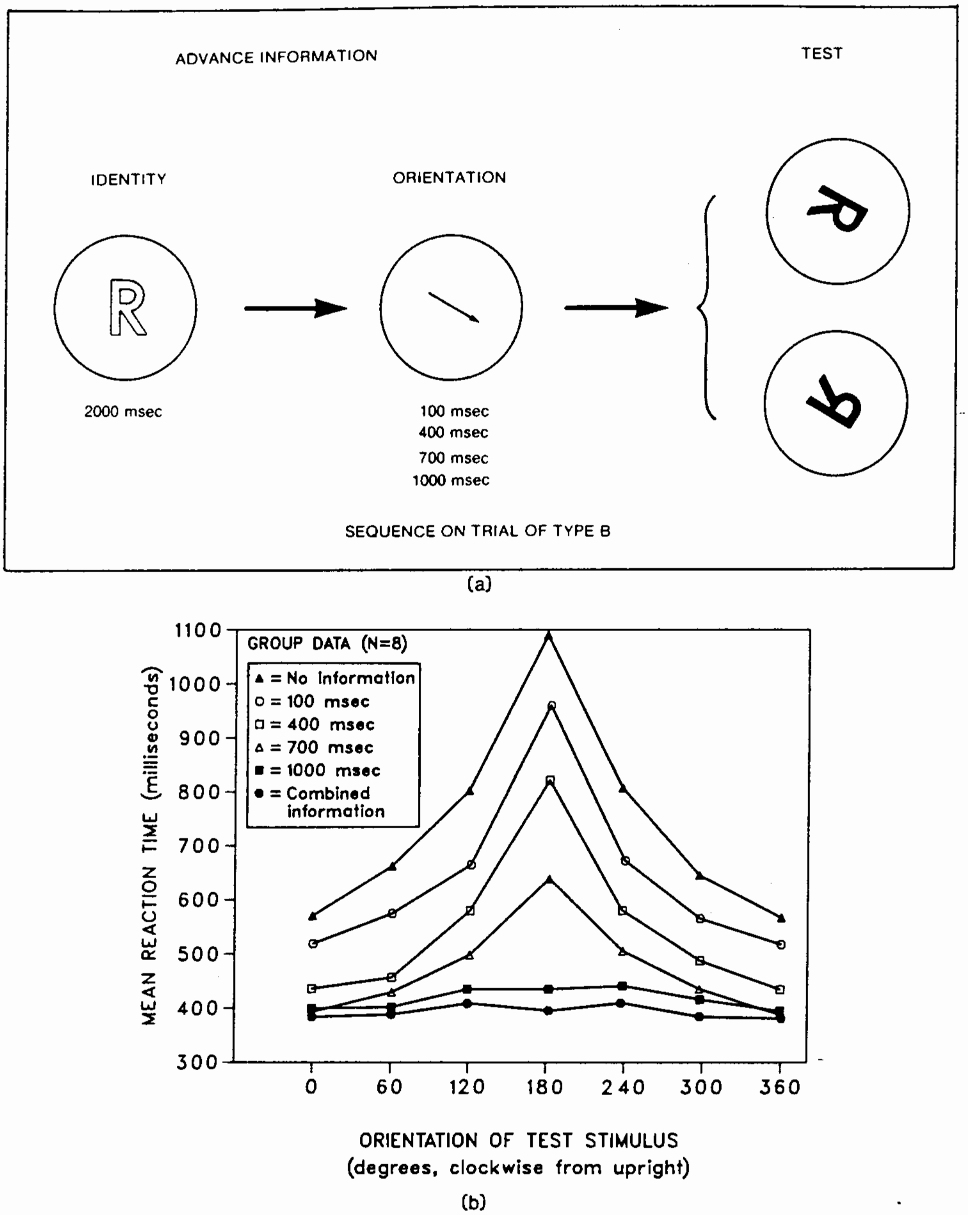


Figure 37.16. Dependence of time to discriminate normal from reflected versions of a specified alphanumeric character, on departure from upright orientation, following advance information as to orientation available for different amounts of time. (a) Example of the sequence of stimuli presented within the circular field on a particular trial: an outline of the character to be tested on that trial (displayed for 2 sec), immediately followed by an arrow indicating the orientation to be tested (displayed for 100, 400, 700, or 1,000 msec), immediately followed by the (normal or reflected) test stimulus. (b) Mean reaction time to indicate whether the test patterns were normal or reflected versions of the previous alphanumeric characters, plotted as a function of clockwise angular departure from the upright orientation, separately for different amounts of preparation time. The increase in reaction time as the orientation departed from upright suggests that subjects imagined disoriented characters rotated toward upright to make the discrimination; and the flattening of this increase when advance information as to orientation was available for longer times suggests that subjects could imagine the normal character rotated into the tilted orientation in advance and thus eliminate the need to imagine a rotation after the test stimulus appeared. (From L. A. Cooper & R. N. Shepard, The time required to prepare for a rotated stimulus, *Memory and Cognition*, 1973, *1*. Reprinted with permission.)

which the letter appeared to be upright, the more nonlinearly reaction time increased with departure from upright. Moreover, in their second experiment Cooper and Shepard (1973a) showed that the time to compare a test stimulus to an imagined stimulus when neither was likely to be in the canonical upright orientation, and when the data were averaged over all pairs of orientations, did increase linearly with angular difference. Finally, in other studies by Cooper (1975) and Shepard and Metzler (1971) to be considered in the following sections, subjects identified or compared rotated versions of forms having no standard upright orientation and, in this case, the reaction time functions were strikingly linear.

A plausible alternative to mental rotation in accounting for the general finding of a monotonic increase in reaction time with increasing angular departure from the upright is that subjects analyze the individual features of the pattern in some symbolic or descriptive fashion, and that this becomes more difficult to do as the pattern appears in less familiar orientations. Such an account would predict that the slope of the reaction time function ought to depend on the particular perspective views of the object shown and on its complexity; these predictions are examined in the following sections.

3.1.2. Imagined Rotation of Three-Dimensional Structures. Shepard and Metzler (1971) presented subjects with pairs of perspective line drawings of three-dimensional objects having no canonical orientation. Examples are shown in Figure 37.18(a). These objects were identical or were mirror images of one an-

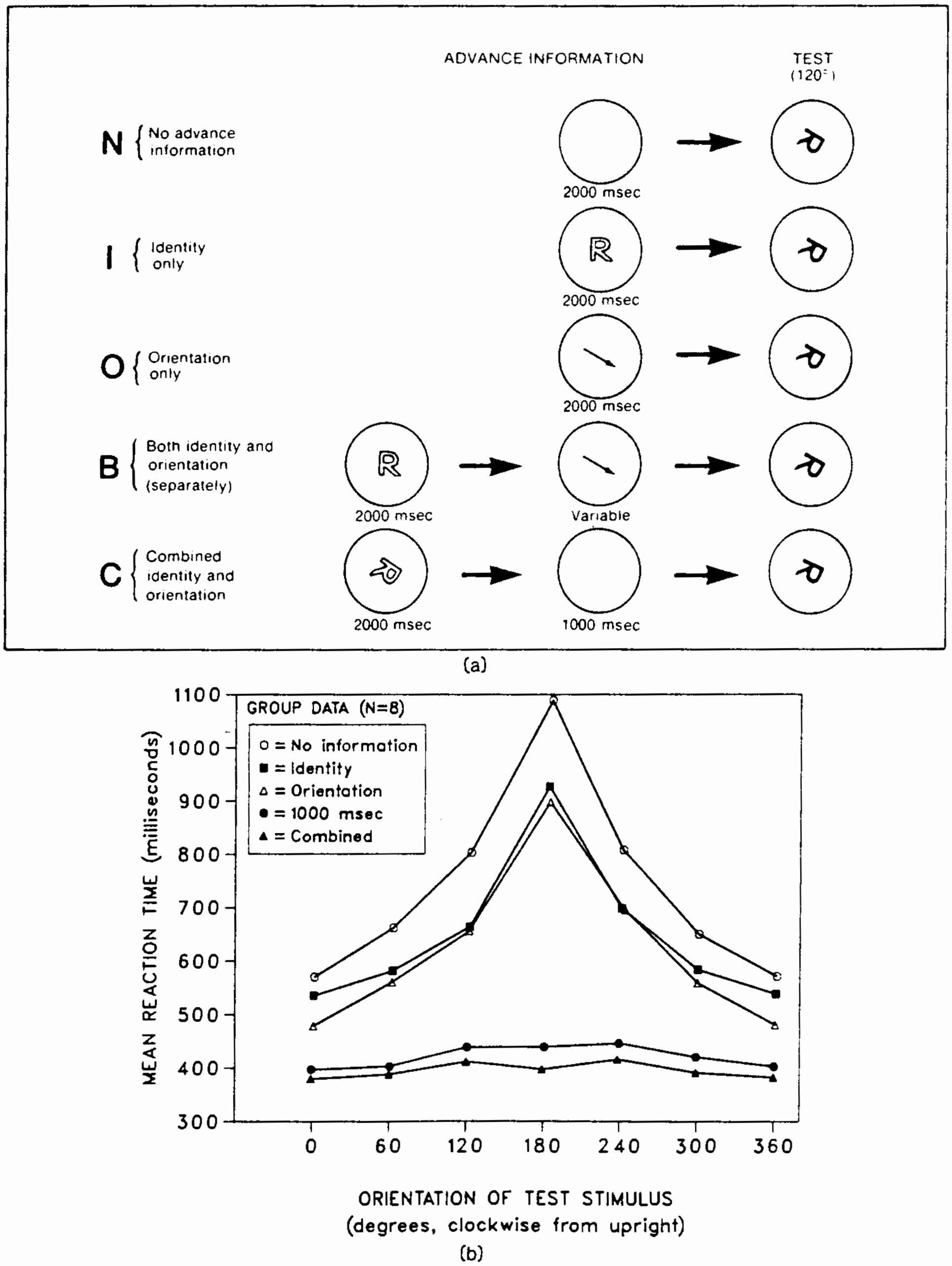


Figure 37.17. Dependence of time to discriminate normal from reflected versions of alphanumeric characters, on departure from upright, following advance information of different types. (a) Examples of sequences on trials that provided different types of advance information as to identity and/or orientation of the ensuing test character. (b) Mean reaction time to indicate for each type of advance information (if any) whether the test stimulus was a normal or reflected version, plotted as a function of clockwise angular departure from the upright orientation. Only when the identity as well as the orientation of the test pattern was indicated in advance did reaction time become independent of the orientation of that pattern. (From L. A. Cooper & R. N. Shepard, Chronometric studies of the rotation of mental images, *Visual information processing*, Academic Press, 1973. Reprinted with permission.)

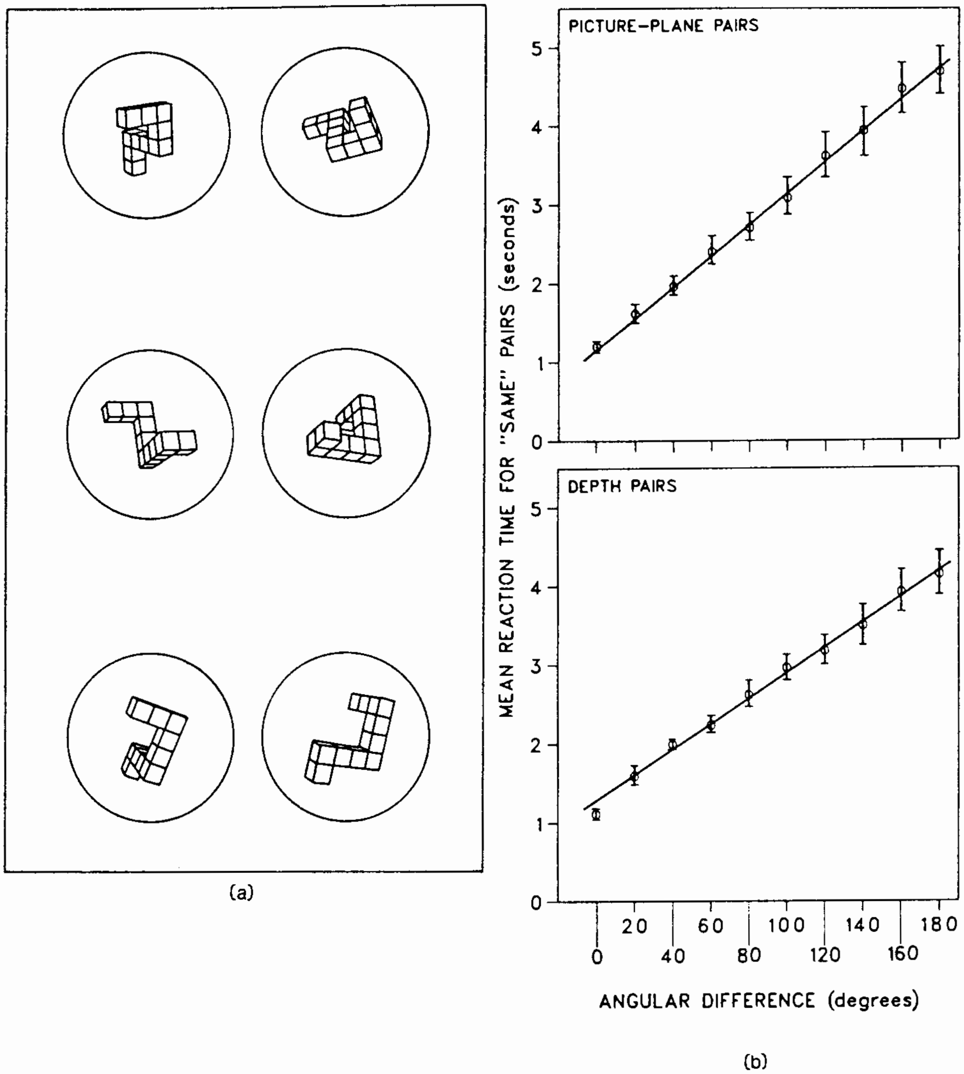


Figure 37.18. Dependence of time to determine that two three-dimensional objects were of the same shape, on the angular difference in their portrayed orientations. (a) Pairs of views of the Shepard-Metzler objects differing by a rotation in the picture plane (top), by a rotation in depth (middle), and by a reflection as well (bottom). (b) Mean reaction time to verify that the two objects were identical in shape plotted as a function of the angular difference in their orientations in the picture plane (above) and in depth (below). The approximate equivalence of the two fitted functions indicates that what subjects imagined to be transforming was the three-dimensional object, and not simply its two-dimensional picture. (From R. N. Shepard & J. Metzler, Mental rotation of three-dimensional objects, *Science, 171.* Copyright 1971 by American Association for the Advancement of Science. Reprinted with permission.)

other, and they differed by rotations either in the picture plane or in depth, about the vertical axis. Both types of rotational disparities occurred in multiples of 20° steps. The time required to verify that the objects were the same increased linearly at the same rate as a function of angular differences for rotations in the picture plane and in depth. See Figure 37.18(b). Evidently, what is imagined as rotating is the three-dimensional object and not the two-dimensional picture, even when the latter would suffice (and would have a simpler relation to the proximal stimulus on the subject's retina). Additional evidence for the equivalence of picture-plane and depth rotations for three-dimensional objects has been reported by Carpenter and Just (1978).

The time required to imagine an object rotated through a certain angle in depth evidently depends almost entirely on that angle and little, if at all, on differences in the pictorial representations of the objects as particular edges disappear or appear from behind others (Metzler & Shepard, 1974). Performance of such tasks also does not require normal visual stereopsis, since the results just described were obtained with pictures, not with the objects themselves (Shepard & Metzler, 1971), and the performance of stereoblind individuals does not differ from the performance of individuals with normal vision (Klein, 1977).

Evidence that the appearances of an object or arrangement of objects can be imagined from other vantage points has been provided in a different way by Pinker and Finke (1980). In an extension of Pinker's work on mentally scanning different two-dimensional projections of a three-dimensional array (described in Section 2.2.3), subjects were asked to draw a presented three-dimensional array as it would appear following a rotation of the whole array (note the relation of this experiment to the "perspective problem" as studied by developmental psychologists, e.g., by Huttenlocher & Presson, 1973; Salatas & Flavell, 1976). Subjects learned the locations of four objects suspended in a clear plastic cylinder, which were positioned so that when the cylinder was rotated by 90°, the objects would appear to form a parallelogram. This "emergent" pattern was not anticipated from the way the objects appeared from the original vantage point. In the first experiment, after their positions had been memorized the four objects were removed and the cylinder was rotated to the 90° position. The subjects were asked to describe and to draw the pattern that the objects were imagined to present, from their vantage point, following the rotation. The subjects' drawings revealed that to a good first approximation the form of the parallelogram had indeed "emerged" in the rotated images.

Pinker and Finke then asked the subjects to rotate the cylinder so that specified pairs of objects would appear to be vertically aligned, in conditions where they either observed the objects or imagined them. The results, shown in Figure 37.19(a), suggested that the subjects had imagined the objects rotated more than they had physically rotated the cylinder. When they then imagined a single object in the cylinder, and attempted to align a cursor with its apparent horizontal extent as the cylinder was rotated, they imagined the object rotated slightly ahead of where it actually would have appeared [at least for rotations up to 270°; see Figure 37.19(b)]. Theories of functional equivalence, while supported by the overall correspondence between changes in the perspective appearance of the imagined and observed objects, would not have predicted this systematic "overshoot" in the imagined rotation (see also the recent related work by Freyd & Finke, in press b; Pinker, Note 7).

3.1.3. Evidence That Imagined Rotations Correspond to the Perception of Actual Rotations. Cooper (1976a), Cooper and Shepard (1973a, Experiment 2), and Metzler and Shepard (1974) all reported experiments meeting two conditions. First, a test stimulus was presented during the course of an imagined rotation. Second, the subjects made fast and accurate discriminative responses only when that test stimulus came on at the orientation in which the subjects were expected to have been imagining the object at that moment in time—that is, at further rotated orientations after longer delays. In addition, Metzler and Shepard (1974) and Cooper (1975) provided evidence that, for a given angular separation between the objects, the subjects could imagine the rotation the shorter or the longer way around the 360° circle, with the times linearly increasing with angular extent of transformation, even beyond 180°. Thus an object imagined to be rotating is imagined in successively more and more rotated orientations.

To explore whether the imagined rotations would also include orientations at which subjects did not expect to be tested, Cooper (1976a) probed subjects whose rates of imagined rotation of random polygons had already been estimated. The subjects were first shown an outline drawing of one such polygon in one of six orientations that they had become familiar with in previous experiments [ranging from 0° to 360° in 60° steps; examples of these polygons are shown in Figure 37.20(a)]. Each subject was instructed to imagine the polygon rotating clockwise as soon as this reference pattern was removed. After a variable delay, a test pattern was presented, and the subject indicated whether it was the congruent or mirror image version of the reference pattern. These test patterns appeared not only at the practiced orientations (i.e., 0°, 60°, 120°, 180°, 240°, and 300°), but also, unexpectedly, at intermediate orientations (30°, 90°, 150°, 210°, 270°, and 330°).

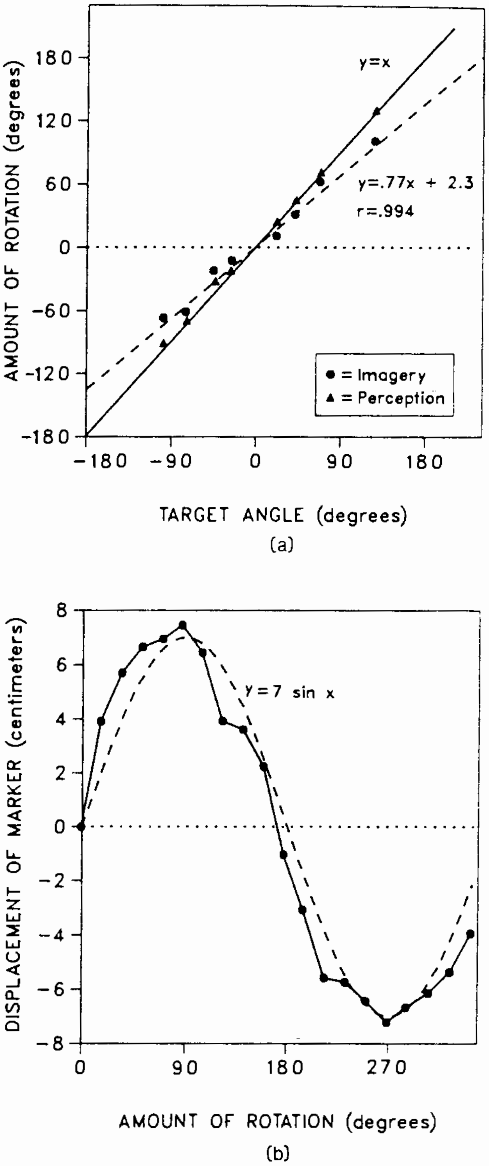


Figure 37.19. Accuracy in imagining the changed appearance of the arrangement of objects suspended in a cylinder that would result from various rotations of the cylinder. (a) Mean angular rotation of the cylinder needed to align pairs of objects vertically that were either imagined or actually observed to be suspended in the cylinder, plotted against the angle corresponding to perfectly accurate performance. (b) Judged apparent horizontal displacements of a single object imagined to rotate with the cylinder through 360°, compared to perfectly accurate performance. The functions suggest, to a first approximation, that a functional equivalence exists between imagined and observed changes in the perspective appearance of objects. (From S. Pinker & R. A. Finke, Emergent two-dimensional patterns in images rotated in depth, *Journal of Experimental Psychology: Human Perception and Performance*, 6. Copyright 1980 by American Psychological Association. Reprinted with permission.)

When a test pattern was presented in an orientation that the subject should be imagining at that moment, as inferred from that subject's previously estimated rate of mental rotation, verification time was independent of the absolute orientation of the test pattern, *even when it appeared at the intermediate, unexpected orientations*. In addition, when a test pattern was presented at some other orientation, verification time increased linearly with increasing angular departure of that orientation from the orientation that the subject was predicted to be imag-

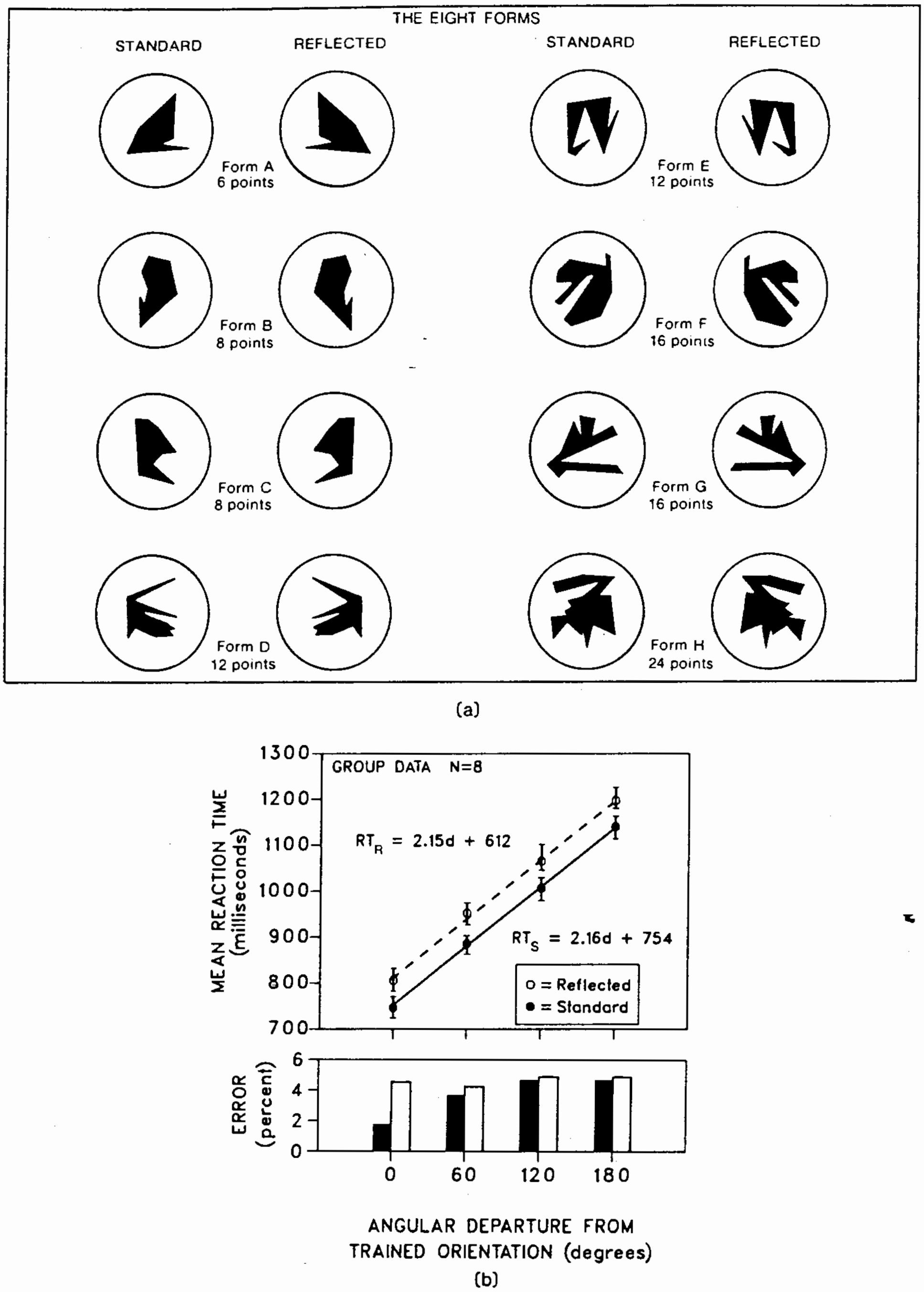
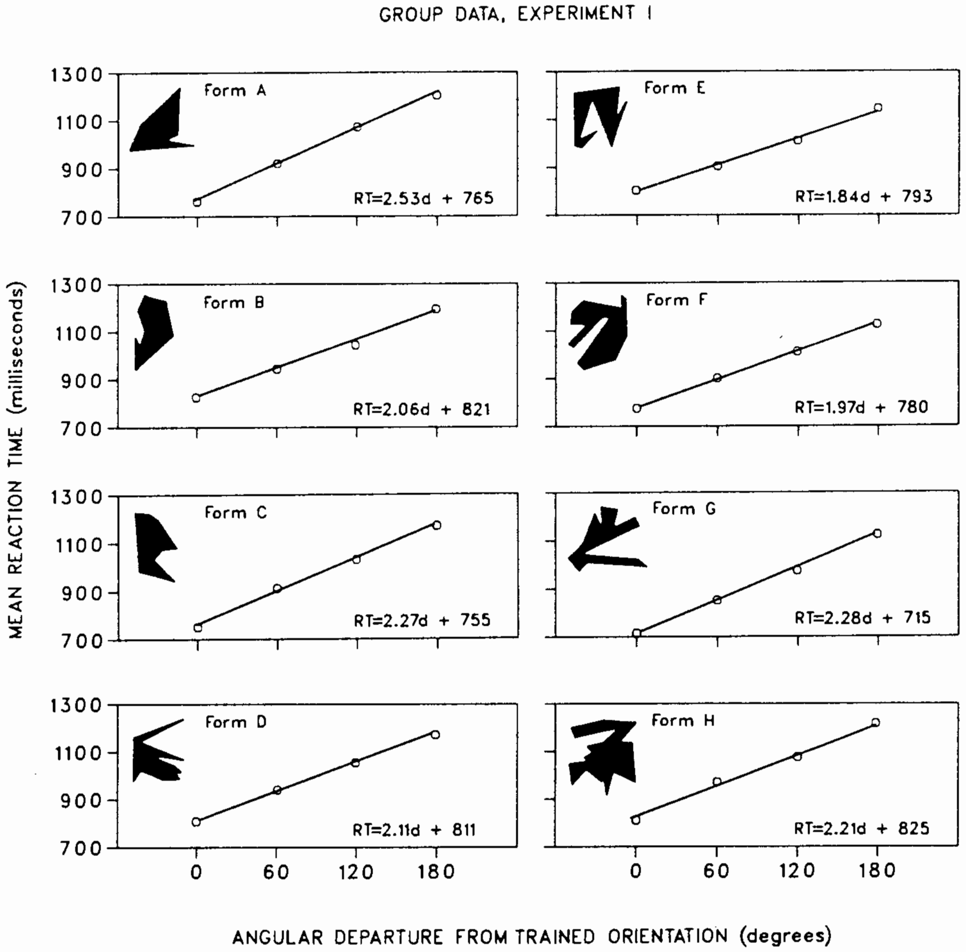


Figure 37.20. Time required to prepare for discrimination of normal from reflected polygons of different complexities when presented in orientations departing from their originally learned orientations. (a) Standard and reflected versions of random polygons constructed with different numbers of vertexes. (b) Mean reaction time to discriminate between two versions of the patterns (averaged across levels of complexity), plotted as a function of angular departure from the orientation at which they were originally learned. (c) Mean reaction time for verifying the standard versions of the individual polygons, plotted as a function of angular departure from learned orientation. The similar slopes of the functions suggest that the rate at which the rotation of each polygon was imagined was independent of the complexity of that polygon. (From L. A. Cooper, Mental rotation of random two-dimensional shapes, *Cognitive Psychology*, 1975, 7. Reprinted with permission.)

ining at that moment. Apparently, in preparing for the successive appearances of the patterns at the expected 60° intervals, subjects were also passing through states of imagining these patterns at the intermediate 30° orientations. These findings indicate that the imagining of a rotation between very different orientations, like the perception of an actual rotation, necessarily passes over a relatively dense trajectory of rotationally intermediate representations.

Additional evidence for the perceptual character of mental rotation comes from more recent work by Corballis and McLaren (1982). They examined the effects of observing patterns undergoing actual rotation on subsequent mental rotation.



(c)

Figure 37.20. *(continued)*

Subjects participating in a version of the Cooper-Shepard task indicated whether rotated letters were normal or reflected. Between presentations of the letters, they looked at a textured disk rotating about the line of sight. Corballis and McLaren found that motion aftereffects, resulting from inspection of the rotating disk, biased the direction of mental rotation for letter orientations. This finding suggests that when one imagines an object in rotation one is engaging the directionally sensitive mechanisms in the visual system that underlie the perception of motion, although it does not rule out other explanations.

3.1.4. Effects of Complexity and Familiarity of Objects on Rate of Imagined Rotation. If the difficulty of imagining an object depended on the complexity of the object, and if the act of imagining the object drew on the same limited resource as the act of imagining the object's rotation, rates of mental rotation would decrease with increasing complexity of the object (e.g., see Pylyshyn, 1978).

Cooper's (1975) random polygons, which are shown in Figure 37.20(a), were constructed by a method developed by Attneave and Arnoult (1956) so as to vary in number of vertexes (which corresponds to perceived complexity; see Attneave, 1957). In one experiment, subjects were trained to discriminate between normal and reflected (i.e., mirror image) versions of each of the patterns, which were presented at any of six orientations differing in 60° steps. During training, each subject saw the patterns at only one orientation, and these training orientations were varied between subjects. The subjects were then shown the patterns in both trained and unfamiliar orientations and indicated whether the patterns were normal or reflected.

As shown in Figure 37.20(b), reaction time averaged across the patterns increased linearly with increasing angular departure from the trained orientation (up to 180°). In addition, the functions for the normal and reflected patterns relating reaction time to angular departure showed identical slopes, with a small intercept difference across orientations. More significant, however, were the negative results for the effect of varying pattern complexity: As shown in Figure 37.20(c), the reaction time functions for each of the patterns showed no evidence of an increase in slope with increasing complexity.

In Cooper's second experiment, subjects were first shown outline drawings of the normal version of the patterns at the trained orientation and were then shown an arrow designating the orientation at which the test patterns would be presented (as in the Cooper & Shepard 1973b study). In this case, the subjects were able to control how long the arrow was displayed, during which time they were to imagine the standard pattern rotating in a prespecified (clockwise or counterclockwise) direction, and to indicate when they had reached the specified orientation and were ready to respond to the test pattern. Preparation time increased linearly with increasing angular departure of the test pattern up to 300° from the trained orientation, as measured along the specified direction of mental rotation. In addition, the reaction time function for judging the test patterns was essentially flat once the subjects had completed their preparatory mental rotation, as in other experiments by Cooper (1976a) and Cooper and Shepard (1973a), and the slope of the function for preparation time did not differ from that for the reaction time function that was obtained in the previous experiment, in which no advance information was provided about the test patterns or their orientations. For each of these results, pattern complexity again had no effect.

To exclude the possibility that subjects imagined the rotation of only a part of each pattern, Cooper and Podgorny (1976) constructed a set of test patterns that were graded perturbations of the standard patterns. See Figure 37.21(a). These test patterns

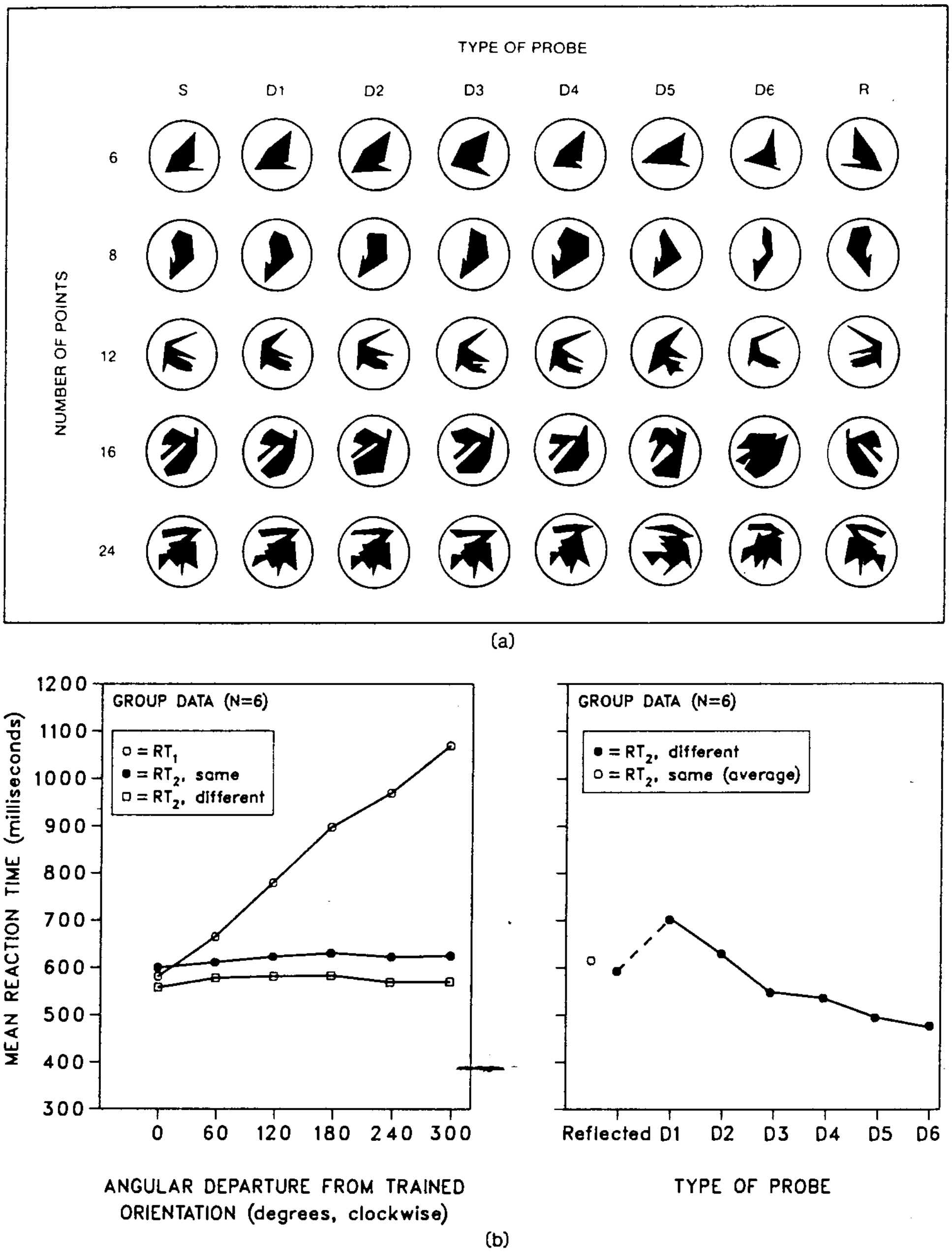


Figure 37.21. Accuracy of mental representation, following imagined rotation, assessed by ability to discriminate between each standard polygon and probe stimuli that differed from it by various small random perturbations. (a) The set of patterns used as test probes for varying levels of pattern complexity, consisting of standard (S) and reflected (R) versions, together with distractor patterns produced by increasing perturbations of the standard (D1–D6). (b) Mean reaction time to prepare for the rotated test probes (RT1) and to indicate whether the probes were identical to the standard patterns (RT2), plotted as a function of their deviation from the trained orientation (left) and as a function of their dissimilarity to the corresponding standard versions (right). The functions suggest that, whereas preparation time depended on pattern complexity, the time required to imagine the pattern rotated, or to discriminate it from similar versions, did not. (From L. A. Cooper & P. Podgorny, Mental transformations and visual comparison processes: Effects of complexity and similarity, *Journal of Experimental Psychology: Human Perception and Performance, 2*. Copyright 1976 by American Psychological Association. Reprinted with permission.)

were constructed by randomly displacing points on the standard patterns in increasing numbers and extents, and they were quantified according to similarity ratings obtained for each level of complexity. The experimental procedure was similar to that in the second experiment of Cooper (1975). Subjects indicated when they had imagined the standard pattern rotated to a specified orientation and then compared what they were imagining to the test pattern presented at that orientation.

The left half of Figure 37.21(b) shows that preparation time increased with increasing angular departure from the orientation of the standard pattern, independently of pattern complexity. Also as before, discrimination time was independent of the absolute orientation of the test patterns. In addition, as shown in the right half of this figure, discrimination time increased with increasing similarity between the standard and test patterns in a way that, once again, was independent of pattern complexity. This supports the subjects' claims that they imagined rotating each pattern as a whole. If the subjects had imagined rotating only a part of the pattern, they would usually not have been prepared to detect a small perturbation in some randomly selected part of the test stimulus.

Even so, mental rotation (like mental synthesis—see the discussion of Thompson & Klatzky, 1978, in Section 1.3.1) undoubtedly does depend on complexity, when perceptual learning of the particular objects has not progressed to the point where the subjects can readily imagine them transformed as a whole. Thus, Yuille and Steiger (1982), using objects of the type introduced by Shepard and Metzler (1971), reduced the effective complexity of those objects by informing the subjects that in the particular task used the subjects would need to attend to only one part of the object. The result was a marked decrease in the slope of the function relating reaction time to angular difference and, hence, a marked increase in the estimated rate of mental rotation. (Yuille & Steiger also *increased* the complexity of the objects by adding cubical blocks to form more elaborate though still redundant structures; however, this manipulation produced an increase primarily in the intercept rather than the slope of the reaction time function, suggesting an increase in comparison time more than in rotation time.) Robertson and Palmer (1983) obtained a similar effect of complexity on rate of mental rotation, using large letters (which could be normal or mirror reversed) that were composed of strings of identical small letters (which could all be normal or mirror reversed together, independently of the larger letter). Estimated rates of rotation were slower when structure had to be preserved at the local as well as the global level. Similarly, Hochberg and Gellman (1977) had previously found that mental rotations were performed more rapidly when relatively novel patterns had distinctive "landmark" features that could serve as salient cues for orientation. Likewise, when Pylyshyn (1979) instructed subjects to imagine line drawings of patterns rotated to verify whether similarly rotated test forms constituted parts of those patterns, he found that the slope of the reaction time functions decreased with increasing "goodness" of the forms. Evidently, as Kosslyn (1980) has argued, when the forms are not well learned, subjects may imagine the rotation of certain good, salient, or landmark features of the forms before they imagine the rotation of other parts. Such a strategy would of course greatly increase reaction time or errors in the kind of task employed by Cooper and Podgorny (1976) and was presumably avoided for this reason by their highly practiced subjects.

Supporting these conclusions, Bethell-Fox and Shepard obtained evidence that the effect of complexity does decrease with the amount of experience the subjects have with the particular forms used (see Shepard & Cooper, 1982, p. 178). In the same connection, Shepard and Cooper (1982) noted that the effects of "pattern goodness" (see Pomerantz & Kubovy, Chapter 36, for a discussion of this concept) on rate of mental rotation reported by Pylyshyn (1979) may have been confounded with differences in encoding time for his test forms (since mental rotations in his study were always initiated from the same orientation). To support a claim that subjects can never imagine an object rotating as a whole, one presumably would have to (1) establish that the object was highly overlearned, and (2) employ a task (such as a combination of those devised by Cooper 1976a; Cooper & Podgorny, 1976; Shwartz, 1979) which cannot be accomplished by mentally transforming only one part of the object at a time.

The various findings just summarized seem to be consistent with the following tentative conclusions. A subject who, through exposure or practice, has become more familiar with a particular object will be more likely to imagine the object transformed as a whole, and at a rate that is less dependent on the complexity of the object. Rates of mental transformation may, however, depend on other structural factors. For example, Metzler (1973) found that subjects were faster at completing an imagined rotation around a natural axis of that object than around an arbitrary axis (see Metzler & Shepard, 1974). An account based on a descriptive analysis of the individual features of an object might seem to provide a better explanation than does the functional equivalence account in such a case. In order to demonstrate this, however, one would have to show that the *perception* of rotation is not similarly facilitated when it takes place around a natural axis of the object.

3.1.5. Comparisons Between Rates for Two- and Three-Dimensional Objects. A further issue for the interpretation of mental rotation arises from the large range in reported rates of mental rotation: Shepard and Metzler (1971) estimated rates on the order of 60° per second whereas Cooper and Podgorny (1976) estimated rates on the order of 500° per second. One might be tempted to attribute this difference to the fact that the objects were three-dimensional in the former case and only two-dimensional in the latter (although the functional equivalence hypothesis has not been shown to entail a dependence of rate of mental rotation on dimensionality). In any case, Podgorny (Note 8) found no difference between rates of mental rotation for three-dimensional Shepard-Metzler forms and for two-dimensional silhouette versions of those forms. Similarly, Cooper and Farrell (see Shepard & Cooper, 1982, pp. 179–181) found equivalent rates of mental rotation for drawings of three-dimensional cubes and for similar drawings of two-dimensional hexagonal patterns (which contained the same number of lines as the drawings of the three-dimensional cubes). Hence, differences in procedure rather than difference in stimulus dimensionality may have been primarily responsible for the marked differences in rates of mental rotation reported in the various experiments.

3.1.6. Comparisons Between Simultaneous and Successive Stimulus Presentations. A procedural difference between the Shepard–Metzler (1971) and Cooper–Podgorny (1976) studies that may be relevant to the reported differences in rate is that the standard and test patterns were presented simultaneously in the former but successively in the latter. Podgorny (Note 8) compared simultaneous and successive methods of presentation for Shepard–Metzler forms and found that mental rotation rates were indeed about 1.5 times faster when the patterns were