The perception of transparent three-dimensional objects

(vision/illusion/visual learning/association)

DALE PURVES* AND TIMOTHY J. ANDREWS

Department of Neurobiology, Box 3209, Duke University Medical Center, Durham, NC 27710

Contributed by Dale Purves, April 9, 1997

ABSTRACT When the proximal and distal elements of wire-frame cubes are conflated, observers perceive illusory structures that no longer behave veridically. These phenomena suggest that what we normally see depends on visual associations generated by experience. The necessity of such learning may explain why the mammalian visual system is subject to a prolonged period of plasticity in early life, when novel circuits are made in enormous numbers.

Information generated by the eyes is ambiguous. Everyday we have to make decisions (about the size and distance of objects, their form, and whether they are moving) based on retinal images that can have two or more meanings (1–4). Indeed, because the complexities of a three-dimensional world are projected onto a two-dimensional receptor sheet, the interpretation of most retinal images is equivocal. The ability to resolve these uncertainties, a talent directly relevant to survival, shows that normally we have little trouble reaching valid conclusions about potentially confusing visual stimuli. But how do we accomplish this? In the 19th century, students of vision were divided on this issue, the two opposing camps being represented by Hering and Helmholtz (1, 5). Hering maintained that the innate analytic abilities of the visual system enabled such determinations to be made more or less *a priori* (the "nativist" position). Conversely, Helmholtz maintained that the correct interpretation of visual stimuli is generally a matter of inferences based on visual experience (the "empiricist" position).

We have reexamined this long-standing debate using visual stimuli generated by wire-frame cubes. Because the proximal and distal elements of such structures are not easily distinguished, an illusory object can be perceived that behaves quite differently from the solid objects we are accustomed to seeing. In addition to their intriguing—and often amusing—nature, these alternative percepts raise the question of whether visual perception is based on the operation of *a priori* rules for processing information supplied by the retina (6) or is better explained in terms of *a posteriori* associations acquired by experience with objects in the real world (1, 4, 7, 8).

Altered Form of a Wire-Frame. When a transparent cube is viewed monocularly, the two most common interpretations of the stimulus alternate, much as when one views a two-dimensional representation (the familiar "Necker cube"; ref. 9). For a transparent cube positioned as in Fig. 1*A*, one interpretation—the correct one in this example—is looking down on the top of the cube; the other is as if looking up at its bottom (Fig. 1*B*). If the transparent cube is seen in its top-down presentation, all six faces appear to be approximately equal in area. When, however, the same retinal image is perceived as if viewed from the bottom, the structure is seen as a truncated pyramid, the proximal faces appearing smaller than the distal faces (Fig. 2). Moreover, when seen in the illusory bottom-up

orientation, the cube appears to be balanced on its distal–inferior vertex, with the surface on which it actually rests rising from the balance point (see Figs. 1 and 2). (Illusory, in this case, means an interpretation of the stimulus that does not accord with the configuration of the object determined by direct measurement.) In short, the observer no longer judges the object to be a cube, despite the unchanged retinal image, knowledge of its actual structure, and the immediately preceding perception of a cube in top-down view.

A first order explanation of these phenomena follows from the geometry of the situation. Because of their greater distance, the angles subtended on the retina by the distal elements of the cube are less than the angles subtended by the proximal ones. When the object is perceived in its top-down (actual) presentation, the visual system "compensates" for this asymmetry of the retinal image such that the structure is seen as a cube (Fig. 2). This adjustment presumably occurs because the visual system associates retinal images that are routinely distorted by the geometry of size and distance with percepts that better represent the actual object. However, when the illusory (bottom-up) interpretation prevails, the usual relationship of the front and back faces of the cube is reversed, such that a different form of the same retinal image is perceived. This alternative perception occurs because the usual compensatory mechanism is now applied inappropriately.

Altered Motion Parallax. A second remarkable phenomenon is apparent if, while viewing a wire-frame cube, the head is moved from side to side. Normally, this strategy is used to ascertain the spatial relationships of objects by motion parallax (Fig. 3). As the head moves one way, objects in the foreground are perceived as shifting in the opposite direction with respect to the background, thus aiding judgments about depth (10) that also are informed by stereopsis, accommodation, vergence, and many other cues. As long as the observer perceives the transparent cube in its actual (top-down) orientation (see Fig. 1), motion parallax is generated by head movements (Fig. 3*A*). When, however, the same retinal image is seen in reversed perspective, motion parallax fails: the object no longer moves laterally in relation to the background but rotates in the direction of the head movement (Fig. 3*B*).

A first order explanation again follows from the geometry of the situation. When the head is moved laterally, the proximal elements of the cube move a greater distance on the retina than the distal elements. Important to note, these changes of the retinal image generated by head movement are the same as those that occur when the cube rotates (see Fig. 3). The visual system normally appreciates that, when the head is moved to assess spatial relationships, the foreground objects are not in fact rotating but are most usefully perceived as shifting laterally with respect to the background. When, however, the observer sees the transparent cube in reversed perspective, the object elements perceived to be nearer (i.e., the distal elements of the cube) move less than the proximal elements; in this case, the visual system interprets the cube to be rotating in the

*To whom reprint requests should be addressed. e-mail: purves@ neuro.duke.edu.
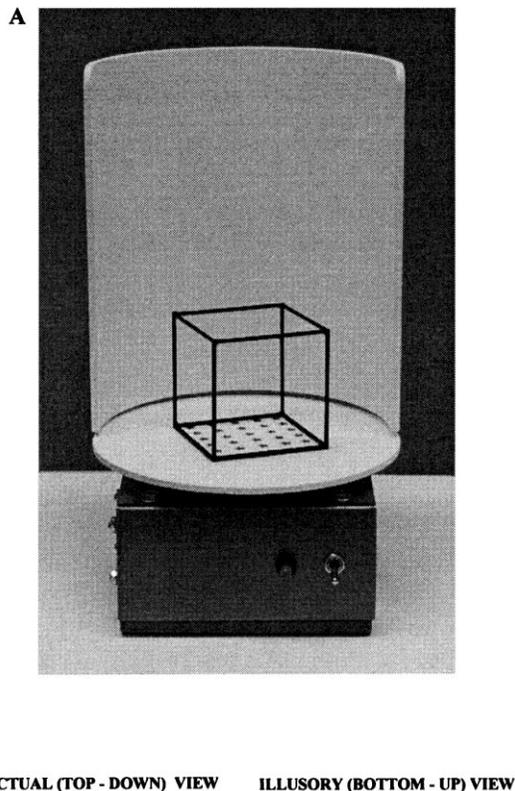
FIG. 1. Example of a transparent cube and the two most common interpretations of its arrangement in space. (*A*) Photograph of the apparatus used for these observations. A cubic frame with sides 5 cm in length was constructed from brass tubing painted black. The cube was cemented to a featureless white surface that could be rotated at variable speed by a small electric motor; a white cowl was attached to the base to provide a relatively featureless background. (*B*) Illustration of the two most common perceptions of a transparent cube oriented as in *A*. (*Left*) The actual (top-down) interpretation. (*Right*) The illusory (bottom-up) interpretation. The front and top, or the front and bottom, faces have been opacified to render the illustrations of the two interpretations unambiguous in two dimensions. Although the two perceptions illustrated here are invariably observed, a number of other interpretations can be briefly seen with prolonged observation. These include the perception of the transparent cube as a two-dimensional design and an interpretation in which the proximal superior and distal inferior vertices both appear to be coming toward the observer as the apices of overlapping pyramids.

direction of the head movement. The illusory perception occurs because this sequence of events signifies rotation when looking at objects that behave veridically. Thus, the visual system associates a particular sequence of changes in the retinal image with a particular perception (motion parallax), predicated on the behavior of conventional (solid) objects.

Altered Direction of Movement. If the surface on which a wire-frame cube rests is made to rotate (see Fig. 1*A*), a third phenomenon becomes apparent. Although the cube seen in its
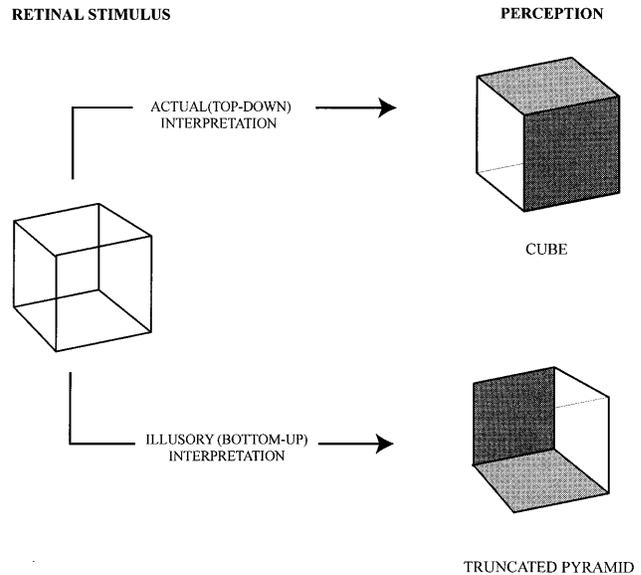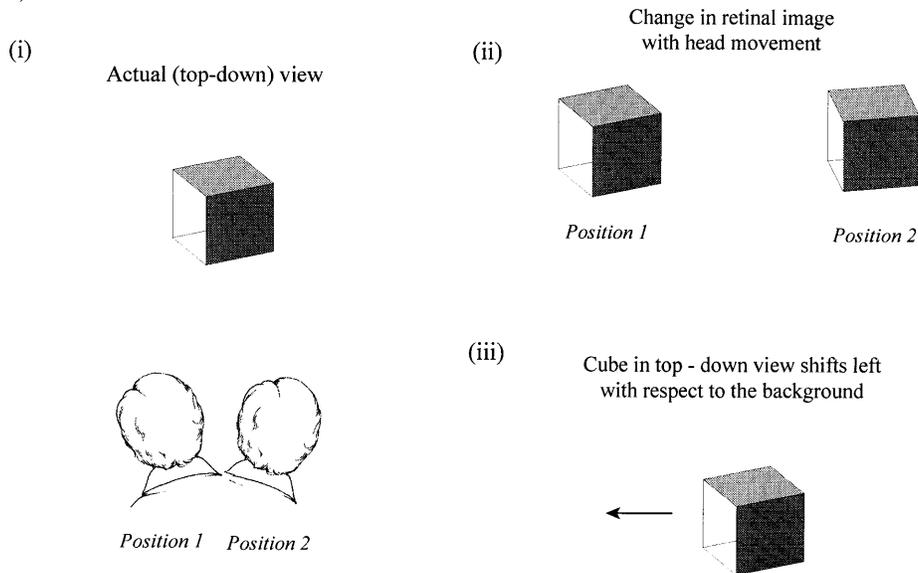

FIG. 2. Altered perception of the form of a transparent cube in response to different interpretations of the same retinal stimulus. When the stimulus is interpreted in its actual (top-down) orientation, the observer perceives a cube. Notice that the perceived cube is drawn without perspective because this is the two-dimensional representation that best depicts the subjective experience. When, however, the stimulus is interpreted in its illusory orientation, a truncated pyramid is seen. Two faces have been opacified to render the illustration unambiguous, as in Fig. 1. Although monocular viewing facilitates their appreciation, the phenomena described here can also be seen in binocular view and at any distance.

actual (top-down) presentation appears to turn normally, when the observer interprets the image to be in the illusory (bottom-up) orientation (see Fig. 2), the direction of rotation immediately reverses. As a result, the cube is perceived to be tumbling in the opposite direction above the rotating surface (Fig. 4) (see also refs. 11 and 12). Moreover, the elements of a pattern on the portion of the rotating surface bounded by the bottom frame of the cube (see Fig. 1*A*) also rotate in an opposite direction. Despite knowledge of the actual arrangement of the object and the law of gravity, this illusory behavior looks every bit as "real" as the veridical rotation of the cube and the surface on which it rests.

The geometrical explanation of these subjectively amazing percepts is again straightforward. When the cube rotates, some elements move to the left while others move to the right. If the elements that move to the left in the actual (top-down) presentation are interpreted as being closer to the observer, the cube rotates in a clockwise direction. If, on the other hand, the contours that move to the right are perceived as being closer to the observer, the cube rotates in a counterclockwise direction. The visual system apparently determines the direction of rotation based on associating the proximal parts of an object moving to the left and the distal parts moving to the right with clockwise rotation and vice versa (compare Fig. 4 *A* and *B*; see also ref. 13). As a result, whenever the interpretation of the cube's orientation changes, the perceived direction of rotation instantly reverses. The other aspects of the illusory behavior (i.e., tumbling above the surface, part of which is appropriated by one of the upright faces of the cube; the reversed rotation of the elements of a pattern bounded by cube) follow from earlier explanations.

Implications. Each of these several observations shows that the perception of a transparent object can be dramatically altered by the observer's interpretation of its arrangement in space. The ambiguous retinal stimulus generated by monocular viewing of a wire-frame cube is similar to that

Neurobiology: Purves and Andrews

Proc. Natl. Acad. Sci. USA 94 (1997)   6519

A) NORMAL MOTION PARALLAX

(i)

Actual (top-down) view

(ii) Change in retinal image with head movement

Position 1   *Position 2*

(iii) Cube in top - down view shifts left with respect to the background

Position 1   *Position 2*

B) ABROGATED MOTION PARALLAX

(i) Illusory (bottom-up) view

(ii) Change in retinal image with head movement

Position 1   *Position 2*

(iii) Cube in bottom - up view rotates counterclockwise

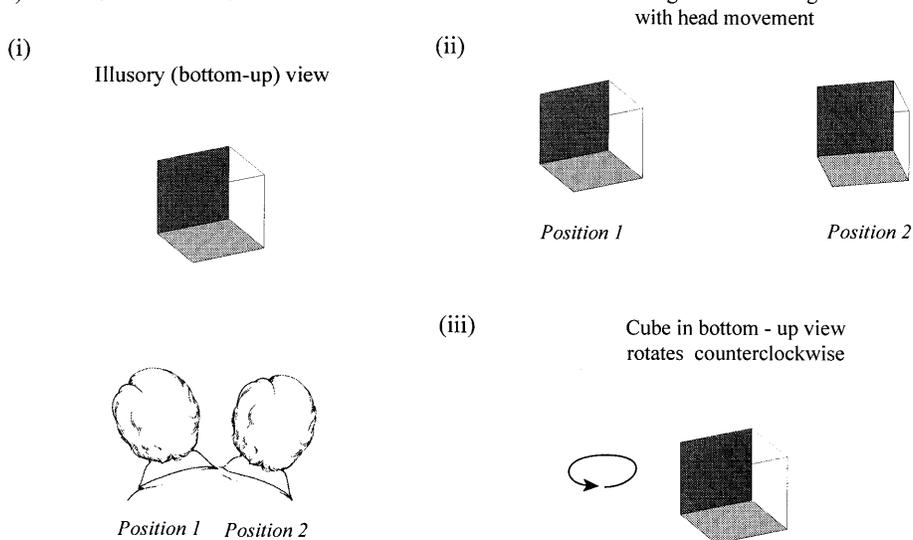*Position 1*   *Position 2*

FIG. 3.   Abrogated motion parallax. (*A*) Illustration of motion parallax. When the head is moved (*i*), an object in the foreground moves in the opposite direction with respect to the background (*ii*). If the foreground object happens to be a cube, the proximal faces move a greater distance to the left than the distal faces, as illustrated. (The conventions are the same as in Figs. 1 and 2). Although the sequence of retinal image changes is geometrically indistinguishable from object rotation, prior experience evidently has taught us to ignore this alternative possibility in favor of perceiving depth through motion parallax (*iii*). (*B*) When the retinal image generated by the transparent cube is interpreted to be in its illusory (bottom-up) orientation (*i*), the familiar response to head movement (i.e., motion parallax) no longer occurs. The proximal and distal faces of the transparent figure now change with respect to each other in a manner opposite that which occurs when the cube is perceived as it actually is (*ii*). In this circumstance, the cube is seen to be rotating in the direction of the head movement (*iii*).

elicited by the familiar two-dimensional Necker cube. As a result, perception shifts between two equally plausible interpretations of the retinal image that interchange the object's proximal and distal elements. The perceptions that arise from the two alternative interpretations of a two-dimensional Necker cube behave identically. In the case of the three-dimensional cube, however, conflating the front and back of the stimulus has dramatic consequences. The aspects of a wire frame cube that change depending upon whether the stimulus is seen in its veridical or illusory presentation include such basic properties as the form of the object, its spatial relationship with other objects, and its direction of movement (Table 1). Beyond the geometrical

explanations already offered, how can the same retinal image give rise to such different perceptual experiences?

As suggested by Hering (1, 5) and more recently by others (e.g., ref. 6), one explanation might be that visual perception is based on *a priori* analytic processes that operate on retinal information. When confronted with an ambiguous image, these processes would continue to operate but could generate more than one perceptual outcome in response to a particular stimulus. This strategy, however, would preclude the routine resolution of visual ambiguity. This point may be best appreciated by considering the resolution of semantic ambiguity. Take, for example, the sentence "The house is on the lake." Like the retinal image of a wire frame cube, the
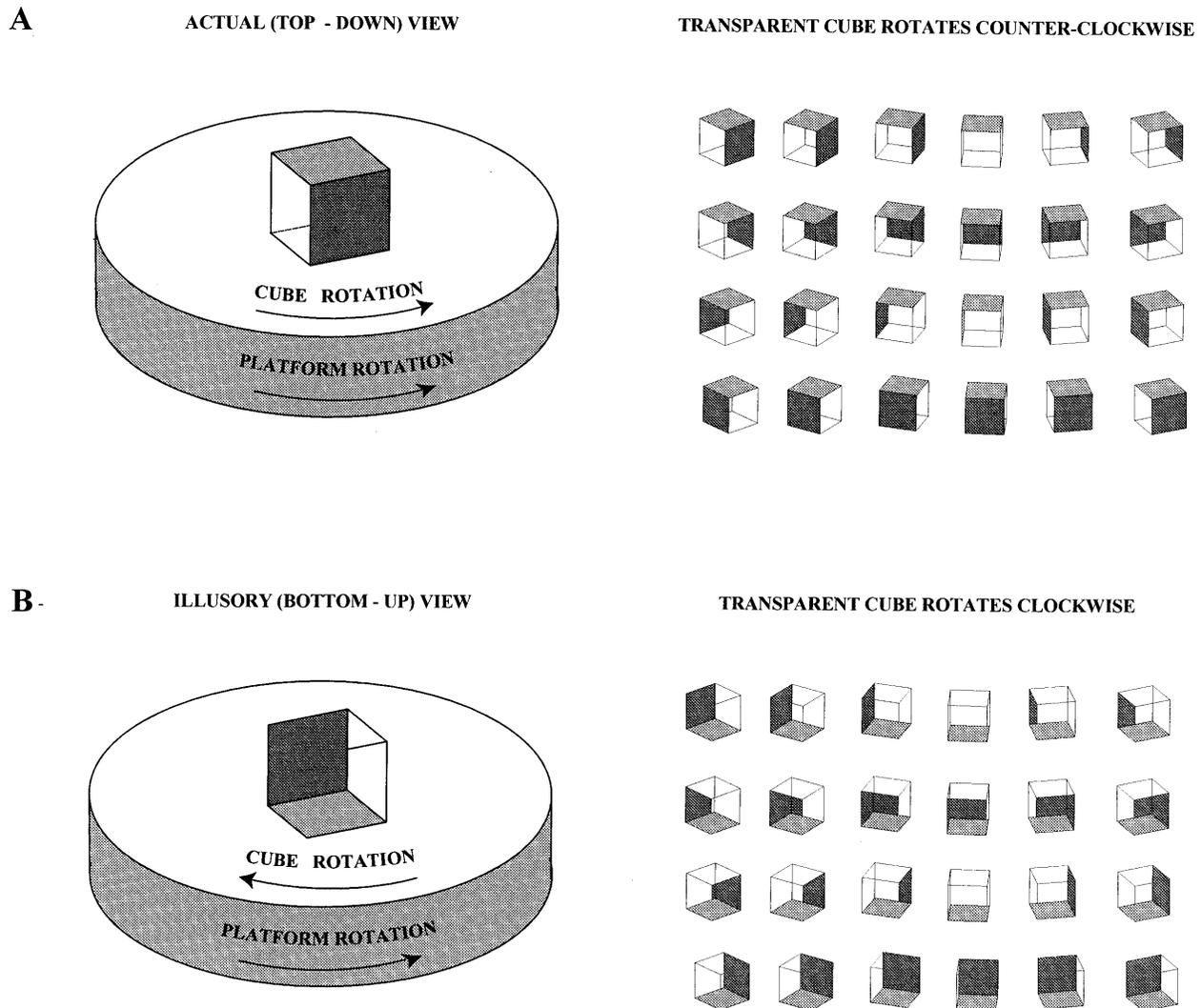
A   ACTUAL (TOP - DOWN) VIEW   TRANSPARENT CUBE ROTATES COUNTER-CLOCKWISE

B -   ILLUSORY (BOTTOM - UP) VIEW   TRANSPARENT CUBE ROTATES CLOCKWISE

FIG. 4.   Reversed direction of movement perceived as a result of the rotation of transparent objects. (*A*) When a solid cube or other three-dimensional object is placed on a rotating surface, it is of course perceived to be turning in the same direction as the surface. The same is true for the transparent cube seen in its actual (top-down) orientation (see Figs. 1 and 2). (*B*) When, however, the illusory (bottom-up) perception supervenes, the apparent direction of rotation is reversed. The explanation of this phenomenon is evident in the series of diagrams on the right, in which a transparent cube is shown in successive "frames," each rotated 15° from the earlier one (as in previous figures, two faces of the cube have been opacified to avoid ambiguity). Comparison of the sequences in (*A*) and (*B*) indicates why the motion of the proximal and distal faces is reversed in the two situations. This difference leads to the reversed, but completely realistic, perception of tumbling counter-rotation as long as the illusory interpretation holds sway. The rate of rotation makes no difference in the perceptions described, as long as the velocity is not so great as to cause blurring. This phenomenon can also be observed by simply rotating a transparent cube held in hand. When the cube is perceived in its illusory (bottom-up) form, turning the hand clockwise gives rise to the bizarre but quite compelling perception that the cube is rotating counterclockwise.

sense is ambiguous in this case because of the multiple meanings of the preposition "on" (in particular, the statement could mean the house is floating on the lake or is simply near its shore). No *a priori* rule can, in principle, determine which of the possible meanings is intended because that information is not contained in the statement. The ambiguity could be resolved arbitrarily by limiting the meaning of the preposition but only at considerable cost to the richness of language. In fact, semantic ambiguity is retained, the correct meaning being sorted out by additional knowledge about context, usage, etc. Likewise in the case of an ambiguous retinal image, the uncertainty is resolved by virtue of additional information. Although ancillary cues such as those provided by stereopsis, or feedback from vergence and/or accommodation may often indicate the correct meaning of a retinal image, they are of limited effectiveness in determining spatial relationships among objects that are more than a few meters away from the observer (1, 3). Indeed, our

observations make plain that such ancillary cues cannot resolve the ambiguities presented by a transparent cube; if they could, we would never see the illusory perceptions we describe.

The most plausible source of the additional information needed to resolve visual ambiguity is prior experience. Such experience could be derived from cues associated with other aspects of the scene, information from other sensory modalities (e.g., tactile experience with objects), from motor feedback, or even from associations established during phylogeny (14–16). Although the associational consequences of visual stimulation may be quite predictable, the "rules" in this conception are empirical. The visual system must accumulate by experience the associations elicited by an ambiguous retinal image and by the same token must eventually learn which set of associations best represents the actual object (i.e., the veridical percept). Because the generation of such associations is deeply ingrained in the nervous system, we are usually

Neurobiology: Purves and Andrews

Proc. Natl. Acad. Sci. USA 94 (1997)   6521

Table 1. Summary of observations

Retinal image	Interpretation	Behavior of object
Stationary; proximal and distal elements differ in size	actual	appears as cube
	illusory	appears as truncated pyramid
Changes as observer moves head from side to side	actual	cube appears to move in direction opposite to head movement against background (motion parallax)
	illusory	cube appears to rotate in the same direction as head movement
Changes as object rotates	actual	cube appears to rest on surface and rotates in same direction as support
	illusory	cube appears to tumble in opposite direction as support

unaware of the visual puzzles they routinely solve. The virtue of the transparent figures we have used here is to present a type of ambiguity with which we have had little or no experience, thus forcing the observer to be aware of a process that we normally take for granted.

The analogy between ambiguous visual and linguistic information is also helpful in thinking about the development of the ability to resolve uncertainties by prior experience. The basic circuitry for understanding and producing speech sounds is present very early (17, 18), presumably being "hard-wired," much as the circuitry that subserves classical visual receptive field properties (19–21). Whether in the context of language or vision, such circuitry provides the wherewithal to trigger the associations that indicate the correct meaning of a stimulus. The importance of experience in resolving ambiguity in the visual world is underscored by well documented clinical cases in which the proper interpretation of visual stimuli has to be learned again—often with great difficulty and limited success—when sight is restored in adults after blindness since childhood (22–25). (The etiology in such cases is typically bilateral destruction of the corneas by trauma or infection.) When vision is restored, these individuals invariably report difficulty understanding the visual world, despite the fact that they had been normally sighted in early life and that their postoperative visual acuity is reasonably good. Over weeks or months or in some cases longer, most of these patients learn to correctly interpret the meaning of various visual stimuli.

That experience has profound effects on the organization of the visual system in humans and other mammals is well known (26, 27). Moreover, a variety of evidence has shown that most brain circuitry is constructed postnatally (28–34) and is subject to the influence of neural activity (35–37). For example, in the developing rodent brain, different functional regions of cortex grow in proportion to their degree of metabolic and electrical activity (35, 36). Despite a wealth of information about neural development, the purpose of the large number of activity-dependent connections established postnatally has remained unclear. The observations we describe here suggest an answer to this puzzle. If the inherent uncertainty of many—perhaps most—retinal images can only be resolved by learning about actual objects, the influence of postnatal visual activity may serve primarily to establish the neuronal associations that enable appropriate interpretations of otherwise ambiguous information. In the context of receptive field properties alone, it is difficult to imagine why the visual system should remain plastic for a prolonged period in postnatal life, particularly because this malleability entails substantial jeopardy from the debilitating effects of visual deprivation (26, 27, 38). In the context of forming the neuronal associations needed to interpret inevitably ambiguous retinal images, such plasticity makes good sense.

1. von Helmholtz, H. L. F. (1924) *Helmholtz's Treatise on Physiological Optics*, transl. Southall, J. P. C. (George Banta Publishing, Menasha, WI), Vol. I-III.
2. Duncker, K. (1938) in *Source Book of Gestalt Psychology*, ed. Ellis, W. H. (Routledge, London), pp. 161–172.
3. Rock, I. (1984) *Perception* (Freeman, New York).
4. Gregory, R. L. (1990) *Eye and Brain: The Psychology of Seeing* (Princeton Univ. Press, Princeton), 4th Ed.
5. Turner, R. S. (1994) *In the Eye's Mind: Vision and the Helmholtz-Hering Controversy* (Princeton Univ. Press, Princeton).
6. Marr, D. (1982) *Vision* (Freeman, San Francisco).
7. Nakayama, K., Shimojo, S. (1992) *Science* **257,** 1357–1363.
8. Nakayama, K. & Shimojo, S. (1995) *Visual Cognition: An Invitation to Cognitive Science*, eds. Kosslyn, S. M., Osherson, D. N. (MIT Press, Cambridge, MA), 2nd Ed., pp. 1–70.
9. Necker, L. A. (1832) *Phil. Mag. J. Sci.* **1,** 329–337.
10. Rogers, B. & Graham, M. (1979) *Perception* **8,** 125–134.
11. Peterson, M. A. & Shyi, G. C.-W. (1988) *Percept. Psychophys.* **44,** 31–42.
12. Masin, S. C. (1993) *Foundations of Perceptual Theory* (North–Holland, Amsterdam).
13. Ittelson, A. & Ames, W. H., Jr. (1950) *J. Psychol.* **30,** 43–62 (1950).
14. Tinbergen, N. (1969) *Curious Naturalists* (Doubleday, Garden City, NY).
15. Alcock, J. (1993) *Animal Behavior: An Evolutionary Approach* (Sinauer, Sunderland, MA).
16. Fantz, R. L. (1963) *Science* **140,** 296–297.
17. Eimas, P. D., Siqueland, E. R., Juscyzk, P. & Vigorito, J. (1971) *Science* **171,** 303–306.
18. Miyawaki, M., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J. J. & Fujimura, O. (1975) *Percept. Psychophys.* **18,** 331–340.
19. Hubel, D. H. & Wiesel, T. N. (1963) *J. Neurophysiol.* **26,** 994–1002.
20. Hubel, D. H. & Wiesel, T. N. (1974) *J. Comp. Neurol.* **158,** 267–294.
21. Stryker, M. P. & Sherk, H. (1975) *Science* **190,** 904–906.
22. von Senden, M. (1960) *Space and Sight*, transl. Heath, P. (Methuen, New York).
23. Gregory, R. L. & Wallace, J. G. (1963) *Exp. Psych. Soc. Monograph.* **2,** 1–46.
24. Valvo, A. (1971) *Sight Restoration After Long-Term Blindness: The Problems and Behavior Patterns of Visual Rehabilitation*, eds. Clark, L. L. & Jastrzembska, Z. Z. (American Foundation for the Blind, New York), pp. 1–5.
25. Sacks, O. (1995) *An Anthropologist from Mars: Seven Paradoxical Tales* (Alfred A. Knopf, New York).
26. Wiesel, T. N. (1982) *Nature (London)* **299,** 583–591.
27. Hubel, D. H. (1988) *Eye, Brain, and Vision* (Freeman, New York).
28. Cragg, B. C. R. (1975) *J. Comp. Neurol.* **160,** 147–166.
29. Pomeroy, S. L., LaMantia, A.-S. & Purves, D. (1990) *J. Neurosci.* **10,** 1952–1966.
30. LaMantia, A.-S., Pomeroy, S. & Purves, D. (1992) *J. Neurosci.* **12,** 976–988.
31. Riddle, D., Richards, A., Zsuppan, F. & Purves, D. (1992) *J. Neurosci.* **12,** 3509–3524.
32. Bourgeois, J.-P. & Rakic, P. (1993) *J. Neurosci.* **13,** 2801–2820.
33. Purves, D., Riddle, D., White, L. & Gutierrez, G. (1994) *Curr. Opin. Neurobiol.* **4,** 120–123.
34. Purves, D., White, L., Zheng, D., Andrews, T. & Riddle, D. (1995) in *Individual Development Over the Lifespan: Biological and Psychosocial Perspectives*, ed. Magnusson, D. (Cambridge Univ. Press, Cambridge, UK), pp. 162–178.

35. Riddle, D. R., Gutierrez, G., Zheng, D., White, L., Richards, A. & Purves, D. (1993) *J. Neurosci.* **13,** 4193–4213.
36. Zheng, D. & Purves, D. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 1802–1806.
37. Purves, D. (1994) *Neural Activity and the Growth of the Brain* (Cambridge Univ. Press, Cambridge, UK).
38. Horton, J. C. (1992) in *Adler's Physiology of the Eye*, ed. Hart, W. M. (Mosby, St. Louis), pp. 728–772