13
Inferential versus Ecological Approaches to Perception

Dennis R. Proffitt

Perception is the most transparent of all human faculties. Perception is effortless. It just happens. Unlike perception, acts of thinking, remembering, speaking, and reasoning often require some effort and planning. Large individual differences in abilities are found for the other faculties, but not for perception. People become famous for being great thinkers, but there are no great perceivers in history. Because of the ease and automaticity of perception, its dazzling complexity is often overlooked. It is only when attempting to explain how perception happens that the incredible difficulty of the feat becomes apparent.

What is the problem? What is perception, and why is it difficult? These questions have typically been answered by representing the problem of perception as is depicted in figure 13.1. Here and throughout this chapter, only visual perception will be discussed; however, this problem representation generalizes to the other sense modalities as well.

In the world there is some object that is perceived. This physical object is called the *distal stimulus*, and to be seen it must be illuminated. Some of the light that strikes it is absorbed and some is reflected, and of the light that is reflected some gets into the observer's eyes. The projected image formed on the back of the eye consists of an array of light having at each point some intensity value and wavelength. This projected image is called the *proximal stimulus*. The proximal stimulus causes receptor cells in the eye to change their activity, and this, in turn, causes a change in the activity of the neurons to which they synapse. This activity flows back through visual tracks to various regions within the brain. As a consequence of all of this—somehow—perception occurs. The percept consists

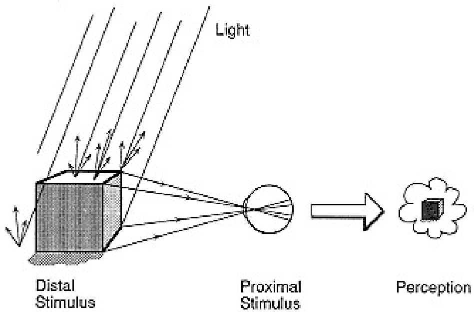

Figure 13.1
The traditional representation of the problem of perception. A physical object (distal stimulus) is illuminated, and some of this light is reflected into the eye causing an image of the object (proximal stimulus) to be formed on the retina. Photoreceptors respond to this proximal stimulus, thereby evoking perceptual processes that culminate in an awareness of the object (perception).

of an awareness of the object. Following Restle (1979), in figure 13.1 a cloud is drawn around the percept to indicate its mental status.

The conundrum inherent within this representation is that perceptions seem to bear a far closer resemblance to distal stimuli than to the proximal stimuli upon which they are based. For instance, three-dimensional objects project two-dimensional retinal images, and yet perceptions of objects are three-dimensional. Physical objects have constant properties, such as size, shape, and color (spectral reflectance), whereas proximal images have varying properties. The size of the image on the retina varies with distance, shape varies with object orientation, and color (reflected intensity and wavelength) varies with the intensity and spectral distribution of illumination. Proximal stimulation cannot be the sole informational basis for perception. Something must be added to sensory information to achieve the perceptions we form.

Any information that goes beyond that given in proximal stimulation must be brought to the occasion of perception by the perceiver. Most accounts of perception assert that the visual system *infers* the perceptual world on the basis of both sensory information and assumptions, biases, and knowledge inherent to the perceiver. The reason for postulating inferential processes is the insufficiency of proximal stimulation. There is more to perception than meets the eye.

There is, however, another point of view on the nature of perception that takes exception to everything just stated. Called the *ecological approach* by its creator, Gibson (1979), this theory claims that no inferences are required to account for perception because the effective information for perception is fully sufficient to specify what is perceived. To see how this can be, we need to return to the representation of the problem of perception.

What is the problem? For Gibson, figure. 13.1 completely misrepresents the problem of perception. For him, the purpose of perception is not to achieve a mental representation of distal objects. The purpose of perception is to control purposive actions. The information for perception is not the retinal image; rather, it is to be found in the flow of optical information that occurs at a moving point of observation. Gibson argued that perceptions can be based entirely on optical information if, and only if, the observer is allowed to move and explore the environment.

The purpose of this chapter is to describe the main characteristics of both traditional inferential theories of perception and Gibson's ecological approach. Following this discussion an attempt will be made to clarify why it is that both approaches continue to have their influence today, even though they seem quite incompatible. It will be argued that their incompatibility stems, in part, from the fact that proponents of each approach are asking quite different questions. The manner in which a problem is represented determines the form of its solution. It will also be argued that the differences between these positions are, at a metaphysical level, deep and irreconcilable.

Inferential Approaches to Perception

The development of the inferential approach to perception can be traced to Hermann von Helmholtz. Although many early proposals can be found in philosophy, Helmholtz's influence on contemporary theorizing is clear and direct. Helmholtz (1867/1925) coined the term *unconscious* *inference* to describe the processes by which the perceptual system uses inductive inference to derive perceptual interpretations from incomplete sensory information.

As one of the greatest and most versatile scientists to have ever lived, Helmholtz made important and lasting contributions to such diverse fields as medicine, anatomy, and physiology on the one hand, and physics and mathematics on the other. He also wrote prolifically on the philosophy of science, and in this domain he argued strongly for empiricism, believing that the source of all knowledge—both for the individual and for science—is rooted in empirical experimentation. Science and perception are both deemed to entail inferential processes but of quite different sorts. Science, he noted, is guided by deductive inferences, executed consciously and derived from carefully designed and controlled experiments. Perceptual inferences are unconscious inductions based upon incomplete and inexact experience. They are acquired through happenstance. Helmholtz (1894/1971, p. 505) wrote:

> How young children first acquire an acquaintance with or knowledge of the meaning of their visual images is easily understood if we observe them while they busy themselves with playthings. Notice how they handle them, consider them by the hour from all sides, turn them down or try to break them. This is repeated every day. There can be no doubt that this is the school in which the natural relations among the objects around us are learned, along with the understanding of perspective images and the use of the hands.

Helmholtz proposed that through experiences such as these, the child comes to internalize two things: (1) a knowledge of geometrical optics and (2) implicit assumptions about the nature of the world.

As an example of this sort of approach, consider how one perceives the shape of a book. As represented in figure 13.2, the projected image of a book is ambiguous. The laws of geometrical optics place constraints on what the distal object could be, but still perception is underspecified. An indefinite number of differently shaped objects could project the identical image. However, in experience, the most frequently encountered objects consistent with this projected image have been rectangular solids. Rectangularity is a pervasive constraint in the artifacts that people create, and thus it would be reasonable for the perceptual system to acquire a bias to construe visual images as being rectangular objects whenever such

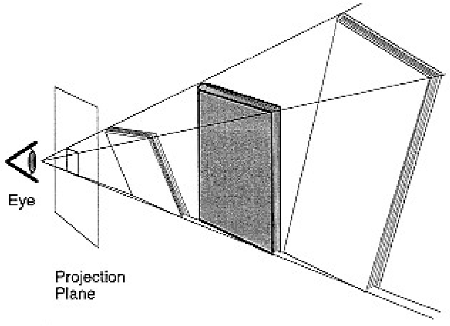

Figure 13.2
An illuminated book projects an image into the eye, and for expository purposes, onto a projection plane in front of the eye. In addition to the book, there are an indefinite number of objects of different shapes and sizes that could project an identical image. Two of these are depicted.

perceptions are possible. This bias would reflect an implicit assumption that rectangular objects were the likely cause of visual images consistent with rectangularity.

Helmholtz's position has become the mainstream in perceptual theorizing (Proffitt & Kaiser, 1998). Within this approach, the problem of perception is grounded in the inherent ambiguity of optical information. Using the optical information available, an internalized geometry, and an appreciation for regularities that have been encountered in experience, the perceptual system is said to make unconscious inferences about the external world. Perceptions are the conclusions of these inferences. In the words of Gregory (1978), a contemporary proponent of this position, "The senses do not give us a picture of the world directly; rather they provide evidence for the checking of hypotheses about what lies before us. Indeed, we may say that the perception of an object is an hypothesis, suggested and tested by the sensory data" (p. 13). The work of Adelbert Ames, Jr. (Ittelson, 1968) stands as one of the best examples of perception viewed in this manner.

Ames had a magician's appreciation for the ambiguity of optical information. He knew that through careful contrivance he could evoke illusions of the most elaborate sort. Of these, the most famous is the distorted room demonstration. As depicted in figure 13.3, a full-size room was constructed such that with the exception of the front wall, all of the other walls, the ceiling, and the floor were nonrectangular shapes. In addition, the walls had windows that were similarly nonrectangular. The dimensions of the room and its windows were determined so that when viewed from the outside through a viewing hole in the room's front door, the optical information was consistent with a rectangular room. And, indeed, a rectangular room is what people saw. Ames's purpose in constructing his demonstrations was to convince people that what they perceived was due not only to reality, but also to what they brought to the act of perception. In the case of the distorted room demonstration, he argued that people bring assumptions about how rooms ought to appear given their past experience. These assumptions include a strong proclivity to suppose that rooms and windows are rectangular.

Kubovy (1986) provided a somewhat different twist to the rectangularity bias as it is applied to the Ames distorted room demonstration. He noted that all of the surfaces in an Ames distorted room meet at edges that project fork or arrow junctions such as those depicted in figure 13.4. These junctions are consistent with Perkins's laws. Discovered independently by Perkins (1972, 1973) and by Shepard (1981), these laws state the conditions under which people will perceive rectangularity when confronted with fork and arrow junctions. The first of Perkins's two laws states that a fork junction will be perceived to be the projection of the vertex of a rectangular solid if and only if each of the three angles forming the Y configuration is greater than 90°. The second law states that an arrow junction will be perceived to be the vertex of a rectangular solid if and only if each of the small interior angles is less than 90° and together they sum to more than 90°. These laws will correctly detect rectangularity if the object is, in fact, rectangular and is not viewed too peripherally (Kubovy, 1986). These laws do not, however, assure physical rectangularity as is evidenced by the distorted room demonstration. Kubovy proposed that the reason an Ames distorted room is perceived to be rectangular is because the presented fork and arrow junctions are consis-

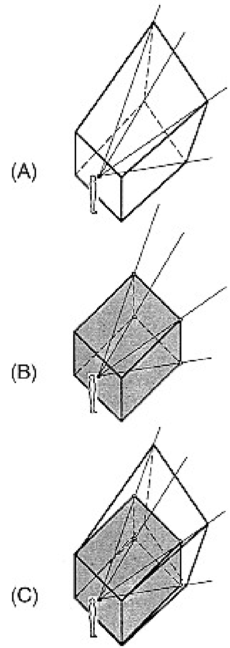

Figure 13.3
An Ames distorted room. Panel A shows a person viewing an Ames room through its viewing hole. Lines of sight are drawn to the room's vertices. Panel B depicts the rectangular room that is perceived. Panel C superimposes the perceived room onto the actual distorted room. Notice that the lines of sight for the perceived and actual room correspond.

tent with Perkins's laws. Kubovy also provided a drawing of an unfamiliar object like that in figure 13.5. Because the junctions in this depiction conform to Perkins's laws, it is perceived to be rectangular. Perkins's laws are a specific instance of how a strong rectangularity bias is operative in perception. It is generally assumed that this bias is based upon the prevalence of rectangularity in the artificial world, and thus, that it has been internalized through experience.

That Perkins's laws are deeply internalized is further supported by a remarkable set of studies by Enns and Rensink (1991). In a visual search

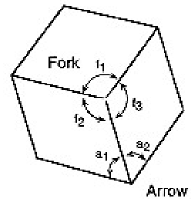

Figure 13.4
Fork and arrow junctions as described by Perkins's laws. For the Y-shaped fork junction, all three angles are greater than 90°. For the arrow junction, the two interior angles are each less than 90°, and together these two angles sum to greater than 90°.

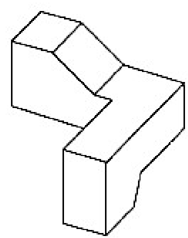

Figure 13.5
An unfamiliar object that appears rectangular in those regions that conform to Perkins's laws.

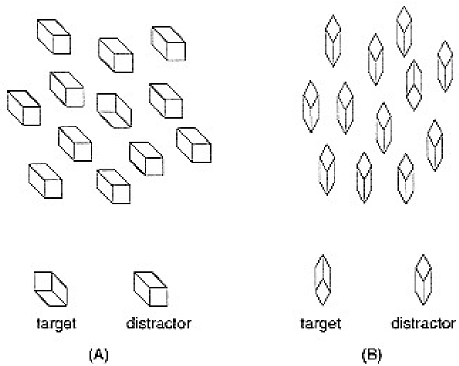

Figure 13.6
A and B presenting arrays of elements in which there is one target that differs from all of the other elements in the array. In (A), the target and distractors conform to Perkins's laws, and the target can be found preattentively. In (B), the target and distractors do not conform to Perkins's laws,and the target is much more difficult to find.

task, they presented figures containing either fork or arrow junctions. As depicted in figure 13.6, target and distractors were identical except that the target was oriented at a 180° angle relative to the array of distractors. When the junctions obeyed Perkins's laws (figure 13.6A), search occurred preattentively, meaning that the time to find the target was unaffected by the number of distractors present, whereas configurations that violated Perkins's laws (figure 13.6B) required slower, more effortful search.

The most comprehensive account of perception viewed as an inferential process is found in Rock's (1983) book, *The Logic of Perception*. In this book, Rock surveyed much of the literature on perception, and everywhere he looked he found evidence for reasoning, problem solving, inference, and knowledge-based assumptions in perception. Consider, for example, lightness perception.

A surface's reflectance value is perceived as its lightness. Dark surfaces absorb more light than do light ones. Given this fact, it might seem that the perceptual system would need only to register the amount of light coming from a surface to determine its lightness. Ignoring such additional complications as surface orientation, the problem is that the luminance emanating from a surface depends upon two things: the amount of light illuminating it and the reflectance value of the surface. If, however, one is presented with two equally illuminated surfaces, then relative lightness could be perceived on the basis of the invariant ratio of luminance emanating from the two surfaces (Wallach, 1948). For instance, if one surface reflects 80% of the light striking it and another surface reflects 40%, then the ratio of their reflected luminance will be 2:1 regardless of how much they are illuminated.

The complication that arises with this ratio account is that surfaces are not always illuminated equally. For example, part of a surface may have a shadow cast upon it. Gilchrist, Delman, and Jacobsen (1983) showed that in perceiving lightness, the visual system must first categorize the edges in the scene. Edges fall into two classes: those due to differences in illumination such as shadows and those due to differences in reflectance values. Rock argued that in perceiving lightness, the perceptual system must go through a multiple-stage process in which inferences are first made in edge classifications, followed by inferences that make use of these edge classifications in determining perceived lightness. When looking at a shadow, for example, the edges of the shadow are detected and an inference is made that it is, in fact, a shadow. Given this inference, the differences in luminance emanating from either side of the edge are attributed to differences in illumination rather than reflectance. As exemplified in this example, an important aspect of Rock's account is that perceptions play a causal role in subsequent perceptions. Perceiving the nature of edges is a necessary precondition for perceiving lightness.

Another example of such perceptual interdependencies is found in traditional accounts of size perception. Rock, like most other theorists, assumed that in order to perceive size, perceived distance must be taken into account. A well-known example of this notion is a demonstration first reported by Emmert (1881). He had observers form an afterimage by looking at a bright light, and then he instructed them to look at near and far surfaces. The afterimage appeared to be localized on whatever surface was being inspected, and thus it was perceived to be located at different observer-relative distances. The visual angle of the afterimage remained constant, whereas its apparent distance varied as the observer looked about. Emmert observed that the apparent size of the afterimage was larger when looking at far surfaces as opposed to nearer ones. From this observation he formulated what has become know as *Emmert's Law*, which states that the perceived size of an afterimage is proportional to its apparent distance. In general, many theorists, as did Rock, assumed that perceived size depends upon perceived distance (c.f. Gogel, 1990, 1993). Notice that the perceptual rules that relate perceived size and distance apply to psychological variables, not to physical ones.

Rather than appealing to perceived variables, there is a greater tendency today to analyze a scene into optical variables that can be geometrically related to distal properties of the scene. This is especially true within the growing field of computational vision, where there is an effort to extract properties of physical objects from their projected images. Computational accounts also differ from Rock's in another way. Instead of postulating that the perceptual system follows rules of inference, computational models instantiate inferential rules without necessarily following them (Epstein, 1995). That is, these models perform as if they were making inferences, even though their algorithms do not embody the inferences themselves. Consider an example. Ullman (1979, 1983) showed how the three-dimensional structure of a rotating object could be derived from its transforming two-dimensional projection so long as the object was assumed to be rigid. The algorithm that he developed does not have within it any reference to the rigidity assumption; rather, the algorithm produces correct interpretations of three-dimensional structure if and only if the images that are presented to it are projections of rigid rotations.

Poggio, Torre, and Koch (1985) looked at a number of problems in vision from a computational point of view and concluded that they were all *ill-posed problems*. They defined this distinction as follows: "A problem is well-posed when its solution exists, is unique and depends continuously on the initial data. Ill-posed problems fail to satisfy one or more of these criteria" (p. 315). Viewing perception as an ill-posed problem motivates a search for intelligent resources capable of making educated guesses in interpreting inputs. Poggio, et al. wrote, "The main idea for 'solving' ill-posed problems, that is for restoring 'well-posedness,' is to restrict the class of admissible solutions by introducing suitable *a priori* knowledge" (p 315). Ullman's (1979) use of a rigidity assumption in his account of extracting three-dimensional structure from motion is a good case in point. If objects can deform as they rotate, then extracting structure from motion is an ill-posed problem. Ullman's rigidity assumption restores well-posedness to the problem. In so doing, his algorithm will yield accurate descriptions of an object's configuration so long as the object is not deforming as it rotates. If the object is deforming, then the algorithm will provide an inaccurate description of its form. This is the hallmark of educated guesses: they are correct with a statistical probability no greater than the likelihood that their assumptions are correct.

Another example of an a priori constraint is Nakayama and Shimojo's (1992) principle of generic sampling. This principle states that the perceptual system assumes that a given object is not being viewed from an accidental vantage point. For example, when looking at a drawing of a square, one perceives it to be a two-dimensional configuration and not one end of a three-dimensional box viewed from a unique vantage point normal to its surface. Almost all perspectives on a box will show more than one of its sides. Only when viewed from a small number of accidental points of view would a box project as a square. The principle of generic views states that the perceptual system assumes that its current vantage point is not an accidental one. Rock (1983) proposed a coincidence explanation principle that attributed to the perceptual system the same sort of assumption.

All of these a priori constraints are consistent with the notion that the perceptual system possesses internalized knowledge about environmental regularities that are usually true. Marr (1982) called such internalized regularities *natural constraints*, meaning that they derive from a knowledge about what sorts of conditions are most likely to occur in the world. This is precisely what Helmholtz had in mind. Through experience—either of the individual or the species—the perceptual system comes to be imbued with knowledge about what is most likely to be present in the world given the evidence extracted from optical information. Perception is an educated guess. It is usually correct, but it is fallible. Illusions such as those seen in the Ames demonstrations are a symptom of its fallibility.

The Ecological Approach to Perception

For Gibson, the positions described in the foregoing discussion are simply muddled, their problem being that they began with a flawed representation of the problem of perception. If perception is represented in the manner depicted in figure 13.1, then it is, indeed, an ill-posed problem. Indeed, the problem is so ill-posed, Gibson argued, that no amount of inference and a priori knowledge will allow a successful restoration of well-posedness. The perceptual system cannot acquire an appreciation for the regularities that uphold in the world without the ability to have perceptions that seemingly require an internalization of these regularities to begin with. Gibson (1979. p. 253) wrote:

> Knowledge of the world must come from somewhere; the debate is over whether it comes from stored knowledge, from innate knowledge, or from reason. But all three doctrines beg the question. Knowledge of the world cannot be explained by supposing that knowledge of the world already exists. All forms of cognitive processing imply cognition so as to account for cognition.

Gibson's solution to this paradox was to propose that the information available in optical information is fully sufficient to support perception. There is no need for inference and a priori knowledge because nothing needs to be added to what is given in visual information. Some background into the position is required in order to see how this argument can be made.

Gibson saw his position as having developed from two antecedents. The first was Gestalt psychology, from which he acquired an appreciation for the systems approach to perception and the role of relational variables in specifying perceptual constancies. In regard to the systems approach, the Gestalt psychologists believed that perceptions were irreducible wholes. From a systems perspective, the laws that govern a system cannot be determined from an analysis of its constituent parts. In Gibson's ecological approach, the organism and environment comprise an irreducible system. Relational variables are mathematical relationships that can be extracted from visual information and that specify some persistent property in the environment. Wallach's (1948) aforementioned luminance ratio for specifying surface lightness over changes in illumination is such a relational variable. In Gibson's use, these variables came to be called *higher-order units* of perception.

The second influence was American functionalism as it developed from William James. Like Gestalt psychology, James's functionalism took a systems approach. Perception depended not only upon immediately given information, but also on the context of space and time in which it occurs. Time is clearly of importance for James, with experience being described as a stream of consciousness. Both persistence and change over time are essential properties of experience. Of particular relevance to Gibson's approach is the functionalist's program of understanding biological processes in terms of utility. From this perspective, the purpose of perception is to control actions. The veracity of perceptions is to be evaluated on the basis of whether they lead to appropriate actions, not on whether they correspond to reality objectively defined. The pragmatic definition of truth—truth is what works—makes sense only when embedded within a systems analysis that includes both the organism and the environment.

A clear statement of the functional theory of truth is found in Will's (1978, p. C7) commentary on baseball's Hall of Fame. The hall contains, "a plaque honoring the one American whose achievements of mind rank with those of Aristotle, Newton, Hegel and Einstein." This individual is Alexander Cartwright, and he is credited with setting the distances between bases at 90 feet. Will quoted the sports journalist, Red Smith:

> Ninety feet between bases represents man's closest approach to absolute truth. The world's fastest man cannot run to first base ahead of a sharply hit ball that is cleanly handled by an infielder; he will get there only half a step too late. Let the fielder juggle the ball for one moment or delay his throw an instant and the runner will be safe. Ninety feet demands perfection. It accurately measures the cunning, speed, and finesse of the base stealer against the velocity of a thrown ball. It dictates the placement of infielders. That single dimension makes baseball a fine art—and nobody knows for sure how it came to be. (p. C7)

Setting the bases at 90 feet defines a relationship between the relevant surface layout of the baseball field and the behavioral potential and purposes of the ballplayers.

Applying functionalist notions to the content and meaning of perception, Gibson coined the term *affordance* to refer to the functional utility perceived in the visual world. More will be said about affordances, but first his account will be described, beginning with the environment to be perceived.

The environment can be described at many levels depending upon size scale (atomic to light years), time scale (instants to millennia), and purpose. It is the latter constraint that is easy to forget. All descriptions take their form relative to some purpose. For example, a description of a baseball field will take quite different forms depending upon whether the intent is to convey information of geological or baseball-playing relevance. Gibson argued that traditional theories of perception describe the environment in physical or geometrical terms as opposed to ecologically relevant ones.

The environment, as perceived by an organism, is a habitat. A habitat cannot be described without accounting for its relationship to the organism for which it is home. A given environment can be a habitat for a host of different species, and what makes it a habitat differs somewhat for each. An ecological description of the environment implies the mutuality between the organism's way of life and those aspects of the environment that afford these behaviors.

From this perspective, the perceptual environment consists of three things: a medium, substances, and surfaces. The medium of earth is air. Light passes through air, and locomotion is possible through it. Substances are substantial matter through which locomotion is not possible. Media and substances interface at surfaces, and for perception this is where all the action is. Perception informs the organism about surface layout. It is surface layout, not abstract space or geometry, that is perceived. In perceiving the layout of surfaces, the organism perceives the medium and substances that define them.

In order to perceive surface layout, two things must occur. First, the environment must be illuminated, and second, the organism must be allowed to move and explore it. Illumination begins with essentially unstructured parallel rays of light emanating from the sun. This light is scattered somewhat by the earth's atmosphere, but until it strikes surfaces, it contains no structure or information. Upon contact with surfaces, some of the light is absorbed and some is reflected. Reflected light is structured by surfaces, and thus it contains information about them. The problem for the perceptual system is to pick up this information.

In order to perceive the information that is in light, the organism must move and explore the environment. Gibson was in complete agreement with those who argued that the information available in a retinal image is insufficient to support perception. Gibson's response to this insufficiency was not, however, to postulate inferential processes inherent in the perceiver, but rather to argue that the retinal image is not the informational basis for perception. He proposed that optical flow—the change in optical structure that occurs at a moving point of observation—is the effective informational basis for perception.

Consider again the situation of viewing a book as depicted in figure 13.2. The optical information that is present at a single vantage point is ambiguous, as the figure shows; however, if the observer moves his or her head so as to obtain multiple perspectives on the book, then this ambiguity is eliminated. Figure 13.7 shows three images of a book. Each of these images is, by itself, ambiguous. Taken together, however, the three rotated perspectives of the book are sufficient to define the unique three-dimensional structure of its visible surfaces (Ullman, 1979, 1983). The information specifying the book's form consists of invariants extracted from its transforming image. Invariants are mathematical relationships that remain constant as other aspects of optical structure change. Returning to figure 13.7, notice that one of the book's corners has been colored in. In the leftmost image of the book, this corner projects an angle of over 90°; in the middle panel, this angle projects an angle of

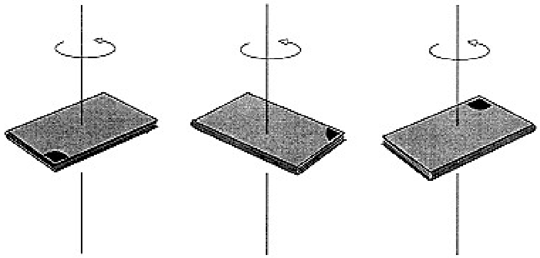

Figure 13.7
Three rotated images of a book. Notice how rotation causes a change in the projected angle for the highlighted vertex of the book.

less than 90°; and in the rightmost panel the angle is again larger than 90°. As the book is observed from different vantage points, this and all other angles change in their projected extent. It can be shown mathematically, however, that these changes in projected structure could be caused by only one rigid three-dimensional structure: a rectangular solid. The mathematics needed to prove this assertion are too complicated to explain here. See Todd (1995) for a review of the literature on perceiving structure from motion.

Consider a second example, of the Ames distorted room depicted in figure 13.3. Ames was able to construct illusory demonstrations such as this because he was able to restrict the viewers' vantage point to a single perspective. Observers looked at the distorted room through a peephole. If the door to the room were opened and observers were permitted to walk about and observe the changing optical structure the room provided, then they would see it accurately as a distorted room. The distorted room projects an image consistent with rectangularity only when viewed from a unique point of observation at the peephole. From every other vantage point, the room projects an optical structure that is totally inconsistent with rectangularity. Figure 13.8 shows this inconsistency from a second vantage point. As is the case generally, three different perspective images of a scene are sufficient to extract the unique three-dimensional structure of its visible surfaces.

Whenever an observer moves, every aspect of a projected scene changes. The changing optical structure that is projected to a moving point of observation is optical flow. Gibson suggested that optical flow contains within it two sorts of information. *Perspective structure* specifies what is changing, whereas *invariant structure* specifies the properties of the scene that remain constant over the change. Perspective structure informs the perceiver about his or her locomotion and changing position relative to the scene. Invariant structure specifies surface layout, including the size, shape, and slant of the surfaces that compose it. Because invariant structure specifies what is constant over change, the extraction of invariant structure requires change. Change is brought about through locomotion or the motions of objects themselves. The optical flow that results is informative about both the invariants of surface layout and the

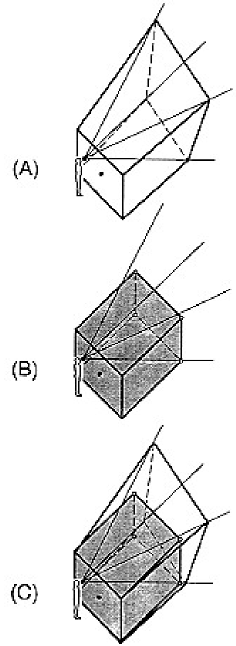

Figure 13.8
An Ames room viewed from a second vantage point. Unlike Figure 3, the lines of sight to the vertices of the perceived and actual room no longer correspond. The second vantage point does not support the perception of a rectangular room.

observer's changing position relative to it. *The perception of the environment and of the self come together*. One is impossible without the other.

Appreciating this mutuality between the perceiver and the environment is essential to understanding Gibson's position. The perception of the world cannot be separated from a perception of self. Consider the perception of size.

As was discussed in the previous section, traditional accounts of perception maintain that distance must be taken into account when perceiving size. Because visual angle varies inversely with distance, it is often supposed that size perception depends upon a prior perception of distance. Gibson stated that this was not so. Size, he argued, could be perceived directly without taking into account distance.

Figure 13.9 depicts an observer looking at an object that has a height, *h*, and that is some distance away, *d*. The altitude of the observer's eye is *i*. If the perceptual system can determine the position of the horizon, then the object's height can be determined as a fraction of the observer's eye height (Sedgwick, 1986). The horizon corresponds to the straight-ahead position in the visual field and is represented in the figure as a line parallel to the horizontal ground at an altitude equal to the observer's eye height. The visual angle from the bottom of the object to the horizon is , and that from the top of the object to the horizon is . The distance from the top of the object to the horizon is *j*. Referring to the figure, notice that

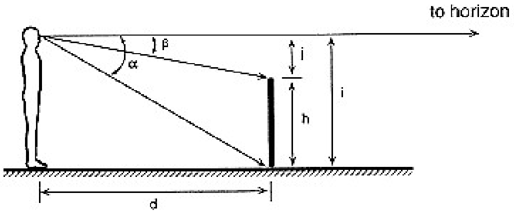

Figure 13.9
An observer looking at an object. The labeled dimensions of this situation are defined in the text.

$i/d = \tan$,

and thus,

$d = i/\tan$.

Similarly,

$j/d = \tan$,

and thus,

$j = d \tan$.

Since

$h = i - j$,

then substituting for j,

$h = i - d \tan$,

and substituting for d,

$h = i - i (\tan \quad / \tan \quad)$.

If the object is taller than eye height, then j is added to i, rather than being subtracted.

Critical to this formulation is determining the location of the horizon, but fortunately, its position is given quite robustly in optical information. All projected horizontal lines converge in depth to the horizon. Moreover, whenever the observer moves forward, there is a discontinuity in optical flow such that texture elements above and below the horizon move up or down, respectively. That is, as one moves forward, everything above eye height flows overhead and everything below travels beneath.

The important thing to notice about the final equation is that size can be determined entirely on the basis of optically given visual angles, and . Distance need not be taken into account. Again, perception of the environment implies a perception of self. Size is perceived relative to the size of one's body (i).

Another source of information about size is found in texture gradients, as depicted in figure 13.10. For a surface consisting of a relatively uniform texture, there is a compression of projected texture with distance. Even though the projected density of texture increases with distance, the amount of texture occluded at the base of an object is invariant over

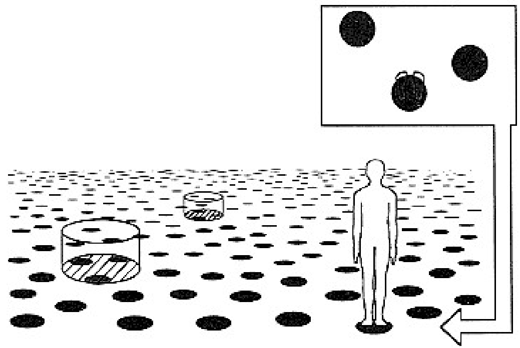

Figure 13.10
A texture gradient. Two cylinders of equal size occlude an equal amount of texture at their bases. The observer can scale the size of these objects to the size of his or her feet by looking down and noticing how much texture they occlude.

displacements on the surface. That is, objects of equal size occlude an equal amount of texture at their base. Texture gradients provide information about relative size but not absolute size unless there is an object of known size on the ground surface. Fortunately, all one needs to do is look down at one's feet in order to scale the texture to this familiar standard.

These are but a few examples of how Gibson's approach can be applied to the perception of environmental properties. In redefining the effective information for perception—from the retinal image to optical flow—Gibson found it to be not nearly so deficient as had been previously thought.

However, the informational basis for perception is not equivalent to the content of perception. We do not perceive information. Instead, we perceive the world and our relationship to it. The content of perception is the functional utility of the surfaces and objects that are encountered. These functional utilities relate the dimensions of objects and surfaces to the behavioral potential of the organism. Gibson coined the term *affordances* to describe these functional utilities.

An affordance specifies what an organism can do with the objects and surfaces that are encountered in the environment. Any given object possesses an indefinite number of affordances; those that are perceived at any moment depend upon the intent of the perceiver. Consider the book you are reading. Its surface layout affords being held. This affordance relates the size and shape of the book to that of your hands. Moreover, the book could also be thrown like a frisbee, used as a club with which to squash a fearsome bug, or placed under the front of a slide projector to raise its projection. The number of possible uses for a book is indefinite, and they do not depend upon its objective conceptual meaning. Gibson (1979, p. 129) wrote: "An affordance cuts across the dichotomy of subjective-objective and helps us to understand its inadequacy. It is equally a fact of the environment and a fact of behavior. It is both physical and psychical, yet neither. An affordance points both ways, to the environment and to the observer." Consider one final example, that being the perception of geographical slant.

The earth's surface is rarely flat, and departures from horizontal are perceived as geographical slant. The magnitude of a hill's slant can be determined from such optical information as texture gradients, motion parallax, and binocular disparity. Even though there is sufficient information to objectively derive slant, people grossly overestimate the inclination of hills in the world (Proffitt, Bhalla, Gossweiler, & Midgett, 1995). When, for example, people stand in front of a 5° hill, they will estimate its inclination to be about 20°, and a 30° hill will be judged to be over 50°. Be that as it may, such overestimations do not cause people to stumble as they commence to walk up or down a hill. Proffitt et al. found that overestimations are only evidenced in explicit judgments of hill slant and that a motoric index of perceived slant is far more accurate. The visual guidance of actions requires accuracy, whereas the conscious awareness of slant is modulated by a person's behavioral potential.

Perceived steepness provides information about the affordances of hills, about whether they can be ascended or descended and with what degree of difficulty. Summarizing the Proffitt et al. findings, a 10° hill is very difficult to ascend for a long distance, and consistent with this it looks very steep. People judge 10° hills to be about 30°. A grassy 30° hill is near the limit of what can be ascended and is too steep to descend due to biomechanical asymmetries in our ascending/descending walking potential. Consistent with this asymmetry, hills steeper than about 25° are judged to be steeper when viewed from the top than from the bottom. Finally, hills appear steeper when we are tired than when we are not. Perceived steepness is not invariant with respect to distal slant alone, but rather it preserves the relationship between locomotor effort and distal. slant. Thus, this basic dimension of surface layout—the earth's topography—is perceived as a relationship between the distal inclination of the ground and our behavioral potential.

Concluding Remarks

The differences between the inferential and ecological approaches to perception are profound. By one account, perception is an educated guess; by the other it is a direct pickup of information. In philosophical parlance, the inferential approach is a form of *idealism*, meaning that perceptions are ideas formed by the perceiving mind about the physical world. The ecological approach, on the other hand, is a variant of *realism*, in which perceptions are viewed as corresponding directly to what is in the world. At this level of evaluation, the differences between these two views are irreconcilable.

At another level, however, these approaches can be seen to complement one another. Gibson's approach asks the question, *What* is perception? This question is answered by asserting that perception is an ecological description of the environment based upon the information in optical flow. The inferential approach asks a different question: *How* is perception achieved? Clearly, the answer to these two questions must begin with an analysis of the available information and a search for algorithms that can constrain this information to allow for unique perceptual interpretations.

Answers to what-versus-how questions correspond to what Marr (1982) called computational as opposed to algorithmic theories. A computational theory addresses the question of what the goal of the computation is, that is, what the system is attempting to do and why. Gibson's answer seems the appropriate one: the goal of perception is to discover the affordances of the environment. The algorithmic theory attempts to provide a representation of the system's inputs and outputs and of the algorithms that map one into the other. Of Gibson's approach, Marr (1982, p. 29) wrote:

> In perception, perhaps the nearest anyone came to the level of computational theory was Gibson (1966). However, although some aspects of his thinking were on the right lines, he did not understand properly what information processing was, which led him to seriously underestimate the complexity of the information-processing problems involved in vision and the consequent subtlety that is necessary in approaching them.

At least to me, Marr's criticism does not seem fair. Given that Gibson's goal was to provide a theory about what is perceived, problems of information processing were not his concern. Information processing is the province of those seeking to understand how perception is achieved at the level of algorithmic theory.

Marr introduced his distinction between the computational and algorithmic level of analysis with an example of how one might understand a cash register. The computational theory of the cash register would entail a description of what the device does. For example, it needs to accumulate prices in a manner unaffected by the order in which items are presented to it. Moreover, sorting items into groups and paying for each group separately should not affect the total. Enumerating such constraints on the price-totaling process results in a definition of the group-theoretic constraints on addition. Note, however, that nothing has been asserted about how these constraints are actually operationalized in the interior workings of the cash register. The actual device might be a computer or an abacus, because either is capable of performing addition. Neither device possesses internalized knowledge about the group-theoretic constraints on addition, even though both have been constructed in such a way that they cannot violate these constraints.

Gibson's approach does not instruct one on how to build a perceiving machine. Attempts to simulate perception require the evocation of a priori constraints such as those discussed earlier. Recall Ullman's rigidity constraint as applied to an algorithm designed to recover three-dimensional structure from motion information: given that the distal object is assumed to be rigid, then three rotated views of the object are sufficient to derive the three-dimensional form of its visible surfaces. However, the rigidity assumption is no more internalized by Ullman's algorithm than is the group theory of addition internalized by the cash register. In both cases, information processing takes a form that yields veridical outcomes if and only if the assumptions apply. Present a multiplication problem to a cash register, and it will compute a wrong answer. Similarly, present a deforming object to Ullman's algorithm and it will fail.

Whether perception is viewed as entailing inferential processes or not depends upon what sort of problem one is inclined to study. If one is interested in specifying the logic by which optical information is transformed into representations of the world, then indeed, logic will be required. If one wishes to understand what is perceived, then logic becomes unnecessary since we do not perceive logic; we perceive the world. At this level of analysis, the approaches seem compatible; they simply address different questions.

At a metaphysical level, however, the approaches make profoundly different assertions about the nature of mind and of being. By the inferential approach, the perceptual system must guess about the nature of the external world. For Gibson, the nature of the world reveals itself directly in experience. In both cases, what is known reflects upon the knower. In the first case, the knower augments optical information with inferences, assumptions, and a priori knowledge. Within the ecological approach, the knower and the known form an irreducible whole. Knowledge of self and of the world must necessarily come together.

References

Emmert, E. (1881). Grossenverhaltnisse der Nachbilder. *Klinische Monatsblaetter fuer Augenheilkunde, 19*, 443–450. (Translation: Size proportions of after images. *Clinical Monthly Newsletter for Opthalmology*.).

Enns, J. T., & Rensink, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review, 98*, 335–351.

Epstein, W. (1995). The metatheoretical context. In W. Epstein & S. Rogers (Eds.), *Perception of space and motion: Handbook of perception and cognition* (2nd ed., pp. 1–22). San Diego, CA: Academic Press.

Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.

Gilchrist, A., Delman, S., & Jacobsen, A. (1983). The classification and integration of edges as critical to the perception of reflectance and illumination. *Perception & Psychophysics, 33*, 425–436.

Gogel, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception & Psychophysics, 49*, 105–123.

Gogel, W. C. (1993). The analysis of perceived space. In S. C. Masin (Ed.), *Foundations of perceptual theory* (pp. 113–182). Amsterdam: Elsevier.

Gregory, R. L. (1978). *Eye and brain: The psychology of seeing* (3rd ed.). New York: McGraw-Hill.

Helmholtz, H. von (1925). *Handbook of physiological optics* (Vol. 3, J. P. C. Southall, Trans.). New York: Dover. (Original work published 1867.)

Helmholtz, H. von (1971). The origin and correct interpretation of our sense impressions. In R. Kahl (Ed.), *Selected writings of Hermann von Helmholtz* (pp. 501–512). Middletown, CT: Wesleyan University Press. (Original work published 1894.).

Ittelson, W. H. (1968), *The Ames demonstrations in perception*. New York: Hafner.

Kubovy, M. (1986). *The psychology of perspective and renaissance art*. Cambridge, England: Cambridge University Press.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Nakayama, K., & Shimojo, S. (1992). Experiencing and perceiving visual surfaces. *Science, 257*, 1357–1363.

Perkins, D. N. (1972). Visual discrimination between rectangular and nonrectangular parallelepipeds. *Perception & Psychophysics, 12*, 396–400.

Perkins, D. N. (1973). Compensating for distortion in viewing pictures obliquely. *Perception & Psychophysics, 14*, 13–18.

Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature, 317*, 314–319.

Proffitt, D. R., Bhalla, M., Grossweiler, R., & Midgett, J. (1995). Perceiving geographical slant. *Psychonomic Bulletin & Review, 2*, 409–428.

Proffitt, D. R., & Kaiser, M. K. (1998). The internalization of perceptual processing constraints. In J. Hochberg (Ed.) *Perception and cognition at century's end: Handbook of perception and cognition, 2nd edition*. San Diego: Academic Press. (Will be published this year. Do not yet have page numbers.)

Restle, F. (1979). Coding theory of the perception of motion configurations. *Psychological Review, 86*, 1–24.

Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.

Sedgwick, H. A. (1986). Space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Vol. 1, Sensory processes and perception* (pp. 21.1–21.57). New York: Wiley.

Shepard, R. N. (1981). Psychophysical complementarity. In M. Kubovy and J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 279–341). Hillsdale, NJ: Erlbaum.

Todd, J. T. (1995). The visual perception of three-dimensional structure from motion. In W. Epstein & Rogers (Eds.), *Perception of space and motion: Handbook of perception and cognition* (2nd ed. pp. 201–226). San Diego, CA: Academic Press.

Ullman, S. (1970). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Ullman, S. (1983). Recent computational studies in the interpretation of structure from motion. In J. Beck & A. Rosenfeld (Eds.), *Human and machine vision* (pp. 459–480). New York: Academic Press.

Wallach, H. (1948). Brightness constancy and the nature of achromatic colors. *Journal of Experimental Psychology, 38*, 6–13.

Will, G. F. (1978, April 2). The joy of baseball. *Washington Post*, p. C7.