Mental Imagery and the Visual System

What is the relation between mental imagery and visual perception? Recent work suggests the two share many of the same neural processes in the human visual system

by Ronald A. Finke

People often report that they can form mental images of an object that resemble the object's actual appearance. The act of constructing such images often produces visual sensations that seem quite realistic. Imagine, for example, that you are looking at an elephant. Does it have a curved trunk? What color are its tusks? How big are its eyes? Most people contend they attempt to answer such questions by "inspecting" a mental image in much the same way as they would inspect a real elephant.

These informal observations about imagery naturally lead one to consider the extent to which imagery and visual perception might be related. They suggest in particular that mental imagery may involve many of the same kinds of internal neural processes that underlie visual perception, a possibility that would have important theoretical and practical implications. If it could be established, for instance, that mental imagery shares with visual perception common neural mechanisms in the human visual system, one could begin to establish just how imagery may interact with visual perception. This would make it possible to explore the various ways imagery could function to facilitate, enhance or even substitute for visual perception.

For the past 10 years my colleagues and I have been developing techniques for investigating the functional relation between mental imagery and visual perception. Because experimental subjects can often guess what ought to happen in an imagery experiment, we have striven to make our techniques precise enough to reveal subtle correspondences between imagery and perception. Our work has revealed that mental images display a much richer variety of visual properties than had been previously thought, but also that imagery differs from perception in certain respects.

Through introspection one can recognize that features of a mental image formed at a small size or a far distance are harder to distinguish than features of an image formed at a large size or a near distance. Try, for example, to imagine an ant on a newspaper several feet away and then on the tip of a toothpick directly in front of your eyes. You should be able to mentally "see" many more of the ant's features (such as its head and body segments) when you imagine it at close range.

Stephen M. Kosslyn of Harvard University explored this relation between image size and feature resolution by employing simple reaction-time techniques. He found that the features of an imagined animal, such as the eyes and ears of a cat, could be detected more quickly when subjects were instructed to fashion relatively large images or assume a relatively close vantage. The experiments were inspired by the common observation that features of real physical objects can be detected faster when they are viewed from a closer distance.

More recently Howard S. Kurtzman and I have done experiments at Cornell University to measure precisely how well the features of objects can be resolved, or distinguished, in imagery and in perception. We were particularly interested in how the size of the features, their spacing and their position in the visual field affected resolution, or the ability to distinguish among details. We predicted that across all these variations visual resolution in mental imagery should match the resolution in perception.

Resolution in visual perception falls off continually as one observes an object at locations progressively farther from the point of eye fixation. The amount of detail that can be distinguished is not the same in all directions, however. As a rule resolution decreases more slowly along the horizontal axis of the visual field than along the vertical axis, and more slowly below the point of fixation than above it. It is also known that bar gratings become harder to resolve as the gratings become increasingly finer—more precisely, as their fundamental spatial frequency increases.

Our method for measuring limits of resolution in mental imagery was based on certain techniques common to visual psychophysics. Initially we showed our subjects a flat disk whose upper half was filled with a series of vertical bars and whose lower half was filled with a series of horizontal bars. The bars in both gratings were the same width. We then instructed our subjects to form a mental image of the disk and to project the image on the center of a screen directly in front of them. On the screen eight lines extended radially out from the center. The subjects indicated how far they could

WATERCOLOR LANDSCAPE was painted by a blind Scotswoman who works from her mental images. The artist, Carolyn James, suffers from a particularly acute form of the eye disease known as retinitis pigmentosa. Now 42 years old, she was registered blind at 21. To paint she lines up 24 watercolor jars in front of her in a memorized order. She moves from section to section on the paper, determining what she has just finished by detecting the moisture with her fingertips. Each of her watercolors is typically composed of six layers of paint.

look away from their images along each of the lines on the display before they could no longer tell the two halves of the imagined pattern apart. They reported that the gratings appeared fuzzy and indistinct as they were imagined farther into the visual periphery and then could no longer be distinguished beyond a certain point. For comparison, the same judgments were also obtained in a perception condition, in which the same disk was actually projected on the center of the screen in front of the subjects.

We repeated the experiment for each of three disks. The bar widths of the second disk were three times thinner than those of the first; those of the third were three times thinner than those of the second. On the average the fields of visual resolution decreased in size with increasing spatial frequency (decreasing bar width), and they were virtually identical whether the gratings had been imagined or observed. The imagery and the perceptual fields were also very similar in shape: resolution decreased more slowly along the horizontal axis than along the vertical axis, as well as more slowly below the point of fixation than above it.

We then did a control experiment in which we showed a different group of subjects the original set of three disks and asked them to predict the image-resolution fields. Their predictions differed considerably from what we had observed, arguing against a trivial "guessing" explanation of our original findings.

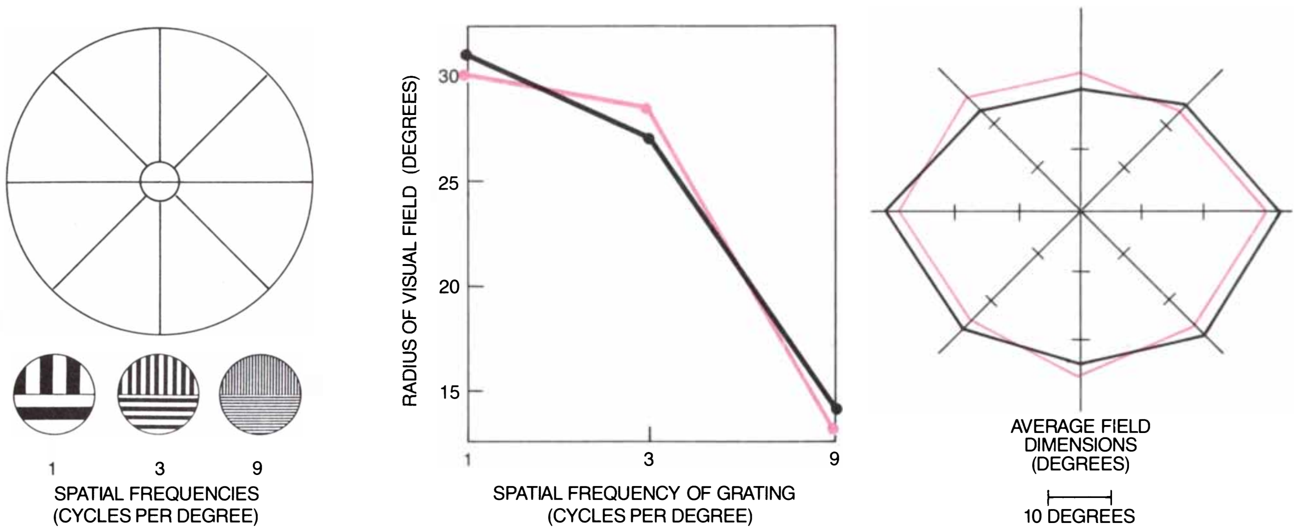
We interpret the results as evidence that pattern discrimination in imagery is constrained in much the same way that it is in visual perception. We propose in addition that these mutual constraints are probably imposed at the pattern-processing levels of the visual system, where the properties of certain neural mechanisms may limit the ability to resolve small or narrowly spaced visual features. Kurtzman and I have done other experiments that are consistent with this result. We found no correspondence between judgments of images and judgments of objects involving differing amounts of visual contrast, or relative brightness, among features. These aspects of perception are thought to be constrained by more primitive kinds of neural processes operating below the levels at which pattern processing takes place.

Our image-resolution findings were based on mental images of flat, two-dimensional patterns. Mental imagery is typically three-dimensional, however; it depicts how objects look in depth as they are viewed from various vantages. When most people imagine their living room, for instance, they can mentally "see" that certain pieces of furniture are in front of others, depending on where in the living room they imagine themselves to be.

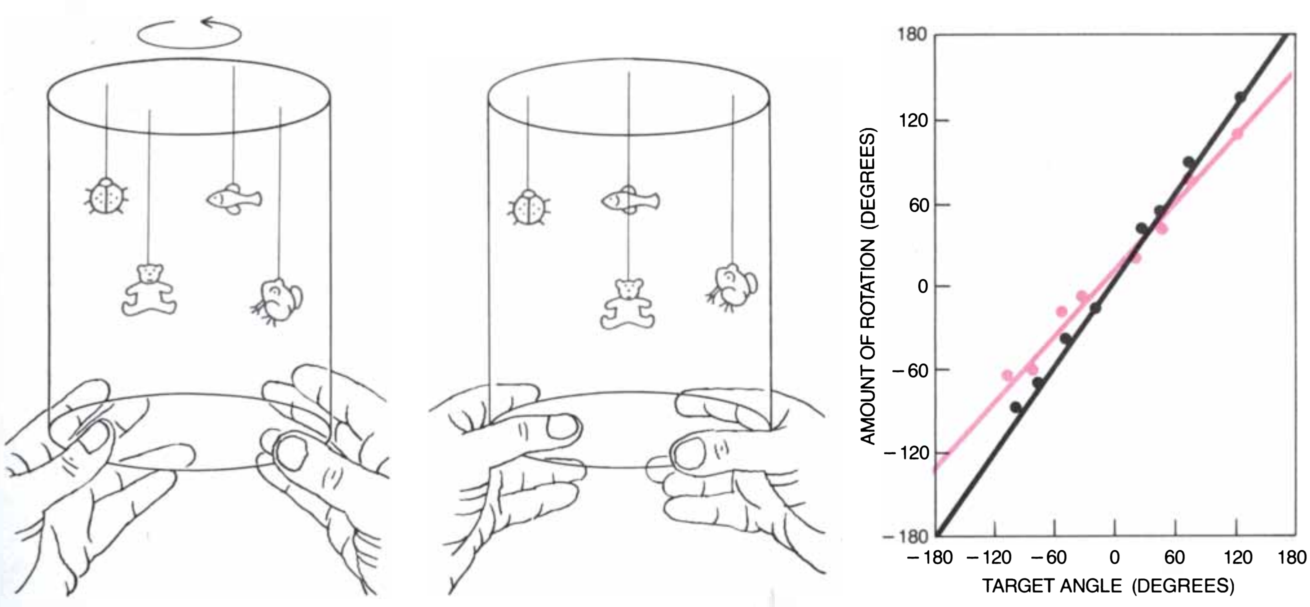
To investigate the three-dimensional properties of images, Steven Pinker and I asked subjects to form and mentally rotate images of a configuration of objects in space. When one actually looks at a three-dimensional configuration of objects from different perspectives, the objects are usually seen to shift their relative position in depth as the viewer moves or as the configuration is rotated. Recall, for example, times when you may have watched

people riding on a merry-go-round. You probably noted that the people would appear to shift their locations in relation to your vantage as the merry-go-round turned, perhaps forming familiar two-dimensional patterns at certain moments in much the same way that a constellation of stars often forms flat, recognizable patterns. In our experiments, which we did at Harvard, we were particularly interested in finding out whether similar kinds of patterns would appear to emerge when subjects imagined looking at a rotated configuration of objects. The results indicate striking similarities.

We asked our subjects first to learn the locations of four small plastic animals suspended at different heights in a transparent cylinder, and then to form mental images of each of them after



CONSTRAINTS on visual resolution, or the ability to distinguish details, were measured by means of the three patterns shown at the bottom left. Experimental subjects were instructed to form mental images of each pattern and to project their images onto the center of a large circular display (*top left*). The subjects then indicated how far they could look away from their images along each of the eight lines on the display before they could no longer tell the two halves of the imagined patterns apart. The visual fields within which the bar gratings of the imagined patterns could be distinguished decreased in size with increasing spatial frequency, or decreasing bar width (*colored line in middle*). On the average these fields were elongated horizontally and were larger below the direction of gaze than above it (*colored shape at right*). Virtually identical results (*black line in middle and black shape at right*) were obtained when the patterns were actually projected on the display, indicating that similar constraints are imposed on feature resolution in imagery and in perception.



THREE-DIMENSIONAL PROPERTIES of mental images were explored with a transparent cylinder. Subjects were told to learn the locations of four small plastic animals suspended at different heights in the cylinder and then, after the animals had been removed, to form mental images of them. As the empty cylinder was rotated 90 degrees, the subjects rotated their mental images (*left*). Although the appearance of the objects in the original viewing position did not suggest that a parallelogram with each animal at a corner would emerge, the subjects detected such a pattern. Their drawings of the pattern, however, showed small but systematic distortions. An explanation for the occurrence of these distortions was suggested by another experiment, in which the subjects rotated the cylinder manually in order to align pairs of the imagined animals vertically (*middle*). To align the imagined animals, subjects rotated the cylinder by a consistently smaller amount (*colored line at right*) than they did in a similar test with the animals physically present (*black line at right*). In other words, the subjects had mentally rotated their images ahead of their manual rotation of the cylinder.

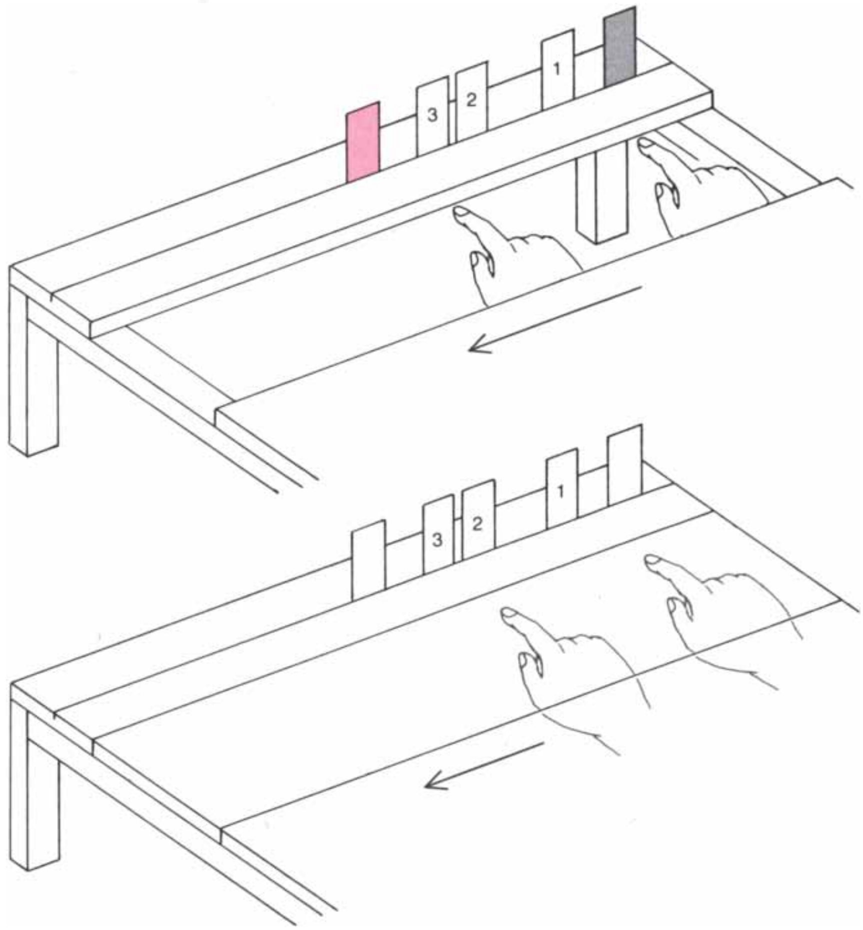
they were removed. We next rotated the empty cylinder 90 degrees and instructed the subjects to draw the imagined configuration as it now appeared from the new vantage. If they had imagined and rotated the animals with perfect accuracy as we rotated the cylinder, the animals would have seemed to form a parallelogram.

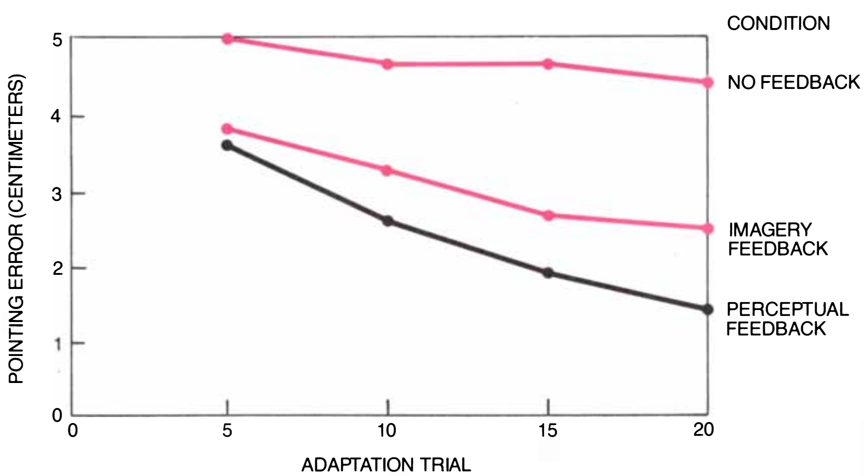
In every case the subjects' drawings revealed that their rotated mental images had depicted a pattern closely resembling the parallelogram, even though the appearance of the objects from the original viewing direction did not suggest that this particular geometric form would emerge. Curiously, there were small but systematic distortions in the drawings. The nature of the distortions suggested that the mental rotations had been less than perfectly accurate.

An explanation of these small distortions was suggested by another experiment, in which the subjects manually rotated the cylinder in order to align pairs of the imagined animals vertically. To our surprise we found that the subjects aligned the imagined animals by consistently rotating the cylinder less than was necessary when the animals were physically present. In other words, the subjects had mentally rotated their images ahead of their manual rotation of the cylinder. This tendency to advance an image ahead by small amounts accounted for the minor distortions we had found in our subjects' drawings of the emergent patterns. The experiment thus strengthened our contention that people can accurately imagine the visual perspectives offered by three-dimensional displays. Moreover, it enabled us to measure properties of mental images that naive subjects would not ordinarily expect.

Showing that subjects cannot guess the outcome of an imagery experiment does not, however, rule out the possibility that their performance could be based on unconscious knowledge about changes in the visual appearance of objects—knowledge that could indirectly influence judgments about images. One method for addressing this problem is to have subjects imagine events so atypical or unnatural that the events could not have been previously experienced. If under these conditions behavioral responses obtained from imagery still correspond to those obtained from perception, the imagery performance could not be attributed to the influence of earlier perceptual experiences.

In a series of experiments I carried out at the Massachusetts Institute of





FUNCTIONAL VALUE of imagery was assessed by considering the role images might play in prism adaptation. Optical prisms displace the apparent location of an object. Subjects wearing special glasses containing such prisms were instructed to point to a target (*colored marker in top row*). As a result of the effect of the prisms they at first pointed about five centimeters to the right of the marker (*gray marker*). Since the prisms displace everything in the field of view, however, once the individuals had extended their arms they could see their error and correct for it in successive attempts (*markers 1–3*). The markers were then utilized by a second group of subjects known as the imagery subjects. They too wore the special glasses, but the space between them and the table holding the markers was covered by a board so that they could not see the fingers of their extended arms (*second row*). Their task was to imagine that they saw their pointing finger arrive under the appropriate error marker as soon as their arms were fully extended. Subjects in a third group, the control group, pointed to the colored marker without being able to observe their errors or being told to imagine them. The graph shows that a significant amount of error reduction ensued when the subjects imagined their errors—almost as much as when they actually saw the errors.

Technology, I attempted to provide this kind of evidence for the functional value of imagery by considering the possible role images might play in prism adaptation. Optical prisms displace the apparent location of objects. A large body of research has shown that people quickly adapt to observing the world through such prisms provided they can move about and note their errors. When the prisms are removed following adaptation, people proceed to make errors of movement in the opposite direction, reflecting changes in their visual-motor coordination. My experiments demonstrated that prism adaptation can occur even when people point at a target and merely imagine they are making errors of movement like those typically induced by displacement prisms.

The subjects in my experiments wore special glasses containing optical prisms. I asked the subjects in one group to extend their right arms and point to a red marker positioned at eye level on a table in front of them. Owing to the effect of the prisms they at first pointed about five centimeters to the right of the marker. Since the prisms displace everything in the field of view, once the individuals had extended their arms they could see their errors and correct for them during a series of subsequent attempts.

I measured and averaged the errors over consecutive groups of trials and then displayed the average error locations with three markers. The markers were for the use of a second group of subjects known as the imagery subjects. These subjects also wore the special glasses, but the area between them and the table holding the markers was covered by a board so that they could not see the fingers of their outstretched arms. I instructed the imagery subjects to point to the red marker while looking through the prisms and then to imagine they saw their pointing finger arrive under the appropriate error marker as soon as their arm was fully extended.

The error markers, in other words, ensured that the imagined errors would correspond to the average pointing errors made by the first group of subjects. I also included a third group of subjects as controls. The individuals in the control group pointed to the red marker without having the benefit of either observing their errors or being told to imagine them.

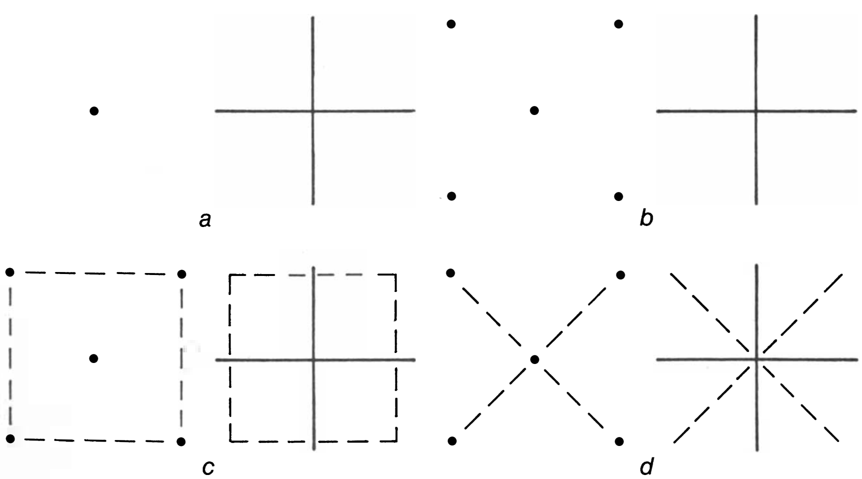
Only the subjects in the perception and imagery conditions showed a significant reduction in pointing error. Moreover, their rates of adaptation were similar. The results for pointing aftereffects, which take place when the glasses are removed, provide additional support for a functional equivalence between observed and imagined errors. Although the aftereffects in the imagery condition were smaller than those in the perception condition, the subjects in both groups pointed to the left of the red marker when normal viewing conditions were restored. I also found evidence of intermanual transfer: the subjects not only pointed to the left with their right hand (the "adapted" hand) but also pointed to the left with their left hand (the "unadapted" hand).

The findings have several implications. First, it is highly improbable that the subjects could have predicted the adaptation and transfer characteristics of prism-induced changes in their visual-motor coordination. It also seems unlikely that they could have had any related kinds of visual experiences providing them with unconscious knowledge of such effects. Second, the findings show that mental imagery can produce certain changes in visual-motor coordination that persist even after the images are no longer formed. They also suggest that the utilization of mental imagery to precipitate such changes may have important practical applications. Professional athletes, for example, often report they find it helpful to rehearse their performance mentally; in the light of these experiments it is reasonable to expect that the success of such techniques depends on the clarity and accuracy with which the performance is imagined.

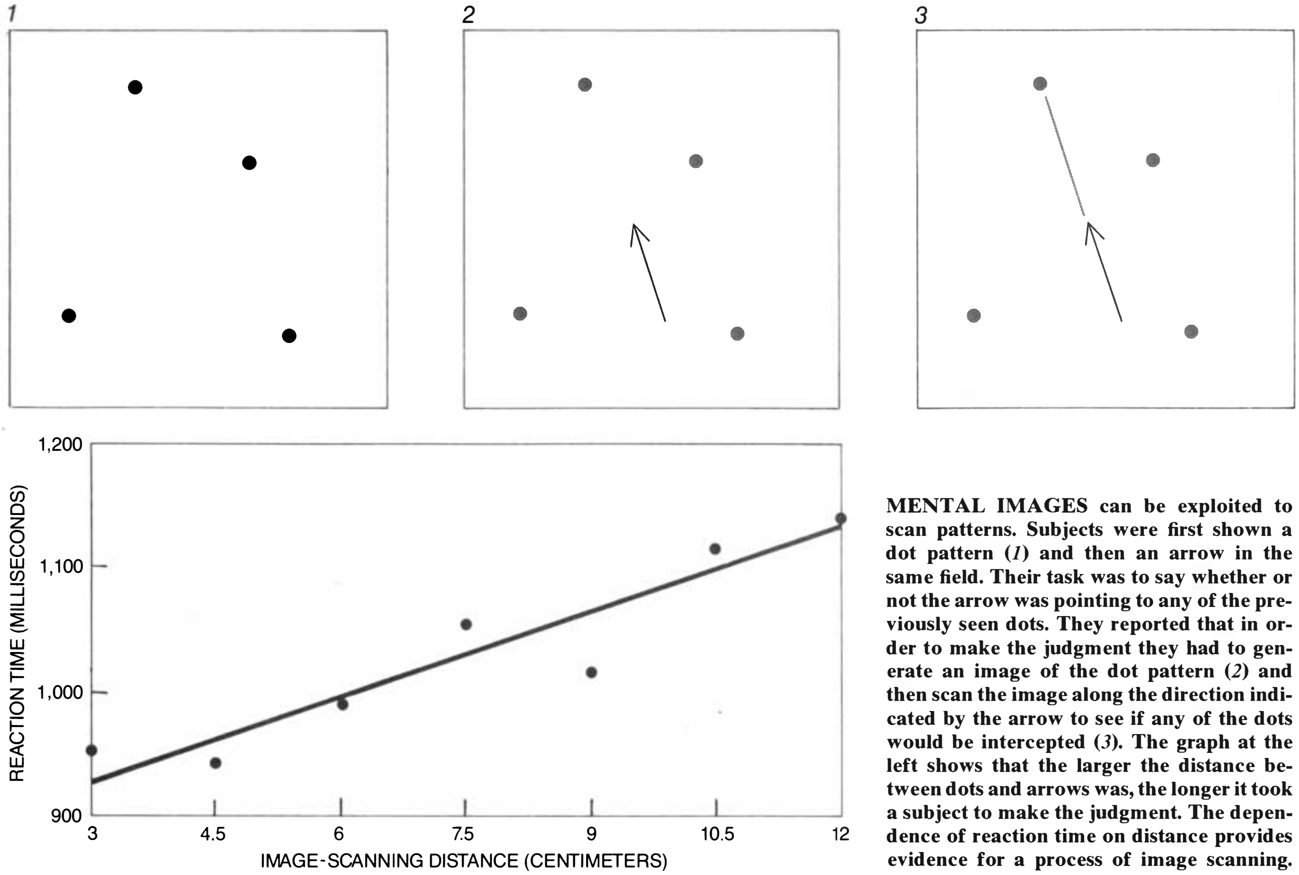
The research findings I have discussed so far illustrate many ways mental images can correspond functionally to physically perceived objects and events. A question of greater practical significance is whether mental imagery can directly facilitate ongoing perceptual processes, assuming that a functional equivalence occurs. Given that similar constraints are imposed on visual resolution in imagery and perception, would it be possible, for instance, to see an object more quickly if an appropriate mental image of the object were formed in advance of its actual appearance?

As proposed some 10 years ago by Ulric Neisser and Lynn A. Cooper, then at Cornell, and Roger N. Shepard of Stanford University, the process of forming a mental image can serve a perceptual anticipatory function: it can prepare a person to receive information about imagined objects. Mental imagery may therefore enhance the perception of an object by causing the selective priming of appropriate neural mechanisms in the visual system. In other words, forming a mental image of an object might initiate certain neural events that are equivalent to those occurring at the moment the object is seen, thereby facilitating the perceptual process.

If an object appears that is different from the one imagined, however, the formation of the image might interfere with the normal operation of the visual system. Suppose, for instance, that you are flying an airplane through



DISCRIMINATION TASK was assisted by mental imagery. Subjects were asked to determine whether a horizontal or a vertical line was the longer. In condition *a* the subjects looked at a fixation point and were then shown the two lines centered at that point. In *b* they were first shown four dots surrounding the fixation point. In *c* they were shown the same dots as in *b* and were asked to form a mental image of a square frame connecting them. In *d* they were asked to form an image of an × through the four dots. The greatest facilitation in the length-discrimination judgments came from imagining the square frame in advance.



MENTAL IMAGES can be exploited to scan patterns. Subjects were first shown a dot pattern (*1*) and then an arrow in the same field. Their task was to say whether or not the arrow was pointing to any of the previously seen dots. They reported that in order to make the judgment they had to generate an image of the dot pattern (*2*) and then scan the image along the direction indicated by the arrow to see if any of the dots would be intercepted (*3*). The graph at the left shows that the larger the distance between dots and arrows was, the longer it took a subject to make the judgment. The dependence of reaction time on distance provides evidence for a process of image scanning.

a cloud. You might see the runway sooner if you were to imagine it in advance at its proper location. If, on the other hand, you imagined the runway at a different location, you might take longer to see it correctly than if you had not imagined it at all.

Some recent experiments I did at Stanford help to clarify the practical relation between image formation and object perception. In one of the experiments subjects indicated whether they saw either a horizontal bar or a vertical bar on a circular screen. These two alternative bars were known as target bars. I told the subjects to form, in advance, a mental image of an identically shaped bar oriented somewhere between the horizontal and the vertical, or to form no mental image. In each trial, therefore, one of the two alternative target bars was superposed on an imagined bar (in those cases where a bar was imagined). I recorded the reaction time for identifying the target bar as a function of the relative alignment between it and the imagined bar. The reaction times for the no-image trials served as a baseline measure.

The subjects made the quickest bar identifications when the imagined bars were closely aligned with the target bars, within a range of about 10 degrees. As the angle between the imagined and the target bars was increased to 45 degrees the identification time also increased. For angles greater than 45 degrees the time decreased once again. In other words, the maximum interference in identification took place when the bars had been imagined to lie exactly in between the two possible target orientations.

Why did the reaction times simply not increase in direct proportion to the degree of misalignment between the imagined and the target bars? One reason is that the subjects' selection of responses could have been based on a comparison between the mental image and the target. If the image matches the target, the response corresponding to the image orientation is quickly selected. If the image and the target differ by 90 degrees, the comparison indicates that the response opposite to the one corresponding to the image orientation is correct. If the imagined bar is in between the two target orientations, however, the comparison becomes confusing and the image interferes with the decision process.

A second experiment supports this explanation. In it I instructed the subjects to indicate as quickly as possible whether either of the two target bars appeared. In the previous experiment they had to distinguish between the two targets; in this case they only had to detect the presence of any target, without having to identify it. The results of this experiment clearly show that mental imagery did not affect simple detection judgments under those conditions. It seems, therefore, that even though image formation may influence the identification of visual patterns, it may not influence the more elementary process of simply detecting any stimulus change.

Additional experiments that Jennifer J. Freyd and I carried out at Stanford provide evidence for a kind of image facilitation that cannot be explained on the basis of response selection. In these experiments we studied the effects of forming a mental image that could serve as a helpful or unhelpful visual context for making difficult length discriminations. We presented our subjects with patterns consisting of two straight lines that formed a simple cross and asked them to indicate which line was longer. At the beginning of certain trials we told the subjects to form an image of an outlined square, which if actually superposed on the center of the line pattern would have enhanced the small differences in the lengths of the lines. During other trials we told the subjects to form an image of an × (the endpoints of which corresponded to the four corners of the imagined square), which would presumably not have been as useful for making the discriminations.

We found that forming a mental image of the helpful context pattern (the square) reduced the time needed to make the length discriminations compared with the time required when the unhelpful pattern (the ×) was imagined or when subjects were not told to form images. Moreover, the effects were similar to those obtained when the same context patterns were actually presented.

We also found that subjects often chose to imagine the helpful context pattern when they were presented with positional cues they could use to imagine either pattern. Since the context patterns themselves could not have biased the selection of the two response alternatives, this type of image facilitation could not have resulted from an internal matching or response-selection process. Instead it may be due to a mental synthesis of real and imagined features at some higher level of the visual system—where the addition of context information can enhance differences among objects that are being compared.

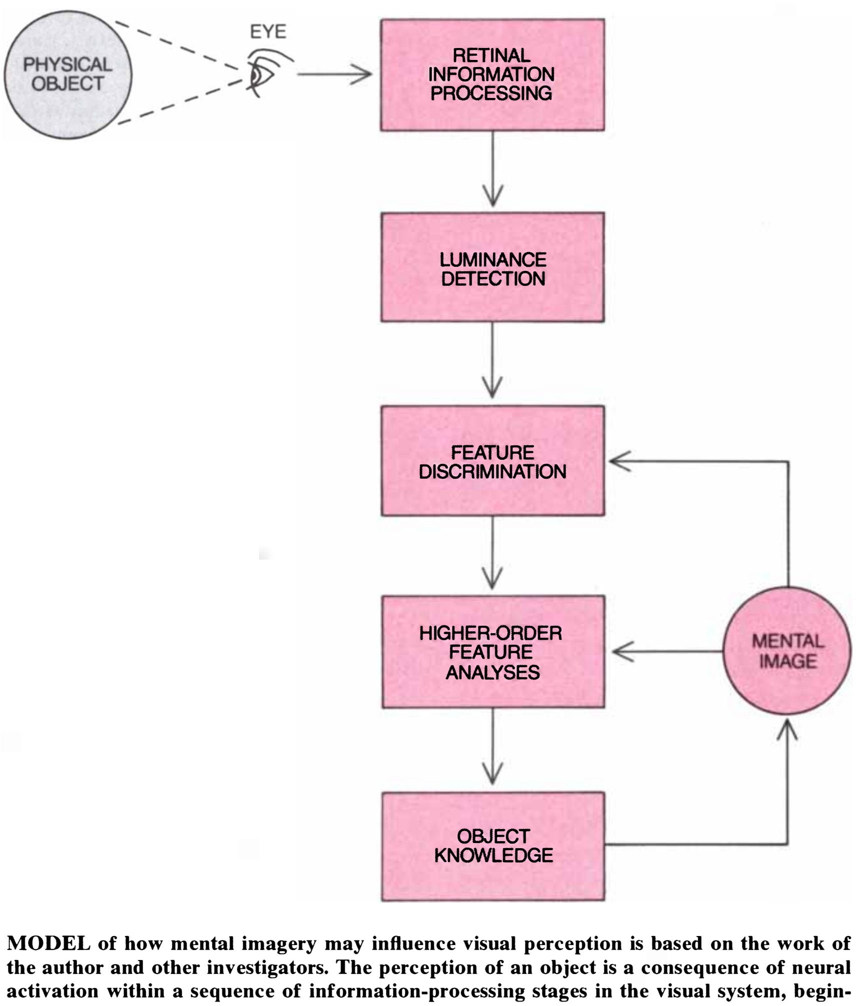
In each of the techniques described up to this point experimental subjects were told explicitly to form some kind of mental image. A possible difficulty with this procedure is that it may encourage the subjects to try to perform as they would in a corresponding perceptual task, thinking that is what they are supposed to do. Although the problem can be largely avoided by attempting to measure subtle or unexpected perceptual effects, an even better way is to show that images can be formed spontaneously for some specific purpose even when no imagery instructions of any kind are given.

The importance of these considerations follows from early studies done by Kosslyn and Pinker on mental-image scanning. They asked subjects to inspect a configuration of objects (such as landmark items drawn on a map), form a mental image of the configuration and "focus in" on one of the objects. The investigators then named a second object and told the subjects to mentally scan along a direct path from the first object to the second. Kosslyn and Pinker consistently found that the time required to complete the mental-image scanning was directly proportional to the original physical distance between the objects, and they therefore concluded that mental images preserve the spatial characteristics of a physical display.

Their findings have been criticized because it would not be hard for experimental subjects to figure out that greater distances should require longer scanning times. Pinker and I have since developed a task that seems to avoid this problem by requiring a subject to form mental images and scan them without explicit directions to do so. After they had inspected a dot pattern, our subjects were shown an arrow and were asked to indicate whether it pointed at any of the previously seen dots. We had predicted that, in order to see if any of the dots would be intercepted, the subjects would have to scan a mental image of the pattern along the direction specified by the arrow.

The experiment turned out to be successful. The decision times increased linearly as distance along the scan path between the arrows and the dots increased. Moreover, nearly all the subjects reported that in order to perform the task they had to form and scan a mental image of the dot pattern. We thus showed that mental-image scanning can be useful whenever it is necessary to anticipate the consequences of moving along a particular path from a given starting point.

Suppose you were trying to figure out where a billiard ball would come to rest on a billiard table after you had aimed it in a certain direction. Even if you could not actually roll the ball across the table or determine the answer mathematically, you could still imagine what would happen by mentally following the motion of the ball and its reflections off the cushions. Cooper and Shepard have reported



MODEL of how mental imagery may influence visual perception is based on the work of the author and other investigators. The perception of an object is a consequence of neural activation within a sequence of information-processing stages in the visual system, beginning at the retinal level. The formation of an image of the object is determined by the knowledge a person has about the features of the object, and presumably it occurs at the very highest levels. Once formed, an image may affect neural mechanisms, at intermediate visual levels, that are responsible for feature discrimination and other more complex types of analyses, thus perhaps modifying perception of the object. Mental imagery probably does not influence visual levels below those concerned with feature discrimination, however.

related findings for the imagined consequences of rotating objects [see "Turning Something Over in the Mind," by Lynn A. Cooper and Roger N. Shepard; SCIENTIFIC AMERICAN, December, 1984].

In the light of these studies it seems reasonable to propose that whenever imagery and perception share common neural mechanisms in the visual system, imagery could facilitate the perceptual processes those mechanisms support. One should therefore seek to determine the lowest visual levels at which such mechanisms may be shared. If visual pattern perception, for instance, is conceived of as involving an orderly sequence of information-processing stages ranging from the lowest to the highest levels of the visual system, one might begin by trying to discover how far down in this sequence image formation can influence the underlying mechanisms.

Starting at the very lowest, or retinal, level, where the most primitive types of information-processing mechanisms are found, one would not expect mental imagery to have much effect. Nor would mental imagery be expected to alter information processing at precortical levels, where mechanisms are responsible for detecting changes in brightness or contrast. Only at the somewhat higher levels responsible for pattern discrimination (as in the visual cortex) does one begin to find evidence that mental images can influence perception. At still higher levels the evidence is strong that imagery can influence perception.

Finally, at the very highest levels one may assume that perceptual processes interact with more abstract processes having to do with knowledge about and understanding of physical objects. Here it is helpful to make a distinction between the form and the function of a mental image. When a person decides to create a mental image of a particular object, the kind of image that can be fashioned depends on the knowledge the person has about the object, such as its size, color and shape. Then once the image is formed it can begin to function in some respects like the object itself, bringing about the activation of certain types of neural mechanisms at lower levels in the visual system. Accordingly, whatever constraints such mechanisms put on the quality of one's perception of the object are also placed on the quality of one's mental imagery. In this way mental images may come to acquire visual characteristics and may serve to modify perception itself.