an Article from **SCIENTIFIC AMERICAN**

AUGUST, 1988 VOL. 256 NO. 8

Perceiving Shape from Shading

Shading produces a compelling perception of three-dimensional shape. One way the brain simplifies the task of interpreting shading is by assuming a single light source

Vilayanur S. Ramachandran

Our visual experience of the world is based on two-dimensional images: flat patterns of varying light intensity and color falling on a single plane of cells in the retina. Yet we come to perceive solidity and depth. We can do this because a number of cues about depth are available in the retinal image: shading, perspective, occlusion of one object by another and stereoscopic disparity. In some mysterious way the brain is able to exploit these cues to recover the three-dimensional shapes of objects.

Of the many mechanisms employed by the visual system to recover the third dimension, the ability to exploit shading is probably the most primitive. One reason for believing this is that in the natural world many animals have evolved pale undersides, presumably to make themselves less visible to predators. "Countershading" compensates for the shading effects caused by the sun shining from above and has at least two benefits: it reduces the contrast with the background and it "flattens" the animal's perceived shape. The prevalence of countershading in a variety of species, including many fishes, suggests that shading may be a crucial source of information about three-dimensional shape.

Painters, of course, have long exploited lighting and shading to convey vivid illusions of depth. Psychologists, however, have not devoted much research to uncovering the mechanisms by which the eye and the brain actually take advantage of shading information. My colleagues and I therefore embarked on a set of experiments intended to reveal what some of the mechanisms might be.

We started out by creating a set of computer-generated displays of simple objects in which subtle variations in shading alone convey the impression of depth. We made sure the images were devoid of any complex objects and patterns, because our goal was to isolate the brain mechanisms that process shading information from higher-level mechanisms that may also contribute to depth perception in real-life visual processing.

Our experiments were based on circular, shaded shapes that create a compelling sensation of depth [*see* a *in upper illustration on page 78*]. The shapes either pop outward like eggs or inward like the cavities of an egg carton. The shapes are ambiguous because the brain does not know from which direction the light is shining. With some effort you can mentally shift the light source to invert the depth of the objects.

Intriguingly, when you mentally reverse the depth of one object, all the other objects in the display reverse simultaneously. This raises an interesting question: Is the propensity for seeing all objects in the display as being simultaneously convex (or concave) based on a tendency to see all of them as having the same depth or is it based on the tacit assumption that there is only one light source? To find out, we created a display in which objects in one row are mirror images of objects in the other row [*see* b]. In this display, when subjects see one row of objects as convex, they always perceive the other as concave.

We drew two conclusions from this simple experiment. First, the derivation of shape from shading cannot be a strictly local operation; it must be a global process involving either the entire visual field or a large portion of it. Second, the visual system seems indeed to assume that only one light source illuminates the entire image. This may be because our brains evolved in a solar system that has only one sun.

Another manifestation of this rule is seen in a complex shape suggesting a white tube lit from the side [*see* c]. The shape nearly always appears convex, perhaps because of subtle cues such as the occlusion of one part of the tube by another, or because of a general tendency to see such shapes as convex. Interestingly, the depth of the two disks superposed on the tube is no longer ambiguous; one is clearly a bump and the other a cavity. Apparently certain features of an object can inform the brain about the direction of illumination, and the depth of other parts of the object are then made to conform to the light source.

The visual system not only assumes a single light source but also tends to assume, naturally enough, that the light comes from above. We vividly demonstrate this effect with a display in which one group of shaded circles is simply the upside-down version of another group [*see lower illustration on page 78*]. Subjects always perceive group *a* as consisting of spheres and group *b* as consisting of cavities. If you turn the page upside down, you will find a striking reversal of depth: the objects in group *b* now appear convex and those in group *a* appear concave. (You can amuse yourself by cutting out the illustration and mounting it on a turntable. How fast can you spin the turntable before you stop seeing the reversals?)

These observations suggest that the brain assumes the sun shines from above. But how does the brain know

VILAYANUR S. RAMACHANDRAN is professor of psychology at the University of California, San Diego, and has a joint appointment as a visiting associate in biology at the California Institute of Technology. He obtained an M.D. at the University of Madras and a Ph.D. in neurophysiology from the University of Cambridge. He has held visiting appointments at several institutions, including the University of Oxford, where he studied the development of binocular neural mechanisms in cats and sheep. His current research goal is to find physiological mechanisms underlying the perceptual effects described in the article.

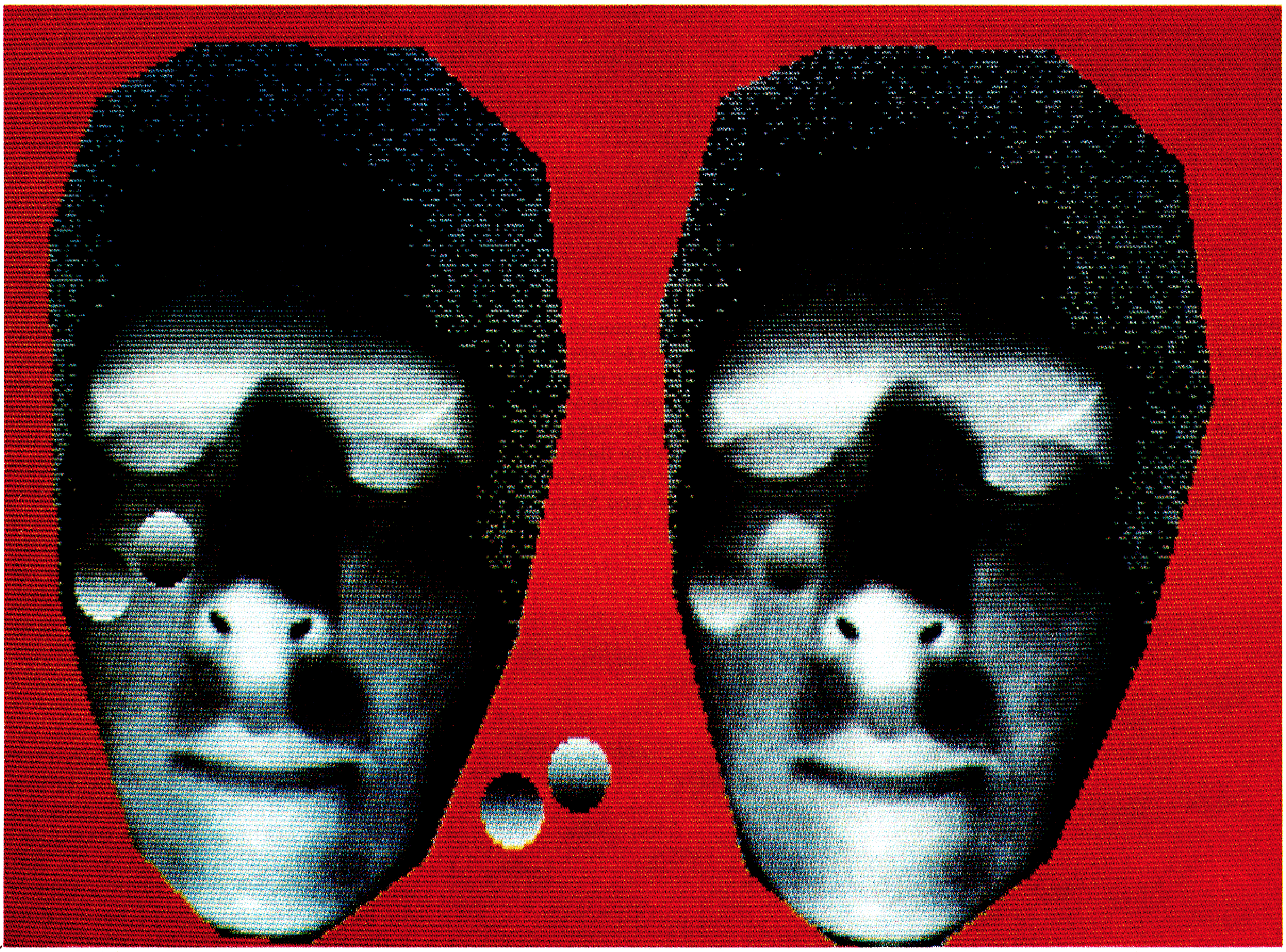
"above" from "below"? Is it the object's orientation in relation to the retina that matters, or is it its orientation with respect to the external world? To appreciate this point try the following experiment. Lie on a couch and let your head hang over the edge so that you are looking at the world upside down. Now ask a friend to stand behind your head and hold the lower illustration on the next page upright. The objects in group *a* will look concave and those in group *b* convex; that is, you get the same effect as you did when you rotated the page. Thus it is the orientation of the object on the retina that matters. Your objective knowledge of up and down does not affect your perception of depth.

Shading by itself generates only a weak impression of three-dimensional shape. To convey a convincing impression of depth the shaded surface must also be enclosed by an outline. Indeed, in many of our displays the luminance variation only roughly approximates the smooth, cosine variation of true shading, and yet the mere presence of a circular outline around the shaded region can generate a compelling illusion of a spherical surface. This raises a new question: What is the exact role of the outline in determining the perception of shape from shading?

To answer the question, we designed a pair of objects that have the same shading but different outlines [*see illustration on page 79*]. Both images have the same luminance gradient: a photocell dragged across each image would register identical variations in the distribution of luminance. Yet the images are strikingly different. The upper image suggests three cylinders lying side by side, whereas the lower image conveys the unmistakable impression of a sheet of corrugated metal. The perceptions seem to depend completely on the contours of the top and bottom edges of the surfaces.

We conclude from these demonstrations that when shading cues are ambiguous, information from borders helps to resolve ambiguity throughout the image. Interestingly, the perceived location of the light source also shifts to conform to the perceived surface. In the upper image on page 79 the light seems to originate perpendicular to the page whereas in the lower image the illumination is from the far left or the far right. It is remarkable that changing an object's boundaries can produce such striking changes in perception.

Our next demonstration shows that even illusory contours will work. A typical example consists of four dark gray disks with a "bite" taken out of each one [*see top illustration on page 80*]. When the disks are in proper alignment, one has the impression of a



HOLLOW-MASK INTERIORS lit from above produce an eerie impression of protruding faces lit from below. In interpreting shaded images the brain usually assumes light shining from above but here it rejects the assumption in order to interpret the images as normal, convex objects. Notice the two disks near the chin still appear as though lit from above: the right disk seems convex and the left one concave. When the disks are pasted on the cheek (*left*), their depth becomes ambiguous. When blended into the cheek (*right*), the disks are seen as being illuminated from below, like the rest of the face.

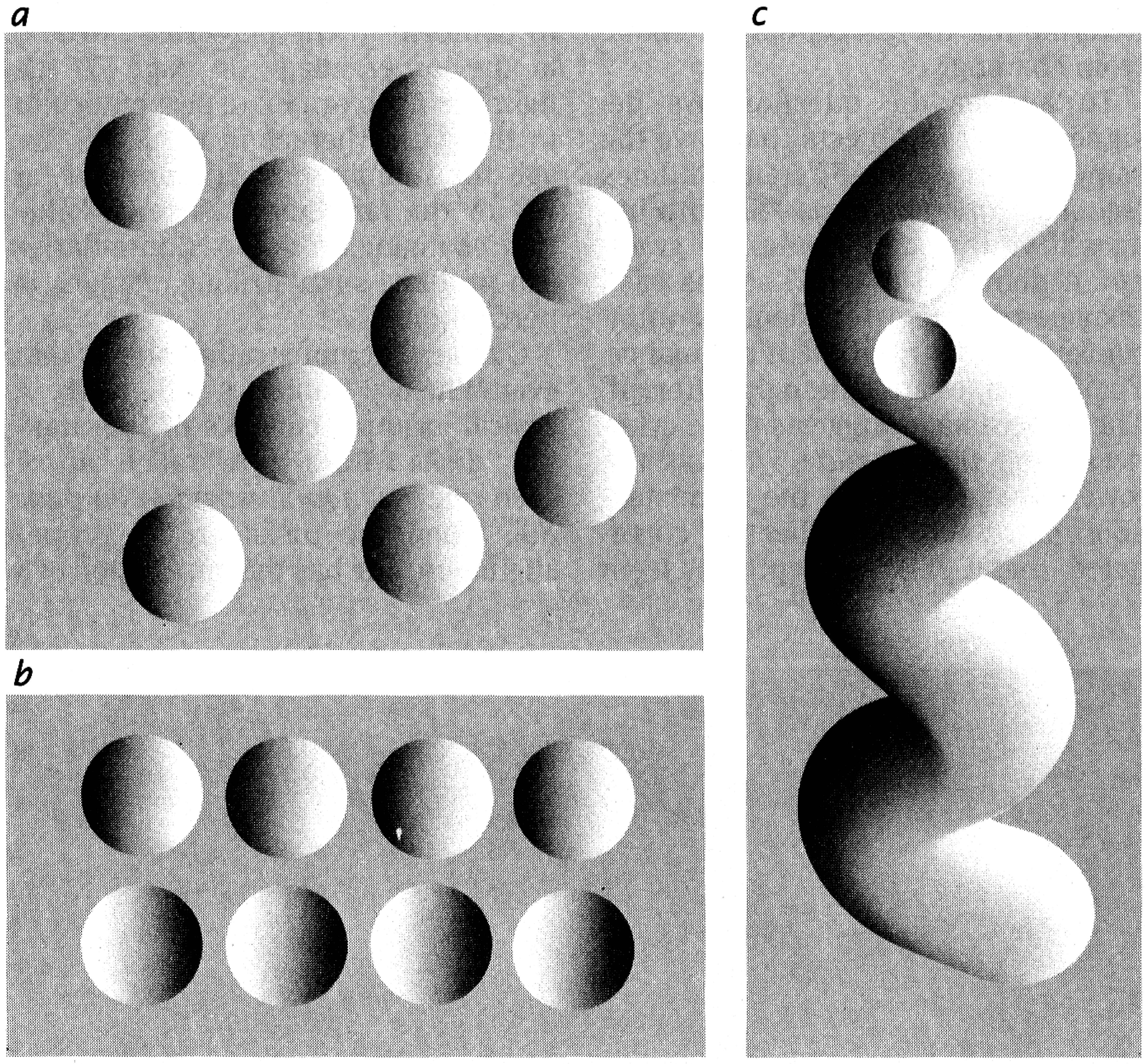
large pale disk at the center partially occluding the gray disks. Indeed, faint lines seem to connect the concave edges of the gray disks, although such lines do not exist physically.

What happens if we replace the background of this display with one in which the luminance varies from top to bottom? The new display looks flat at first, but on prolonged viewing the region inside the illusory disk starts to bulge out toward the observer and may even detach itself from the background to take on the appearance of a floating sphere. Oddly enough, an illusory contour seems to work even better than a real outline. The reason is not entirely clear but the result suggests that the brain regards partial occlusion as stronger evidence for the existence of an object than the presence of a mere outline. After all, the outline might equally well depict a loop of thin wire or a transparent soap bubble.

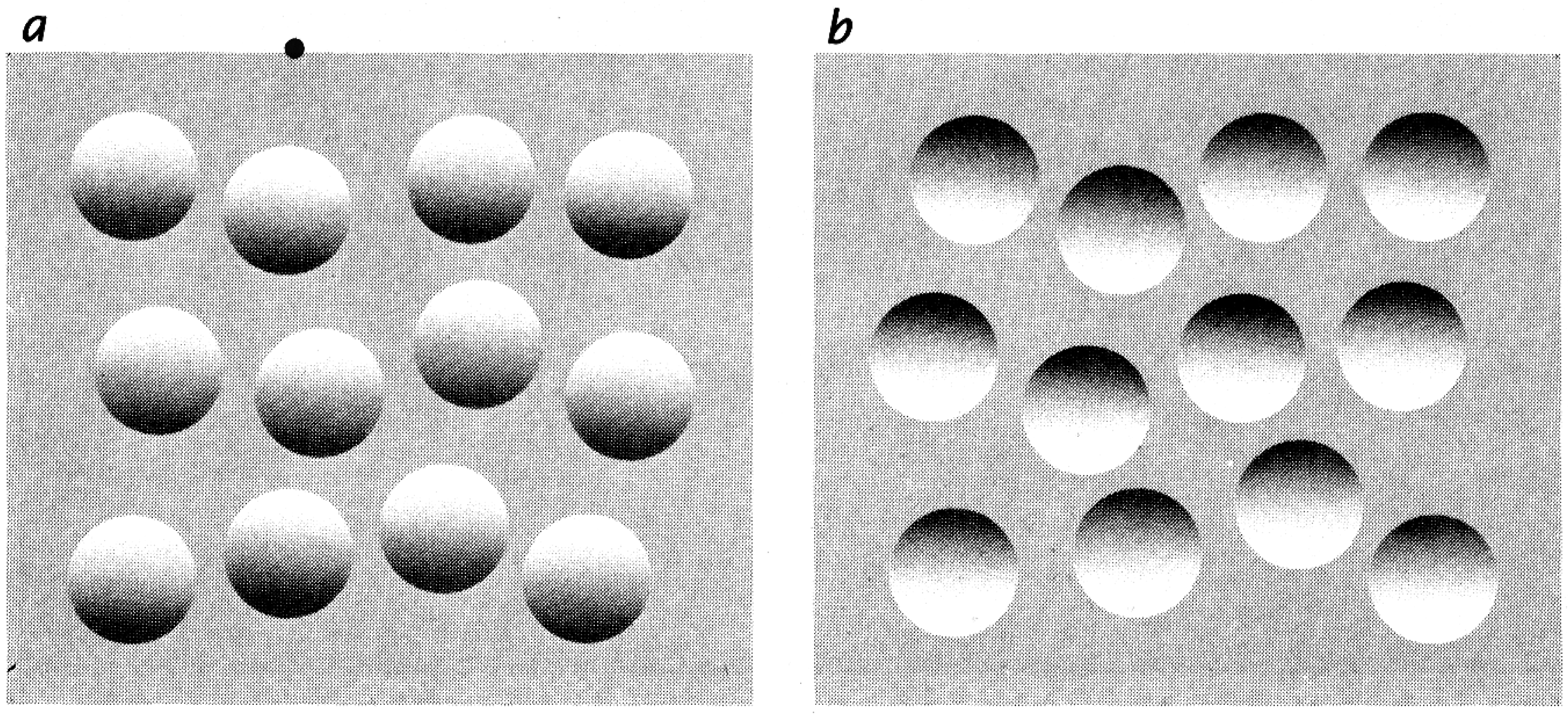
This observation, like the preceding one, demonstrates a direct and powerful interaction between edges, whether real or illusory, and the derivation of shape from shading. If the visual system were making detailed measurements of shading alone to recover surface orientation (as is implied in some artificial-intelligence models of vision), one would not see a sphere in the image, because the shading does not change at all across the illusory border. Yet the visual system perceives a sphere because the shading and the illusory outline mutually reinforce that interpretation.

Another way the visual system delineates objects is by changes in surface reflectance, or the proportion of light reflected by surfaces. A photocell moving across an object's border will usually register an abrupt shift in luminance. What would happen if the outline were defined by a change of color rather than a change of luminance? We took a typical shaded "sphere" and replaced the homogeneous gray background with a colored background in which the luminance gradient matched that of the sphere. The result was dramatic: the illusion of depth dissolved and the sphere appeared flattened, even though its outline was distinctly visible because of the contrast in hue. We concluded that the shape-from-shading system cannot make use of edges defined by color differences. One reason may be that our primitive primate ancestors, which resembled tarsiers, were nocturnal and color-blind; in their twilight world they relied on luminance contrast alone to perceive depth.

These demonstrations imply that the brain recovers information about the shape of objects by combining outlines and shading cues. What does the brain do with the shapes once it has recovered them? An important capacity of perception is the ability to segregate figure from ground. Even in a cluttered scene the visual system can easily decide which features in the image belong together to form objects. In a high-contrast photograph one can see a Dalmatian



SPHERES OR CAVITIES? It depends on where you think the light source is. You can reverse the depth of the objects (*a*) by mentally shifting the light source from left to right. In a second array (*b*) each row by itself is ambiguous, but once you see one row as convex the other row will always appear concave. It is almost impossible to see both as simultaneously convex or concave. The convoluted form (*c*) suggests a white tube lit from the right. The two disks on it seem to conform to the lighting scheme; the top disk is seen as a bump and the bottom one as a cavity. The last experiment was done by the author in collaboration with Dorothy Kleffner and Steven J. Cobb.



BRAIN ASSUMES light comes from above. The objects in group *a* therefore appear convex, whereas those in group *b* appear concave. If you turn the page upside down, the objects will reverse in depth. By turning your head upside down and looking at the page, you can prove it is the orientation of the pattern on the retina that matters.

dog against a dappled background [*see bottom illustration on next page*]. Similarly, one can mentally "lift out" a group of lines that have a particular orientation from a field of lines with a different orientation. On the other hand, it is impossible to segregate a group of mirror-reversed letters from unreversed ones.

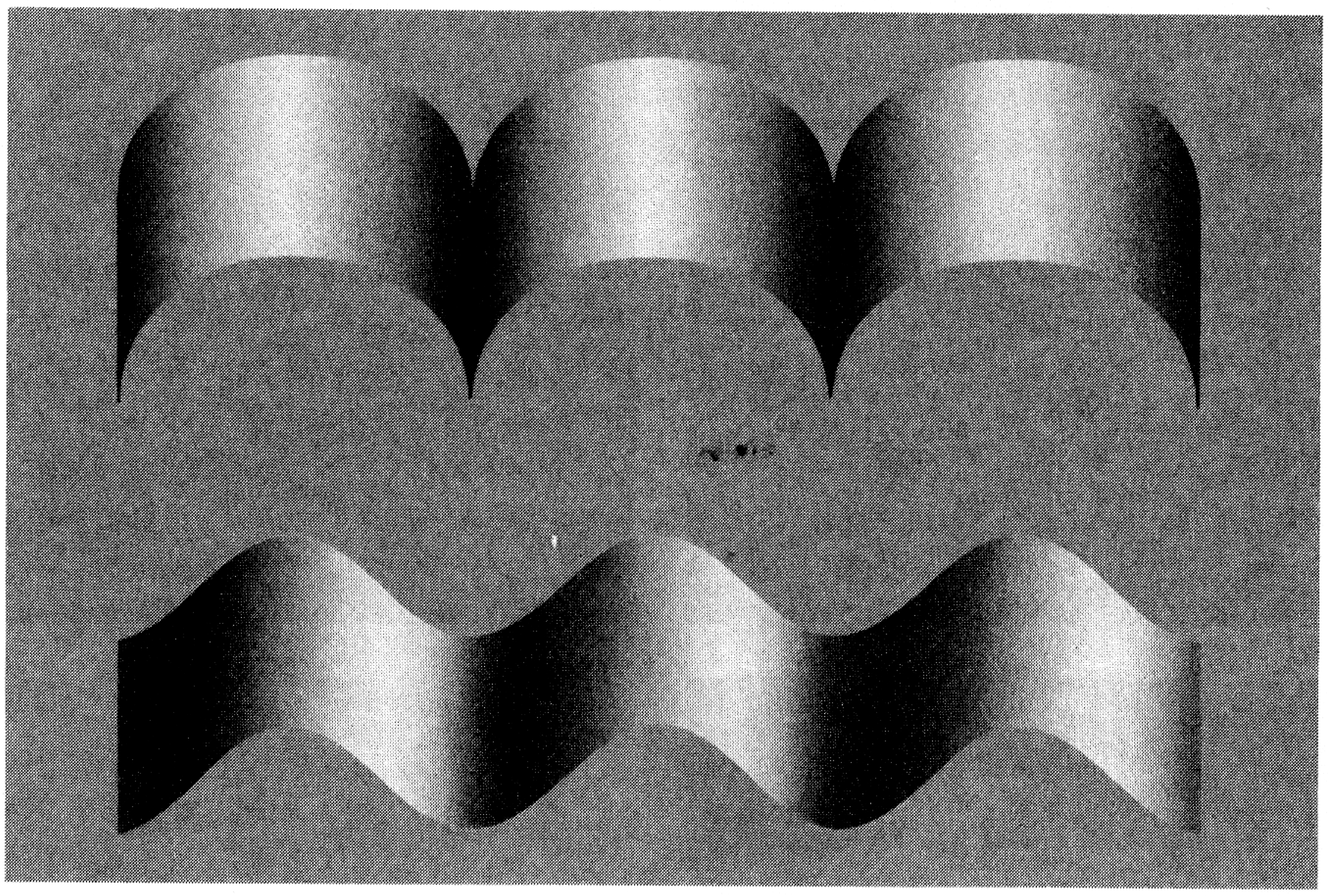
The laws of perceptual grouping were first studied systematically by Anne M. Treisman of the University of California at Berkeley, Bela Julesz of the AT&T Bell Laboratories and Jacob Beck of the University of Oregon. These investigators discovered several important principles. First, they found that an important early stage of visual perception involves extracting certain elementary features, which Julesz calls textons. Examples include oriented edges, color and direction of movement. Once the visual system has extracted the elementary features, similar features are grouped together to form objects. Indeed, Beck suggests that only elementary features, by definition, can be grouped in this way. Presumably, then, alphabetic characters are not elementary features as far as the visual system is concerned.

What about three-dimensional objects, though? Our next several demonstrations show that even shapes defined exclusively by shading can serve as elementary features of visual perception. In an array of cavities interspersed with convex shapes, for example, the convex shapes can mentally be grouped together to form a separate depth plane that is clearly segregated from the concave shapes in the background [*see illustration on page 81*].

When one views this display, it appears as though the visual system passes through several stages of processing. In the earliest stage the system performs computations for defining the three-dimensional shapes, taking several seconds. Once the convex shapes have emerged, one has the distinct impression of being able to "hold on" to them indefinitely in order to group them with similar items in the display. Finally, after the objects are grouped, they are clearly segregated from irrelevant items in the background. The extraction and grouping of textons, then, although usually described as a one-step operation, may in fact involve several distinct perceptual capacities that act together to delineate figure from ground.

We wondered whether the perceptual grouping observed in that display might be the result of some other, more elementary feature than the three-dimensional shape. For example, because the convex shapes differ from the concave ones in the polarity of their bright-to-dark luminance, one might suppose the grouping is achieved by latching on to luminance polarity. To rule out this possibility, we created a display of objects that have the same luminance polarities as those in the preceding display but that do not carry any depth information. It is virtually impossible to achieve perceptual grouping in this display. Even after you have spotted all the targets individually, you will not be able to segregate them from the rest of the objects. Clearly the grouping observed in the preceding display must be based on three-dimensional shape rather than on luminance polarity.

I have pointed out that the illusion of depth is much more powerful when the illumination is from above than when it seems to come from the side. Similarly, lighting from above greatly enhances one's ability to group and segregate images. You can verify this by simply rotating group *a* on page 81 by 90 degrees: the impression of depth will diminish and there will be a considerable reduction in perceptual segregation. This further supports the idea that perceptual grouping must be based on three-dimensional shape. Moreover, these groupings can themselves represent higher-level shapes, such as a triangle. It might be interesting to employ stimuli of this kind to find out whether infants and brain-damaged patients can perceive shape from shading; for example, would an infant respond to spheres arranged to suggest a face?

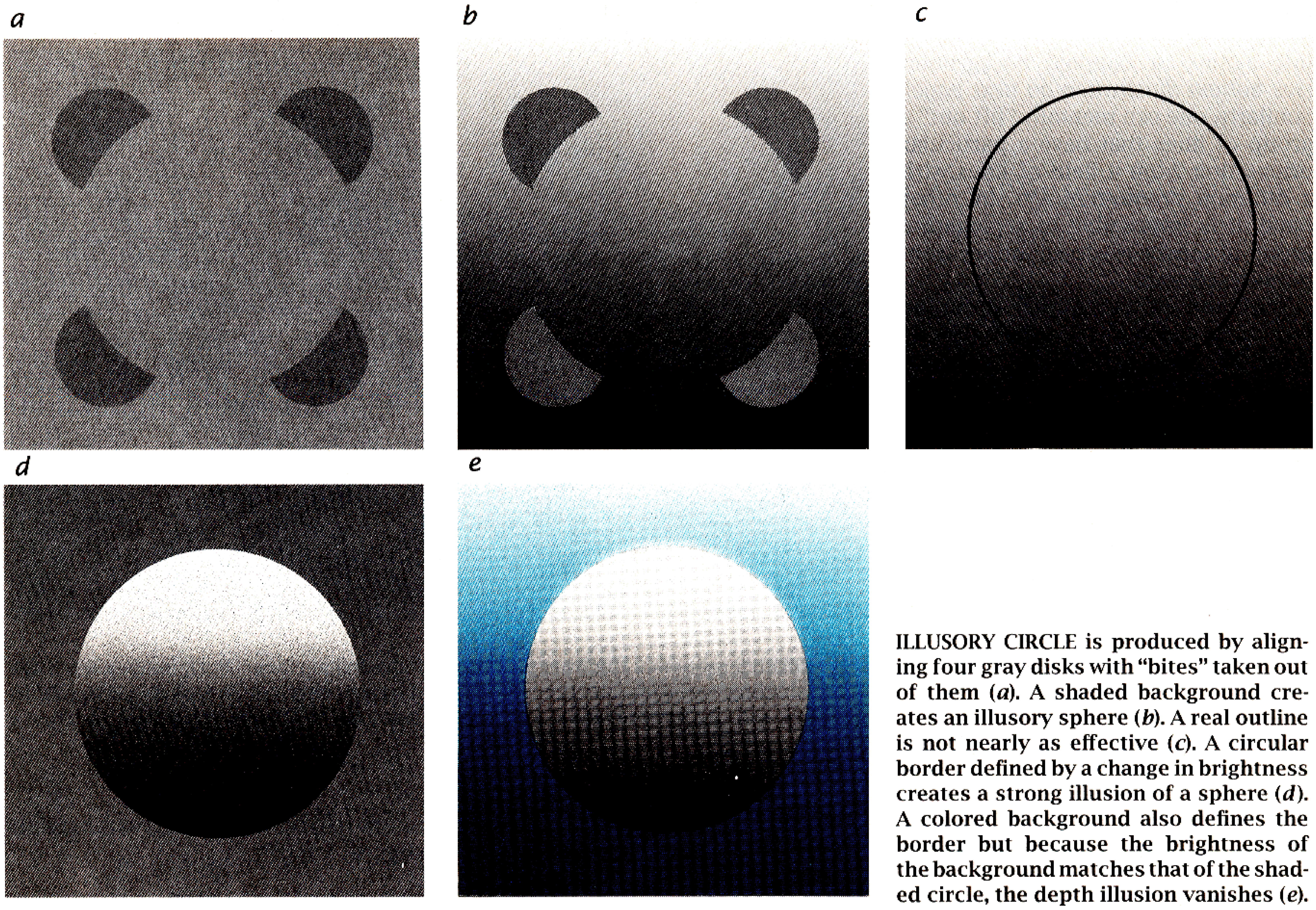


BOUNDARIES influence the interpretation of shaded surfaces. Both images have the same shading variation but the top image suggests three cylinders lit vertically to the page and the bottom one a corrugated metal sheet lit from far left (or far right).

Another remarkable capacity of visual perception is the ability to detect symmetry. This ability extends to fairly complicated shapes, such as plants, faces and Rorschach inkblots. How does the visual system detect symmetry? Does it match all the individual features on one side with those on the other side to determine whether an object is symmetrical? Or does it group features into more meaningful shapes and then look for symmetry in those shapes? Our next demonstration is an attempt to answer these questions.

We compared two arrays of shaded circles [*see top illustration on page 82*]. Subjects usually perceived the left-hand array as spheres and cavities arranged symmetrically about a horizontal axis. Yet a point-by-point examination reveals that the bottom half of the array is not a mirror image of the top half. In fact, it is the array on the right that is truly symmetrical. These results imply that the perception of symmetry is based on three-dimensional shape rather than on the simple distribution of bright and dark areas in the image. You can verify this by rotating the illustration 90 degrees to eliminate the strong impression of depth. You will now see that the right-hand array is more symmetrical than the left-hand one.

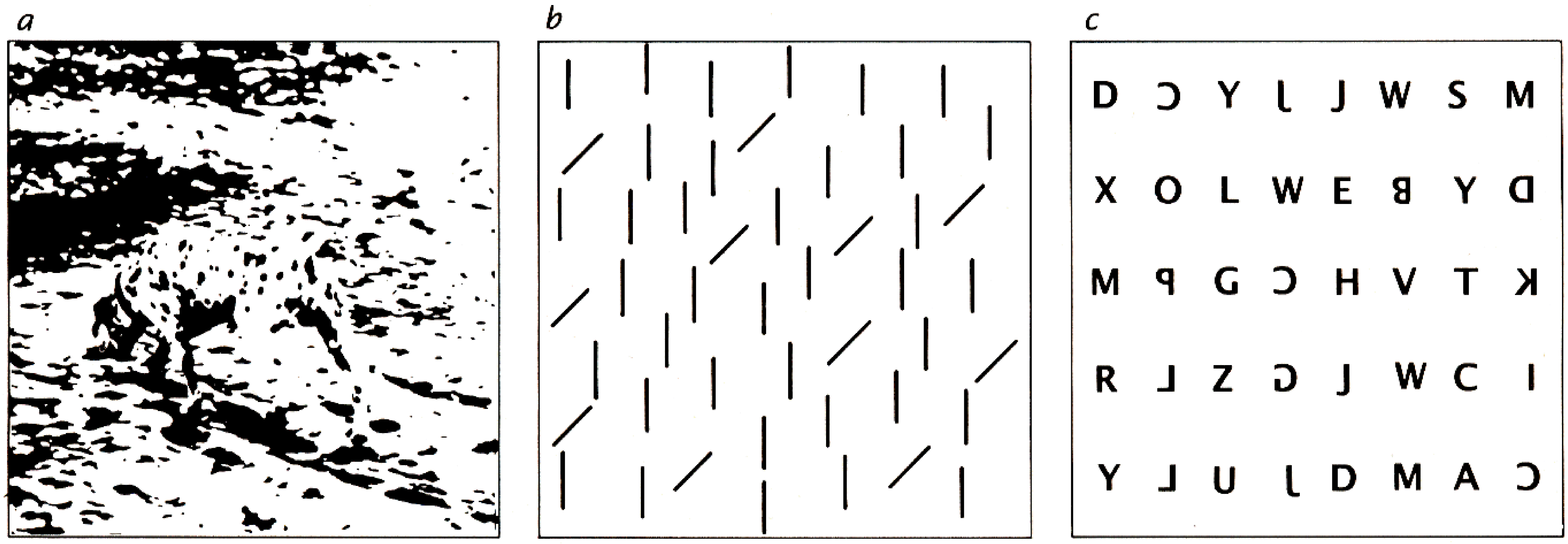
Our observations suggest that shading information is extracted fairly ear-



ILLUSORY CIRCLE is produced by aligning four gray disks with "bites" taken out of them (*a*). A shaded background creates an illusory sphere (*b*). A real outline is not nearly as effective (*c*). A circular border defined by a change in brightness creates a strong illusion of a sphere (*d*). A colored background also defines the border but because the brightness of the background matches that of the shaded circle, the depth illusion vanishes (*e*).

ly in visual processing. Indeed, there may even be neural channels specifically committed to the purpose. Recently Terrence J. Sejnowski and Sidney R. Lekhy of Johns Hopkins University raised the possibility that such cells may exist, based on work with a computer simulation. They began with a "neural network" consisting of three layers of cells: an input layer, a hidden layer and an output layer. Input-layer cells were modeled on the circular, "center surround" receptive fields of cells in a cat's eye. A learning algorithm adjusted the strength of signals passing from cells in one layer to the next, and after 40,000 trials the network could correctly associate shaded shapes with their three-dimensional axes of curvature.

What happened next came as a surprise: the investigators examined the responses of the cells in the hidden layer and found that they responded to bars of various lengths, widths and orientations, bearing an uncanny resemblance to edge-detector cells found in the visual cortex of cats and monkeys. Intriguing as this computer simulation is, its biological relevance is still not clear because the investigators deliberately excluded outlines and other cues known to play a crucial role in human vision. It remains to be seen whether the resemblance between the hidden units and the cor-



PERCEPTUAL GROUPING of elementary features enables one to segregate the shape of a Dalmatian dog from a speckled ground in a high-contrast photograph (*a*). Slanted lines can be grouped and envisioned as occupying a separate plane from vertical lines (*b*). Mirror-reversed letters, however, cannot be visually grouped and segregated from normal letters (*c*).

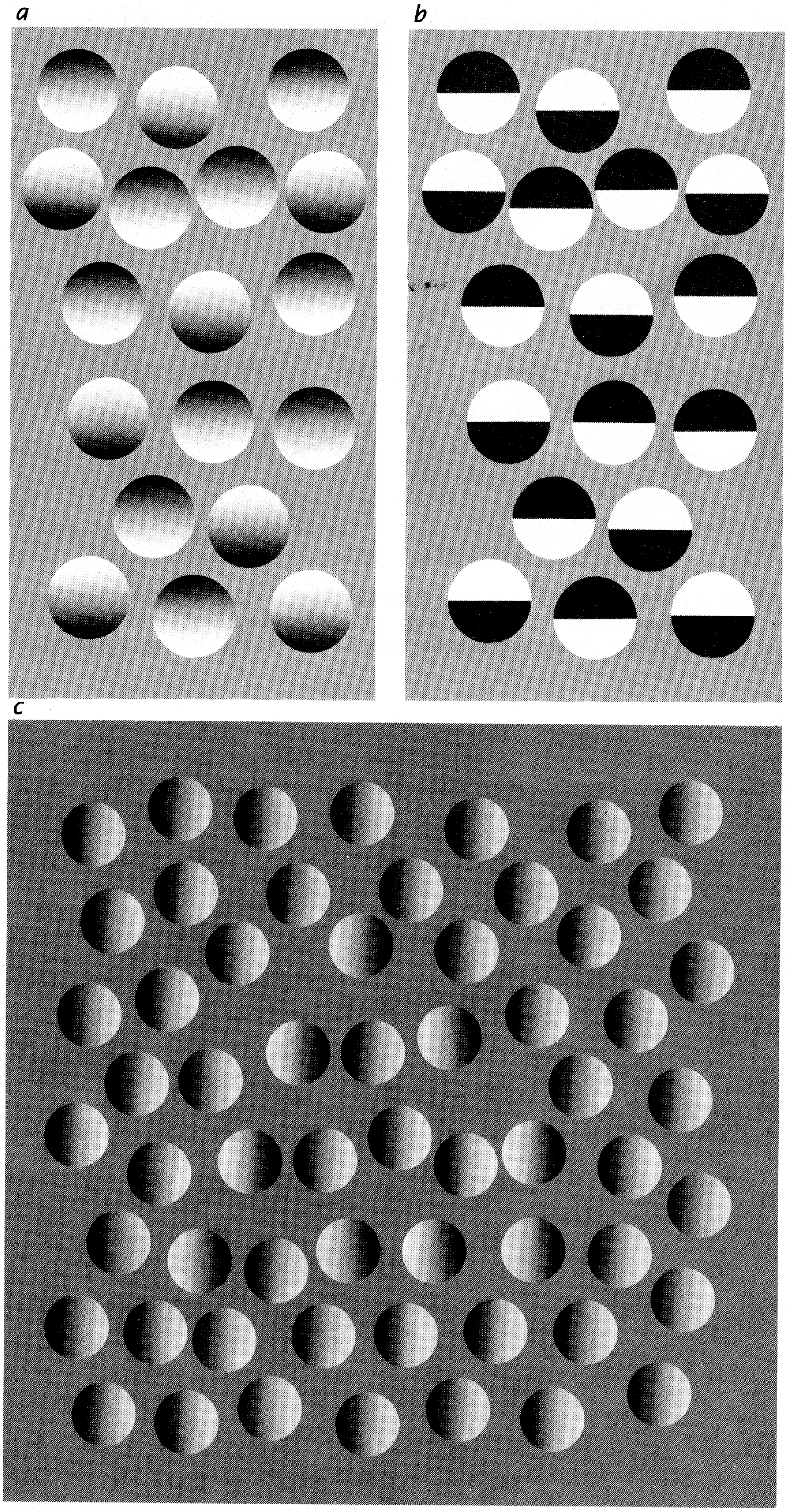
tical edge detectors is merely a coincidence or whether edge-detector cells actually serve to extract three-dimensional shapes from shading.

I have so far considered stationary images, but what about moving objects? In nature it is a reasonably safe bet that anything that moves is either prey or predator. Consequently the visual system appears to have evolved a wide variety of mechanisms for detecting movement. Evidence suggests that the ability to see movement is mediated by specialized groups of brain cells. Can the brain mechanism enabling us to perceive motion also take advantage of information provided by shading? In order to find out we decided to exploit a well-known illusion called apparent motion [see "The Perception of Apparent Motion," by Vilayanur S. Ramachandran and Stuart M. Anstis; SCIENTIFIC AMERICAN, June, 1986].

A simple example of apparent motion is produced by flashing two spatially separated spots of light in rapid alternation. Instead of seeing two lights flashing on and off one usually sees a single light jumping back and forth. To investigate the role of shading cues in human motion perception we created a display that alternated rapidly between one frame showing a shaded convex object above a concave one and a second frame in which the objects are reversed. Eleven naive subjects reported seeing a sphere jumping up and down between two holes in the background.

The result suggests that the brain must first compute three-dimensional shape before it can perceive apparent motion. Indeed, subjects often take tens of seconds to develop a depth impression, during which time they see no apparent motion. It therefore seems unlikely that the apparent motion could be based on some other, more primitive feature of the image. To demonstrate the point more directly we rotated the entire display by 90 degrees. This reduced the impression of depth considerably and led to an almost complete loss of the apparent-motion effect.

The visual system, then, appears to extract a three-dimensional object from shading cues and to perceive movement based on the three-dimensional image, rather than using the "primitive" two-dimensional image directly. Certain cells in the visual cortex of the monkey respond to the apparent motion of simple stimuli such as the flashing spots of light described above. It might be interesting to see



VISUAL SYSTEM can pick out convex shapes from concave ones and group them together (*a*). In an array that conveys the same luminance polarities as the preceding one but no depth information (*b*) it is impossible to visually segregate objects. In an array lit from the side (*c*), perceptual grouping becomes easier when you rotate the picture by 90 degrees. The convex objects stand out and form a triangle. Shading can define such complex shapes for further visual processing. A similar idea has been proposed by Alex Pentland of the Massachusetts Institute of Technology.

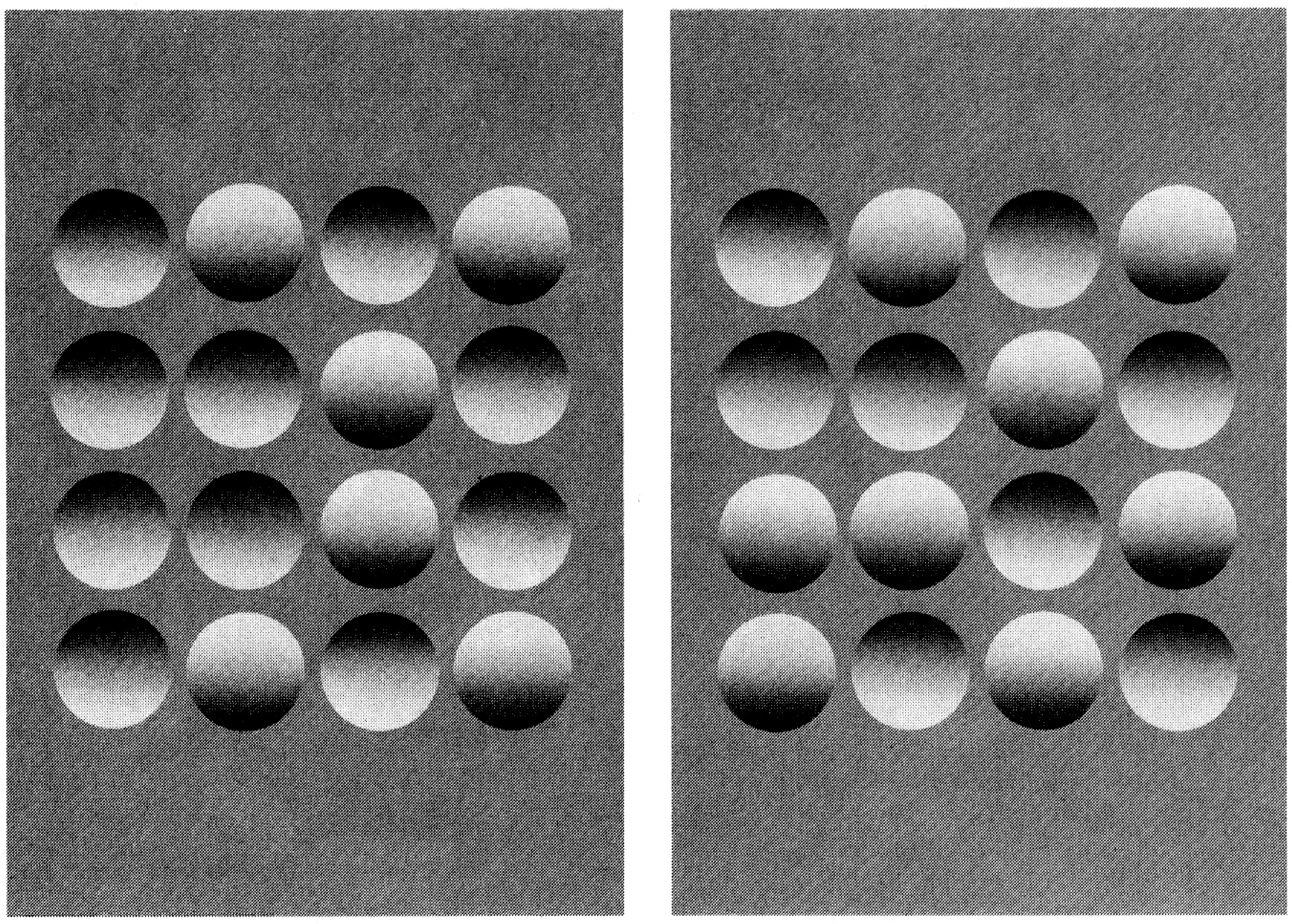
whether these cells would respond to motion based on objects whose shape is perceived from shading.

Clearly, visual perception relies on a constellation of biological processes to arrive at a three-dimensional representation of the world. In order to create this representation the visual system appears to make a variety of simplifying assumptions, such as the rule that there is only one light source. What happens when the visual system tries to construct a coherent scene out of many disparate fragments? Patrick Cavanagh, Diane Rogers-Ramachandran and I recently did a study to try to answer the question.

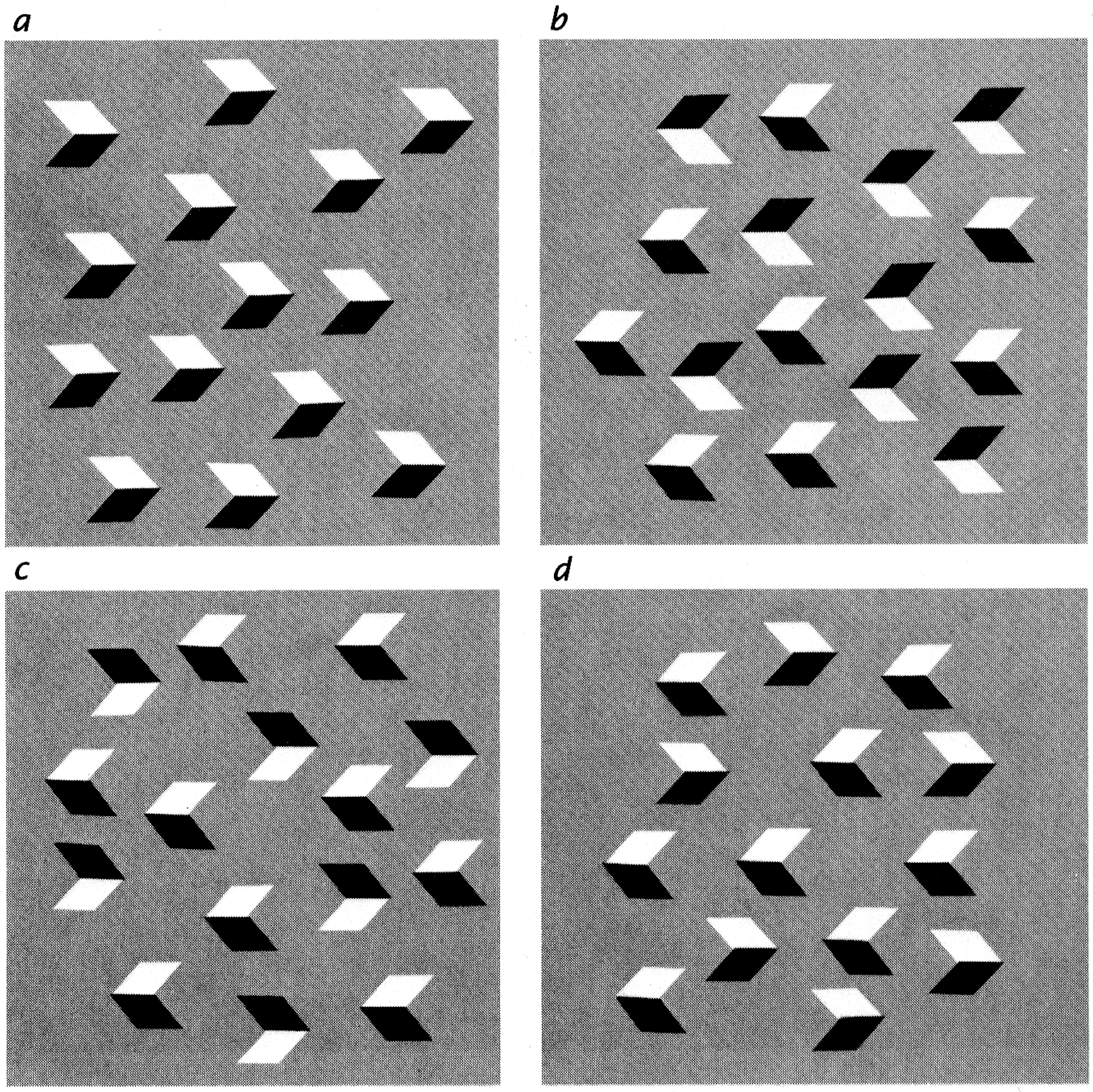
We created simple arrays of randomly placed chevrons, each of which can be viewed as two adjoining faces of a cube [*see bottom illustration at left*]. Array *a* can be perceived as parallel cubes all pointing in the same direction and illuminated by a single light source; the black parallelograms are seen as the shadowed face of the cubes. But equally often the array is perceived as a set of white "gravestones" casting black shadows. By mentally shifting the direction of the light source you can switch from seeing cubes to seeing gravestones. Note that when you see any one figure in the array as a cube you see all others as cubes too. It is impossible, in fact, to simultaneously perceive some figures as cubes and the others as gravestones, because such a perception would violate the single-light-source rule. Interestingly, when the shapes are perceived as cubes, there is a tendency to fill in the missing faces—that is, to perceive illusory surfaces. The illusory surfaces vanish when the shapes are seen as gravestones.

Next we randomly inverted or reversed roughly half of the chevrons in various combinations. These new displays illustrate the subtle interplay of constraints and organizing rules that occurs when the brain tries to create meaningful shapes from isolated fragments. In array *b*, for instance, all the targets usually appear as parallel cubes even though this would be incompatible with a single light source. Apparently when the single-light-source rule cannot be satisfied, it is replaced by a "pointing" rule (or by a rule stating that shapes with similar orientations are in fact parallel surfaces). To avoid conflict the brain simply assumes that the cubes have faces of differing color.

In array *c* you will see a mixture of gravestones and cubes because this



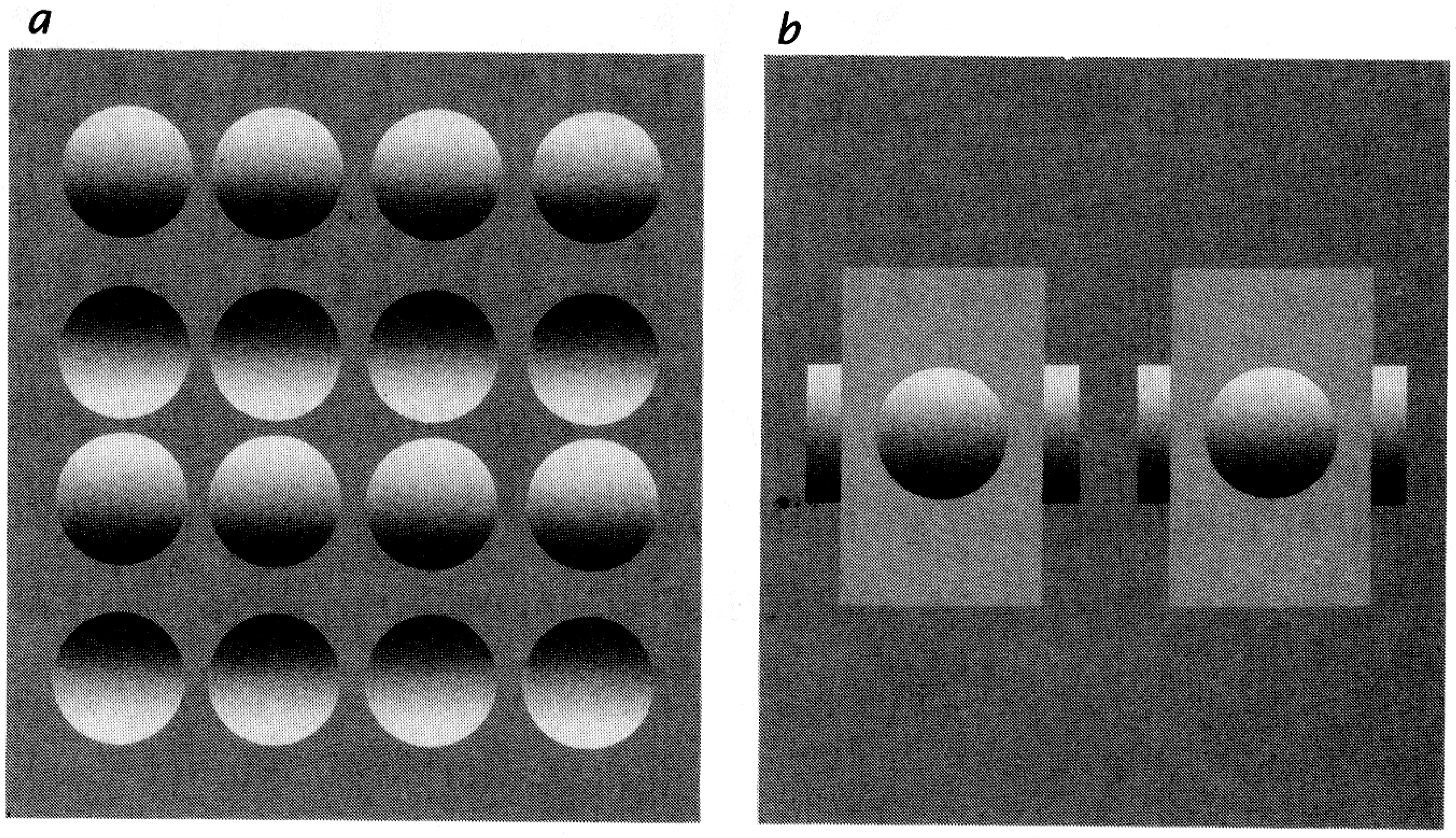
SYMMETRY PERCEPTION occurs after the brain extracts shape from shading. In the left-hand array spheres and cavities seem arranged symmetrically about a horizontal axis. But in two dimensions it is the right-hand array that is truly symmetrical.



CHEVRONS in group *a* appear as illusory cubes or as "gravestones" casting shadows, all lit from the same angle. In group *b* a "pointing" rule seems to override the single-light-source rule and the image is seen as a set of cubes whose faces have different reflectances. Group *c* is always seen as a mixture of cubes and gravestones because this interpretation satisfies the single-light-source rule. Group *d* is ambiguous, and it is difficult to unify the figures into a coherent interpretation.

allows the system to satisfy the single-light-source rule. It is actually impossible to see the display as consisting entirely of cubes or of gravestones, because such a perception would not be compatible with either the pointing rule or the light-source rule. You will also find yourself unifying all the items in the array into a single coherent surface so that it suggests a sculptured metal surface with randomly placed "steps" carved out of it. Whereas an ant crawling on the picture would see only chaotic fluctuations in brightness, the human eye surveys the entire image and knits parallel surfaces together to create spatial order and unity.

In array *d* the figures are neither parallel nor able to satisfy the single-light-source constraint. Hence there is a tendency to see the display as a random collection of flat chevrons pointing in opposite directions. Even though an individual figure in the display can sometimes be seen as a cube or a gravestone, it is difficult to unify all of them into a coherent three-dimensional interpretation.



VISUAL IMAGERY, or higher-level information about objects, profoundly influences the perception of shape from shading. Rows of spheres and cavities (*a*) can also be seen as two blurred stripes visible through 16 holes cut in an opaque sheet. One can no longer see the spherical shapes. Each of two pictures depicts a sphere (*b*), but when they are "fused" in a stereoscopic viewer, the spheres vanish and one sees a circular window cut in a rectangular sheet floating in front of a shaded plane.

In the real world, visual imagery—that is, high-level knowledge about what one is seeing—profoundly affects the perception of shape from shading. Indeed, the interaction between visual imagery and perception is one of the most elusive and enigmatic topics in psychology. To illustrate this point, we created an array of shaded circles that on casual inspection appear to form alternating rows of spheres and cavities [*see illustration on this page*]. The display is susceptible to a radically different interpretation, however: it can be seen as a gray sheet with 16 holes cut out, behind which are two blurred, dark stripes. This perceptual switch causes the shaded circles to lose their spherical shape completely.

The tendency to see stripes rather than spheres and cavities can be enhanced by stereoscopic cues. You can see spheres in *b* on this page, but if you were to "fuse" them binocularly through a stereoscopic viewer, you would see a frame with a circular window standing out clearly from a shaded background. Indeed, it becomes virtually impossible to see the circular shape as a sphere instead of a hole. This implies that the extraction of shape from shading is strongly affected by stereoscopic processing.

The interpretation of shape from shading also interacts strongly with the visual system's knowledge of objects, as is strikingly demonstrated by the opening illustration for this article. In these photographs the hollow insides of face masks are illuminated from above; one would therefore expect them to look hollowed out. But the visual system strongly rejects the possibility of hollow shapes and interprets the images as normal faces lit from below. Thus the visual system overrides the assumption of lighting from above in order to be able to interpret the shapes as normal faces.

Now notice the two small, shaded disks between the chins of the two faces. Even though the light on the faces is assumed to come from below, the disk on the right generally is seen as convex and the one on the left as concave—as though they were both illuminated from above. Perhaps the brain treats these objects as being quite distinct from the faces and therefore, in interpreting their shading, adheres to the more "primitive" rule that they are illuminated from above. When the disks are pasted onto the cheek of one of the faces, however, the depth becomes ambiguous: the right-hand disk can appear concave and the left-hand one convex. Finally, when the outlines of the disks are blended into the cheek, they are always seen as being illuminated from below, like the rest of the face. Consequently the disk at the right suggests a dimple and the one at the left looks like a bump or a tumor.

Our research has revealed a variety of rules that are applied early in the visual processing of shape from shading. We have shown that it is possible to trace the flow of information from the very early stages of shape perception to the final stage, where the information interacts with high-level knowledge of light sources and of the nature of complex, three-dimensional objects. The neurological events mediating the process in human beings are still mysterious, but insights from psychology can help to elucidate what these events may be and how they are organized in the brain. New computational models can also offer plausible mechanisms and help to narrow the search. These developments are launching research on visual perception into a new domain, where it may someday be possible to discover the cellular mechanisms in the brain that enable us to perceive the world visually in three dimensions.

FURTHER READING

THE ROLE OF FRAMES OF REFERENCE IN THE DEVELOPMENT OF RESPONSIVENESS TO SHADING. Albert Yonas, Michael Kuskowski and Susan Sternfels in *Child Development*, Vol. 50, No. 2, pages 495-500; June, 1979.

PERCEPTION OF SURFACE CURVATURE AND DIRECTION OF ILLUMINATION FROM PATTERNS OF SHADING. James T. Todd and Ennio Mingolla in *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 9, No. 4, pages 583-595; August, 1983.

PERCEPTION OF SHAPE FROM SHADING. V. S. Ramachandran in *Nature*, Vol. 331, No. 6152, pages 133-166; January 14, 1988.