

# THE ECOLOGICAL APPROACH TO VISUAL PERCEPTION

Classic Edition

*James J. Gibson*

# TABLE OF CONTENTS

<i>Preface</i>	xi
<i>Introduction</i>	xiii
<i>Introduction to the Classic Edition</i>	xvii

## **PART I**

### **The Environment to be Perceived 1**

- |                                  |    |
|----------------------------------|----|
| 1 The Animal and the Environment | 3  |
| 2 Medium, Substances, Surfaces   | 12 |
| 3 The Meaningful Environment     | 28 |

## **PART II**

### **The Information for Visual Perception 39**

- |   |    |
|---|----|
| 4 The Relationship Between Stimulation and Stimulus Information | 41 |
| 5 The Ambient Optic Array                                       | 58 |
| 6 Events and the Information for Perceiving Events              | 85 |

**x** Contents

7	The Optical Information for Self-Perception	104
8	The Theory of Affordances	119
<b>PART III</b>		
<b>Visual Perception</b>		<b>137</b>
9	Experimental Evidence for Direct Perception: Persisting Layout	139
10	Experiments on the Perception of Motion in the World and Movement of the Self	162
11	The Discovery of the Occluding Edge and Its Implications for Perception	180
12	Looking with the Head and Eyes	193
13	Locomotion and Manipulation	213
14	The Theory of Information Pickup and Its Consequences	227
<b>PART IV</b>		
<b>Depiction</b>		<b>253</b>
15	Pictures and Visual Awareness	255
16	Motion Pictures and Visual Awareness	279
	<i>Conclusion</i>	290
	<i>Appendix 1: The Principal Terms Used in Ecological Optics</i>	294
	<i>Appendix 2: The Concept of Invariants in Ecological Optics</i>	297
	<i>Bibliography</i>	299
	<i>Index</i>	305

## **PART I**

# The Environment to be Perceived

# 1

## THE ANIMAL AND THE ENVIRONMENT

In this book, *environment* will refer to the surroundings of those organisms that perceive and behave, that is to say, animals. The environment of plants, organisms that lack sense organs and muscles, is not relevant in the study of perception and behavior. We shall treat the vegetation of the world as animals do, as if it were lumped together with the inorganic minerals of the world, with the physical, chemical, and geological environment. Plants in general are not animate; they do not move about, they do not behave, they lack a nervous system, and they do not have sensations. In these respects they are like the objects of physics, chemistry, and geology.

The world can be described at different levels, and one can choose which level to begin with. Biology begins with the division between the nonliving and the living. But psychology begins with the division between the inanimate and the animate, and this is where we choose to begin. The animals themselves can be divided in different ways. Zoology classifies them by heredity and anatomy, by phylum, class, order, genus, and species, but psychology can classify them by their way of life, as predatory or preyed upon, terrestrial or aquatic, crawling or walking, flying or nonflying, and arboreal or ground-living. We are more interested in ways of life than in heredity.

The environment consists of the *surroundings* of animals. Let us observe that in one sense the surroundings of a single animal are the same as the surroundings of all animals but that in another sense the surroundings of a single animal are different from those of any other animal. These two senses of the term can be troublesome and may cause confusion. The apparent contradiction can be resolved, but let us defer the problem until later. (The solution lies in the fact that animals are mobile.) For the present it is enough to note that the surroundings of *any* animal include other animals as well as the plants and the nonliving

#### 4 The Ecological Approach to Visual Perception

things. The former are just as much parts of its environment as the inanimate parts. For any animal needs to distinguish not only the substances and objects of its material environment but also the other animals and the differences between them. It cannot afford to confuse prey with predator, own-species with another species, or male with female.

### The Mutuality of Animal and Environment

The fact is worth remembering because it is often neglected that the words *animal* and *environment* make an inseparable pair. Each term implies the other. No animal could exist without an environment surrounding it. Equally, although not so obvious, an environment implies an animal (or at least an organism) to be surrounded. This means that the surface of the earth, millions of years ago before life developed on it, was not an environment, properly speaking. The earth was a physical reality, a part of the universe, and the subject matter of geology. It was a potential environment, prerequisite to the evolution of life on this planet. We might agree to call it a world, but it was not an environment.

The mutuality of animal and environment is not implied by physics and the physical sciences. The basic concepts of space, time, matter, and energy do not lead naturally to the organism–environment concept or to the concept of a species and its habitat. Instead, they seem to lead to the idea of an animal as an extremely complex object of the physical world. The animal is thought of as a highly organized *part* of the physical world but still a part and still an object. This way of thinking neglects the fact that the animal-object is surrounded in a special way, that an environment is ambient for a living object in a different way from the way that a set of objects is ambient for a physical object. The term *physical environment* is, therefore, apt to get us mixed up, and it will usually be avoided in this book.

Every animal is, in some degree at least, a perceiver and a behavior. It is sentient and animate, to use old-fashioned terms. It is a perceiver *of* the environment and a behavior *in* the environment. But this is not to say that it perceives the world of physics and behaves in the space and time of physics.

### The Difference Between the Animal Environment and the Physical World

The world of physics encompasses everything from atoms through terrestrial objects to galaxies. These things exist at different levels of size that go to almost unimaginable extremes. The physical world of atoms and their ultimate particles is measured at the level of millionths of a millimeter and less. The astronomical world of stars and galaxies is measured at the level of light-years and more. Neither of these extremes is an environment. The size-level at which the environment exists is the intermediate one that is measured in millimeters and

meters. The ordinary familiar things of the earth are of this size—actually a narrow band of sizes relative to the far extremes. The sizes of animals, similarly, are limited to the intermediate terrestrial scale. The size of the smallest animal is an appreciable fraction of a millimeter, and that of the largest is only a few meters.

The masses of animals, likewise, are measured within the range of milligrams to kilograms, not at the extremes of the scale, and for good physiological reasons. A cell must have a minimum of substances in order to permit biochemical reactions; living animals cannot exceed a maximum mass of cells if they are all to be nourished and if they are to be mobile. In short, the sizes and masses of things in the environment are comparable with those of the animals.

### ***Units of the Environment***

Physical reality has structure at all levels of metric size from atoms to galaxies. Within the intermediate band of terrestrial sizes, the environment of animals and men is itself structured at various levels of size. At the level of kilometers, the earth is shaped by mountains and hills. At the level of meters, it is formed by boulders and cliffs and canyons, and also by trees. It is still more finely structured at the level of millimeters by pebbles and crystals and particles of soil, and also by leaves and grass blades and plant cells. All these things are structural units of the terrestrial environment, what we loosely call the forms or shapes of our familiar world.

Now, with respect to these units, an essential point of theory must be emphasized. The smaller units are embedded in the larger units by what I will call *nesting*. For example, canyons are nested within mountains; trees are nested within canyons; leaves are nested within trees; and cells are nested within leaves. There are forms within forms both up and down the scale of size. Units are nested within larger units. Things are components of other things. They would constitute a hierarchy except that this hierarchy is not categorical but full of transitions and overlaps. Hence, for the terrestrial environment, there is no special proper unit in terms of which it can be analyzed once and for all. There are no atomic units of the world considered as an environment. Instead, there are subordinate and superordinate units. The unit you choose for describing the environment depends on the level of the environment you choose to describe.

The size-levels of the world emphasized by modern physics, the atomic and the cosmic, are inappropriate for the psychologist. We are concerned here with things at the ecological level, with the habitat of animals and men, because we all behave with respect to things we can look at and feel, or smell and taste, and events we can listen to. The sense organs of animals, the perceptual systems (Gibson, 1966*b*), are not capable of detecting atoms or galaxies. Within their limits, however, these perceptual systems are still capable of detecting a certain

## 6 The Ecological Approach to Visual Perception

range of things and events. One can see a mountain if it is far enough away and a grain of sand if it is close enough. That fact is sufficiently wonderful in itself to deserve study, and it is one of the facts that this book will try to explain.

The explanation of how we human observers, at least some of us, can *visualize* an atom or a galaxy even if we cannot *see* one will not be attempted at this stage of the inquiry. It is not so much a problem of perception as it is of thinking, and there will be more about this later. We must first consider how we can perceive the environment—how we apprehend the same things that our human ancestors did before they learned about atoms and galaxies. We are concerned with direct perception, not so much with the indirect perception got by using microscopes and telescopes or by photographs and pictures, and still less with the kind of apprehension got by speech and writing. These higher-order modes of apprehension will only be considered in Part IV of this book, at the end.

### ***Units of the Ground Surface***

The literal *basis* of the terrestrial environment is the ground, the underlying surface of support that tends to be on the average flat—that is to say, a plane—and also level, or perpendicular to gravity. And the ground itself is structured at various levels of metric size, these units being nested within one another. The fact to be noted now, since it is important for the theory of perspective in Part II, is that these units tend to be repeated over the whole surface of the earth. Grains of sand tend to be of the same size everywhere, and so do pebbles and rocks. Blades of grass are all more or less similar to one another, and so are clumps of grass and bushes. These natural units are not, of course, perfectly uniform like the man-made tiles of a pavement. Nevertheless, even if their repetition is not metrically regular, it is stochastically regular, that is to say, regular in a probabilistic way. In short, the component units of the ground do not get smaller as one goes north, for instance. They tend to be evenly spaced; and if they are scattered, they tend to be evenly scattered.

### **The Time Scale of the Environment: Events**

Another difference between the environment to be described and the world of physics is in the temporal scale of the process and events we choose to consider. The duration of processes at the level of the universe may be measured in millions of years, and the duration of processes at the level of the atom may be measured in millionths of a second. But the duration of processes in the environment is measured only in years and seconds. The various life spans of the animals themselves fall within this range. The changes that are perceived, those on which acts of behavior depend, are neither extremely slow nor extremely rapid. Human observers cannot perceive the erosion of a mountain, but they





**FIGURE 1.1** The structure of the terrestrial earth as seen from above.

In this aerial photograph only the large-scale features of the terrain are shown. (Photo by Grant Heilman)

can detect the fall of a rock. They can notice the displacement of a chair in a room but not the shift of an electron in an atom.

The same thing holds for frequencies as for durations. The very slow cycles of the world are imperceptible, and so are the very rapid cycles. But at the level of a mechanical clock, each motion of the pendulum can be seen and each click of the escapement can be heard. The rate of change, the transition, is within the limits of perceptibility.

In this book, emphasis will be placed on events, cycles, and changes at the terrestrial level of the physical world. The changes we shall study are those that occur in the environment. I shall talk about changes, events, and sequences of

events but not about time as such. The flow of abstract empty time, however useful this concept may be to the physicist, has no reality for an animal. We perceive not time but processes, changes, sequences, or so I shall assume. The human awareness of clock-time, socialized time, is another matter.

Just as physical reality has structure at all levels of metric size, so it has structure at all levels of metric duration. Terrestrial processes occur at the intermediate level of duration. They are the natural units of sequential structure. And once more it is important to realize that smaller units are nested within larger units. There are events within events, as there are forms within forms, up to the yearly shift of the path of the sun across the sky and down to the breaking of a twig. And hence there are no elementary units of temporal structure. You can describe the events of the environment at various levels.

The acts of animals themselves, like the events of the environment they perceive, can be described at various levels, as subordinate and superordinate acts. And the duration of animal acts is comparable to the duration of environmental events. There are no elementary atomic responses.

The natural units of the terrestrial environment and the natural units of terrestrial events should not be confused with the *metrical* units of space and time. The latter are arbitrary and conventional. The former are unitary in one sense of the term, and the latter are unitary in a quite different sense. A single whole is not the same as a standard of measurement.

### Permanence and Change of the Layout

Space and time will not often be referred to in this book, but a great deal will be said about permanence and change. Consider the shape of the terrestrial environment, or what may be called its *layout*. It will be assumed that the layout of the environment is both permanent in some respects and changing in some other respects. A living room, for example, is relatively permanent with respect to the layout of floor, walls, and ceiling, but every now and then the arrangement of the furniture in the room is changed. The shape of a growing child is relatively permanent for some features and changing for others. An observer can recognize the same room on different occasions while perceiving the change of arrangement, or the same child at different ages while noticing her growth. The permanence underlies the change.

Permanence is relative, of course; that is, it depends on whether you mean persistence over a day, a year, or a millennium. Almost nothing is forever permanent; nothing is either immutable or mutable. So it is better to speak of *persistence* under change. The “permanent objects” of the world, which are of so much concern to psychologists and philosophers, are actually only objects that persist for a very long time.

The abstract notion of invariance and variance in mathematics is related to what is meant by persistence and change in the environment. There are

variants and invariants in any transformation, constants and variables. Some properties are conserved and others not conserved. The same words are not used by all writers (for example, Piaget, 1969), but there is a common core of meaning in all such pairs of terms. The point to be noted is that for persistence and change, for invariant and variant, each term of the pair is reciprocal to the other.

### ***Persistence in the Environment***

The persistence of the geometrical layout of the environment depends in part on the kind of substance composing it and its rigidity or resistance to deformation. A solid substance is not readily changed in shape. A semisolid substance is more easily changed in shape. A liquid substance takes on whatever may be the shape of its solid container. The upper surface of a liquid substance tends to the ideal shape of a plane perpendicular to gravity, but this is easily disturbed, as when waves form. When we speak of the permanent layout of the environment, therefore, we refer mainly to the solid substances. The liquids of the world, the streams and oceans, are shaped by the solids, and as for the gaseous matter of the world, the air, it is not shaped at all. I will argue that the air is actually a *medium* for terrestrial animals.

When a solid substance with a constant shape melts, as a block of ice melts, we say that the object has ceased to exist. This way of speaking is ecological, not physical, for there is physical conservation of matter and mass despite the change from solid to liquid. The same would be true if a shaped object disintegrated, changing from solid to granular. The object does not persist, but the matter does. Ecology calls this a *nonpersistence*, a destruction of the object, whereas physics calls it a mere *change of state*. Both assertions are correct, but the former is more relevant to the behavior of animals and children. Physics has sometimes been taken to imply that when a liquid mass has evaporated and the substance has been wholly dispersed in the air, or when an object has been consumed by fire, *nothing* has really gone out of existence. But this is an error. Even if terrestrial matter cannot be annihilated, a resistant light-reflecting surface can, and this is what counts for perception.

Going out of existence, cessation or destruction, is a kind of environmental event and one that is extremely important to perceive. When something is burned up, or dissolved, or shattered, it *disappears*. But it disappears in special ways that have recently been investigated at Cornell (Gibson, 1968a). It does not disappear in the way that a thing does when it becomes hidden or goes around a corner. Instead, the form of the object may be optically dispersed or dissipated, in the manner of smoke. The visual basis of this kind of perception will be further considered in Part II on ecological optics.

The environment normally manifests some things that persist and some that do not, some features that are invariant and some that are variant. A wholly

invariant environment, unchanging in all parts and motionless, would be completely rigid and obviously would no longer be an environment. In fact, there would be neither animals nor plants. At the other extreme, an environment that was changing in all parts and was wholly variant, consisting only of swirling clouds of matter, would also not be an environment. In both extreme cases there would be space, time, matter, and energy, but there would be no habitat.

The fact of an environment that is mainly rigid but partly nonrigid, mainly motionless but partly movable, a world that is both changeless in many respects and changeable in others but is neither dead at one extreme nor chaotic at the other, is of great importance for our inquiry. This fact will become evident later when we talk about the geometry of the environment and its transformations.

### **ON PERSISTENCE AND CHANGE**

Our failure to understand the concurrence of persistence and change at the ecological level is probably connected with an old idea—the *atomic theory of persistence and change*, which asserts that what persists in the world are atoms and what changes in the world are the positions of atoms, or their arrangement. This is still an influential assumption in modern physics and chemistry, although it goes back to Democritus and the Greek thinkers who followed him. There will be more about the atomistic assumption in Chapter 6 on events and how they are perceived.

### ***Motion in the Environment***

The motions of things in the environment are of a different order from the motions of bodies in space. The fundamental laws of motion hold for celestial mechanics, but events on earth do not have the elegant simplicity of the motions of planets. Events on earth begin and end abruptly instead of being continuous. Pure velocity and acceleration, either linear or angular, are rarely observable except in machines. And there are very few ideal elastic bodies except for billiard balls. The terrestrial world is mostly made of surfaces, not of bodies in space. And these surfaces often flow or undergo stretching, squeezing, bending, and breaking in ways of enormous mechanical complexity.

So different, in fact, are environmental motions from those studied by Isaac Newton that it is best to think of them as changes of structure rather than changes of position of elementary bodies, changes of form rather than of point locations, or changes in the layout rather than motions in the usual meaning of the term.

## Summary

The environment of animals and men is what they perceive. The environment is not the same as the physical world, if one means by that the world described by physics.

The observer and his environment are complementary. So are the set of observers and their common environment.

The components and events of the environment fall into natural units. These units are nested. They should not be confused with the metric units of space and time.

The environment persists in some respects and changes in other respects. The most radical change is going out of existence or coming into existence.

# 4

## THE RELATIONSHIP BETWEEN STIMULATION AND STIMULUS INFORMATION

Having described the environment, I shall now describe the information available to observers for perceiving the environment. Only then will we be prepared to consider how they perceive, what the activity of perception consists of, and how they can control behavior in the environment.

For visual perception, the information is obviously in light. But the term *light* means different things in different sciences, and we shall have to sort out the different meanings to avoid confusion. Most of us are confused, including the scientists themselves. The science of light is called *optics*. But the science of vision is also called optics, and the textbooks are not at all clear about the difference. Let us try to distinguish light as physical energy, light as a stimulus for vision, and light as information for perception.

What I call *ecological optics* is concerned with the available information for perception and differs from physical optics, from geometrical optics, and also from physiological optics. Ecological optics cuts across the boundaries of these existing disciplines, borrowing from all but going beyond them.

Ecological optics rests on several distinctions that are not basic in physical optics: the distinction between luminous bodies and nonluminous bodies; the difference between light as radiation and light as illumination; and the difference between radiant light, propagating outward from a source, and ambient light, coming to a point in a medium where an eye might be stationed. Since these differences are fundamental, they should be stated at the beginning. Why they are so important will become clear.

## The Distinction Between Luminous and Illuminated Bodies

Some material bodies emit light, and others do not. Light comes from sources such as the sun in the sky and from other sources close at hand such as fires or lamps on the earth. They “give” light, as we say, whereas ordinary objects do not. Nonluminous objects only reflect some part of the light that falls on them from a source. And yet we can see the nonluminous bodies along with the luminous ones. In fact, most of the things that need to be seen are nonluminous; they are only seen “by the light of” the source. The question is, *how* are they seen? For they do not *stimulate* the eye with light in the same way that luminous bodies do. The intermediate case of luminescent bodies is exceptional.

A terrestrial surface that gives light is usually, although not always, distinguishable from one that does not; it is visibly luminous, as distinct from being visibly illuminated. In physical optics, the case of reflected light is reduced to the re-emission of light by the atoms of the reflecting surface. But in ecological optics, the difference between a luminous and an illuminated surface is crucial. Where a reflecting surface in physical optics is treated as if it were a dense set of tiny luminous bodies, in ecological optics a reflecting surface is treated as if it were a true surface having a texture. There will be more of this later.

## The Distinction Between Radiation and Illumination

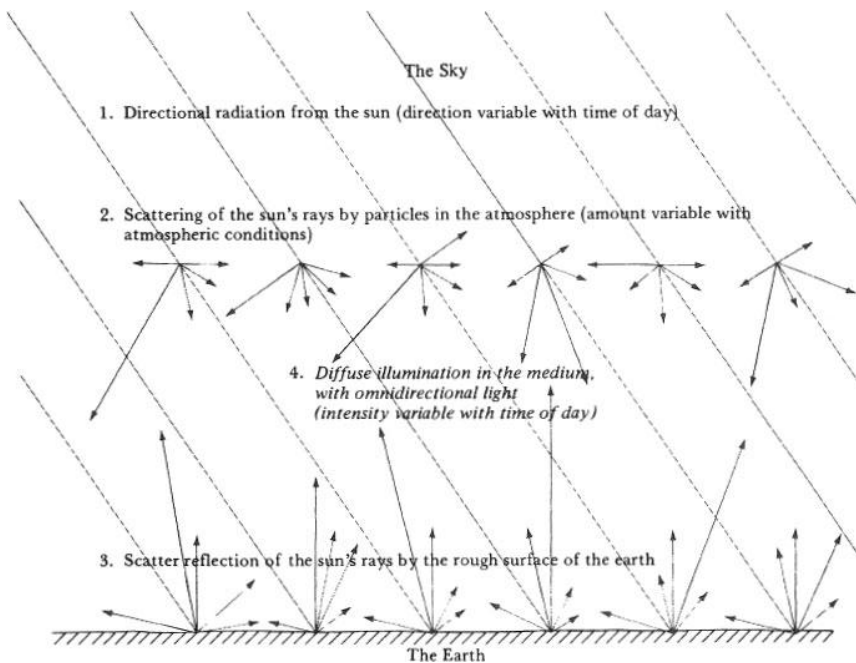
Radiant energy as studied in physics is propagated through empty space at enormous velocity. Such energy can be treated either as particles or as waves (and this is a great puzzle, even to physicists), but it travels in straight lines, or rays. The paths of photons are straight lines, and the perpendiculars to the wave fronts are straight lines. Moreover, light comes from atoms and returns to atoms. They give off and take in energy in quantal units. Matter and energy interact. There are elegant laws of this radiation, both at the size-level of atoms and on the grand scale of the universe. But at the ecological level of substances, surfaces, and the medium, we need be concerned only with some of these laws, chiefly scattering, reflection, and absorption.

### WHY ECOLOGICAL OPTICS?

The term *ecological optics* first appeared in print in an article with that title in *Vision Research* (Gibson, 1961). It seemed to me that the study of light, over the centuries, had not produced a coherent discipline. The science of radiant energy in physics, the science of optical instruments, and the science of the eye were quite different. The textbooks and journals of optics gave the impression of monolithic authority, but there were deep contradictions between the assumptions of the various branches of optics.

When I discovered that even an occasional physicist recognized these cracks in the foundations of the optical establishment (Ronchi, 1957), I ventured to suggest that optics at the level appropriate for perception should have a new name.

In daylight, part of the radiant light of the sun reaches the earth in parallel rays, but another part is scattered by being transmitted through an atmosphere that is never perfectly transparent. This light is even more thoroughly scattered when it strikes the textured ground, by what can be called *scatter reflection*. (This is not to be confused with *mirror reflection*, which is governed by the simple law of equal angles of the incident ray and the reflected ray. Mirror reflection seldom happens, for there are no mirrors on the ground, and even water surfaces, which could act as mirrors, are usually rippled.) The scatter-reflected light is in turn reflected back from the sky. Each new reflection further disperses



**FIGURE 4.1** The steady state of reverberating light in an illuminated medium under the sky.

Although at any point in the air the illumination comes from all directions, the *prevailing* illumination is from the left in this diagram because the direct radiation from the sun comes from the left.



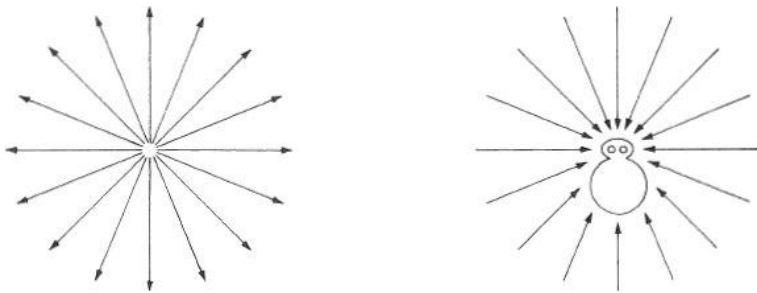
the incident rays. The light thus finds its way into shelters that are not open to the sun, or even to the sky. In semienclosed spaces the light continues to bounce back and forth at 186,000 miles per second. It finds its way through chinks and crevices and into caverns, until the energy is finally absorbed. This light can hardly be thought of as radiation now; it is illumination.

Illumination is a fact of higher order than radiation. In physical optics, experimenters try to avoid what they call stray light in the dark room. But in ecological optics, this light that has gone astray is just what interests us. The optician works with rays of light, rays that diverge in all directions from their source and never converge to a point unless they are focused by a lens. But an organism has to work with light that converges from all directions and, moreover, has different intensities in different directions.

Many-times reflected light in a medium has a number of consequences that, although important for vision, have not been recognized by students of optics. Chief among them is the fact of ambient light, that is, light that surrounds a point, any point, in the space where an observer could be stationed.

### The Distinction Between Radiant Light and Ambient Light

Radiation becomes illumination by *reverberating* between the earth and the sky and between surfaces that face one another. But that term, referring as it does to sound, does not do justice to the unimaginable quickness of the flux or to the uncountable multiplicity of the reflections back and forth or to their unlimited scattering. If the illumination is conceived as a manifold of rays, one can imagine every point on every surface of any environment as radiating rays outward from that point, as physicists do. Every such radiating pencil is completely "dense." One could think of the rays as completely filling the air and think of each point in the air as a point of intersection of rays coming from



**FIGURE 4.2** Radiant light from a point source and ambient light to a point in the medium.

A creature with eyes is shown at the point in the air, but it need not be occupied.

all directions. It would follow that light is ambient at every point. Light would come to every point; it would surround every point; it would be environing at every point. This is one way of conceiving ambient light.

Such an omnidirectional flux of light could not exist in empty space but only in an environment of reflecting surfaces. In any ordinary terrestrial space, the illumination reaches an equilibrium, that is, it achieves what is called a *steady state*. The input of energy from the sun is just balanced by the absorption of energy at the surfaces. With any change in the source, a new steady state is immediately reached, as when the sun goes down or is hidden by a cloud. No matter how abrupt the rise or fall of intensity of the light coming from a lamp, the rise or fall of illumination in the room is just as abrupt. The system is said to be open rather than closed inasmuch as addition of energy to the airspace and subtraction of energy from it are going on all the time, but the *structure* of the reverberation remains the same and does not change. What could this structure be? It is possible to conceive a nested set of solid angles at each point in the medium, as distinguished from a dense set of intersecting lines. The set of solid angles would be the same whatever the intensity of illumination might be (there will be more about this later). They are angles of intercept, based on the environment. The flow of energy is relevant to the stimulation of a retina, but the set of solid angles considered as projections is more relevant to stimulus information.

Consider the differences between radiant light and ambient light that have so far been stated or implied. Radiant light causes illumination; ambient light is the result of illumination. Radiant light diverges from an energy source; ambient light converges to a point of observation. Radiant light must consist of an infinitely dense set of rays; ambient light can be thought of as a set of solid angles having a common apex. Radiant light from a point source is not different in different directions; ambient light at a point is different in different directions. Radiant light has no structure; ambient light has structure. Radiant light is propagated; ambient light is not, it is simply there. Radiant light comes from atoms and returns to atoms; ambient light depends upon an environment of surfaces. Radiant light is energy; ambient light can be information.

## The Structuring of Ambient Light

Only insofar as ambient light has *structure* does it specify the environment. I mean by this that the light at the point of observation has to be different in different directions (or there have to be *differences* in different directions) in order for it to contain any information. The differences are principally differences of intensity. The term that will be used to describe ambient light with structure is an *ambient optic array*. This implies an arrangement of some sort, that is, a pattern, a texture, or a configuration. The array has to have parts. The ambient light cannot be homogeneous or blank. (See the illustrations in Chapter 5.)

What would be the limiting case of ambient light *without* structure? It would arise if the air were filled with such a dense fog that the light could not reverberate between surfaces but only between the droplets or particles in the medium. The air would then be translucent but not transparent. Multiple reflection would occur only between closely packed microspheres, yielding a sort of microillumination of things too small to see. At any point of observation there would be radiation, but without differences in different directions, without transitions or gradations of intensity, there would be no structure and no array. Similarly, homogeneous ambient light would occur inside a translucent shell of some strongly diffusing substance that was illuminated from outside. The shell would transmit light but not structure.

In the case of unstructured ambient light, an environment is not specified and no information *about* an environment is available. Since the light is undifferentiated, it cannot be discriminated, and there is no information in *any* meaning of that term. The ambient light in this respect is no different from ambient darkness. An environment could exist behind the fog or the darkness, or nothing could exist; either alternative is possible. In the case of ambient light that is unstructured in one part and structured in an adjacent part, such as the blue sky above the horizon and the textured region below it, the former specifies a void and the latter a surface. Similarly, the homogeneous area between clouds specifies emptiness, and the heterogeneous areas specify clouds.

The structuring of ambient light by surfaces, especially by their pigmentation and their layout, will be described in the next chapter. Chiefly, it is the opaque surfaces of the world that reflect light, but we must also consider the luminous surfaces that emit light and the semitransparent surfaces that transmit light. As far as the evidence goes, we will describe how the light specifies these surfaces, their composition, texture, color, and layout, their gross properties, not their atomic properties. And this specifying of them is useful information about them.

## Stimulation and Stimulus Information

In order to stimulate a photoreceptor, that is, to excite it and make it “fire,” light energy must be absorbed by it, and this energy must exceed a certain characteristic amount known as the *threshold* of the receptor. Energy must be *transduced*, as the physiologist likes to put it, from one form to another. The rule is supposed to hold for each of a whole bank of photoreceptors, such as is found in the retina. Hence, if an eye were to be stationed at some point where there is ambient light, part of the light would enter the pupil, be absorbed, and act as stimulation. If *no* eye or any other body that absorbs light is stationed at that point, the flying photons in the air (or the wave fronts) would simply pass through the point without interfering with one another. Only *potential* stimulation exists at such a point. *Actual* stimulation depends on the presence of photoreceptors.

Consider an observer with an eye at a point in a fog-filled medium. The receptors in the retina would be stimulated, and there would consequently be impulses in the fibers of the optic nerve. But the light entering the pupil of the eye would not be different in different directions; it would be unfocusable, and no image could be formed on the retina. There could be no retinal image because the light on the retina would be just as homogeneous as the ambient light outside the eye. The possessor of the eye could not *fix* it on anything, and the eye would drift aimlessly. He could not look from one item to another, for no items would be present. If he turned the eye, the experience would be just what it was before. If he moved the eye forward in space, nothing in the field of view would change. Nothing he could do would make any difference in what he could experience, with this single exception: if he closed the eye, an experience that he might call brightness would give way to one he might call darkness. He could distinguish between stimulation of his photoreceptors and nonstimulation of them. But as far as perceiving goes, his eye would be just as blind when light entered it as it would be when light did not.

This hypothetical case demonstrates the difference between the retina and the eye, that is, the difference between receptors and a perceptual organ. Receptors are *stimulated*, whereas an organ is *activated*. There can be stimulation of a retina by light without any activation of the eye by stimulus information. Actually, the eye is part of a dual organ, one of a pair of mobile eyes, and they are set in a head that can turn, attached to a body that can move from place to place. These organs make a hierarchy and constitute what I have called a *perceptual system* (Gibson, 1966b, Ch. 3). Such a system is never simply stimulated but instead can go into activity in the presence of stimulus information. The characteristic activities of the visual system will be described in Chapter 12 of this book.

The distinction between stimulation for receptors and stimulus information for the visual system is crucial for what is to follow. Receptors are passive, elementary, anatomical components of an eye that, in turn, is only an organ of the complete system (Gibson, 1966b, Ch. 2). The traditional conception of a sense is almost wholly abandoned in this new approach. Stimulation by light and corresponding sensations of brightness are traditionally supposed to be the *basis* of visual perception. The inputs of the nerves are supposed to be the data on which the perceptual processes in the brain operate. But I make a quite different assumption, because the evidence suggests that stimuli as such contain no information, that brightness sensations are not elements of perception, and that inputs of the retina are not sensory elements on which the brain operates.

Visual perception can fail not only for lack of stimulation but also for lack of stimulus information. In homogeneous ambient darkness, vision fails for lack of stimulation. In homogeneous ambient light, vision fails for lack of information, even with adequate stimulation and corresponding sensations.

## Do we Ever See Light as Such?

The difference between stimulation and stimulus information can be shown in another way, by considering two contradictory assertions: (1) nothing can be seen, properly speaking, but light; and (2) light, properly speaking, can never be seen. At least one of these assertions must be wrong.

Classical optics, comparing the eye to a camera, has taught that nothing can possibly get into the eye but light in the form of rays or wave fronts. The only alternative to this doctrine seemed to be the naive theory that little copies of objects got into the eye. If all that can ever reach the retina is light in this form, then it would follow that all we can ever *see* is this light. Sensations of light are the fundamental basis of visual perception, the data, or what is *given*. This line of reasoning has seemed unassailable up to the present. It leads to what I have called the sensation-based theories of perception (Gibson, 1966*b*). We cannot see surfaces or objects or the environment directly; we only see them indirectly. All we ever see directly is what stimulates the eye, light. The verb *to see*, properly used, means *to have one or more sensations of light*.

What about the opposite assertion that we *never* see light? It may at first sound unreasonable, or perhaps false, but let us examine the statement carefully. Of all the possible things that can be seen, is light one of them?

A single point of light in an otherwise dark field is not “light”; it specifies either a very distant source of light or a very small source, a luminous object. A single instant or “flash” of such a point specifies a brief event at the source, that is, the *on* and the *off*. A fire with coals or flames, a lamp with a wick or filament, a sun or a moon—all these are quite specific objects and are so specified; no one sees merely light. What about a luminous *field*, such as the sky? To me it seems that I see the sky, not the luminosity as such. What about a *beam* of light in the air? But this is not seeing light, because the beam is only visible if there are illuminated particles in the medium. The same is true of the shafts of sunlight seen in clouds under certain conditions.

One can perceive a rainbow, to be sure, a spectrum, but even so that is not the seeing of light. Halos, highlights on water, and scintillations of various kinds are all manifestations of light, not light as such. The only way we see illumination, I believe, is by way of that which is illuminated, the surface on which the beam falls, the cloud, or the particles that are lighted. We do not see the light that is *in* the air, or that *fills* the air. If all this is correct, it becomes quite reasonable to assert that all we ever see is the environment or facts about the environment, never photons or waves or radiant energy.

What about the sensation of being dazzled by looking at the sun, or the sensation of glare that one gets from looking at glossy surfaces that reflect an intense source? Are these not sensations of light as such, and do we not then see pure physical energy? Even in this case, I would argue that the answer is no; we are perceiving a state of the eye akin to pain, arising from excessive stimulation.

We perceive a fact about the body as distinguished from a fact about the world, the fact of overstimulation but not the light that caused it. And the experiencing of facts about the body is not the basis of experiencing facts about the world.

If light in the exact sense of the term is never seen as such, it follows that seeing the environment cannot be *based on* seeing light as such. The stimulation of the receptors in the retina cannot be seen, paradoxical as this may sound. The supposed sensations resulting from this stimulation are not the data for perception. Stimulation may be a necessary condition for seeing, but it is not sufficient. There has to be stimulus information available to the perceptual system, not just stimulation of the receptors.

In ordinary speech we say that vision depends on light, and we do not need to know physics to be able to say it with confidence. All of us, including every child, know what it is like to be “in the dark.” We cannot see anything, not even our own bodies. Approaching dangers and collisions ahead cannot be foreseen, and this is, with some reason, alarming. But what we mean when we say that vision depends on light is that it depends on illumination and on sources of illumination. We do not necessarily mean that we have to see light or have sensations of light in order to see anything else.

Just as the stimulation of the receptors in the retina cannot be seen, so the mechanical stimulation of the receptors in the skin cannot be felt, and the stimulation of the hair cells in the inner ear cannot be heard. So also the chemical stimulation of the receptors in the tongue cannot be tasted, and the stimulation of the receptors in the nasal membrane cannot be smelled. We do not perceive stimuli.

## The Concept of the Stimulus as an Application of Energy

The explicit assumption that only the receptors of observers are stimulated and that their sense organs are not stimulated but activated is in disagreement with what most psychologists take for granted. They blithely use the verb *stimulate* and the noun *stimulus* in various ways not consistent with one another. It is convenient and easy to do so, but if the words are slippery and if we allow ourselves to slide from one meaning to another unawares, we are confused without knowing it. I once examined the writings of modern psychology and found eight separate ways in which the use of the term *stimulus* was equivocal (Gibson, 1960a).

The concept of the stimulus comes from physiology, where it first meant whatever application of energy fires a nerve cell or touches off a receptor or excites a reflex response. It was taken over by psychology, because it seemed that a stimulus explained not only the arousal of a sensation but the arousal of a response, including responses much more elaborate than reflexes. If all behavior consisted of responses to stimuli, it looked as if a truly scientific psychology

could be founded. This was the stimulus-response formula. It was indeed promising. Both stimuli and responses could be measured. But a great variety of environmental facts had to be called stimuli because a variety of things can be responded to. If anything in the world can be called a stimulus, the concept has got out of hand and its original meaning has been lost. I suggest that we go back to its meaning in physiology. In this book I shall use the term strictly. For I now wish to make the clearest possible contrast between stimulus energy and stimulus information.

Note that a stimulus, strictly speaking in the physiologist's sense, is *anything* that touches off a receptor or causes a response; it is the *effective* stimulus, and whatever application of energy touches off the receptor is effective. The photoreceptors in the eye are usually triggered by light but not necessarily; they are also triggered by mechanical or electrical energy. The mechanoreceptors of the skin and the chemoreceptors of the mouth and nose are more or less specialized for mechanical and chemical energy respectively but not completely so; they are just especially "sensitive" to those kinds of energy. A stimulus in this strict meaning carries no information about its source in the world; that is, it does not specify its source. Only stimulation that comes in a structured array and that changes over time specifies its external source.

Note also that a stimulus, strictly speaking, is temporary. There is nothing lasting about it, as there is about a persisting object of the environment. A stimulus must begin and end. If it persists, the response of the receptor tapers off and ceases; the term for this is *sensory adaptation*. Hence, a permanent object cannot possibly be specified by a stimulus. The stimulus information for an object would have to reside in something persisting during an otherwise changing flow of stimulation. And note above all that an object cannot *be* a stimulus, although current thinking carelessly takes for granted that it is one.

An application of stimulus energy exceeding the threshold can be said to *cause* a response of the sensory mechanism, and the response is an *effect*. But the presence of stimulus information cannot be said to cause perception. Perception is not a response to a stimulus but an act of information pickup. Perception may or may not occur in the presence of information. Perceptual awareness, unlike sensory awareness, does not have any discoverable stimulus threshold. It depends on the age of the perceiver, how well he has learned to perceive, and how strongly he is motivated to perceive. If perceptions are based on sensations and sensations have thresholds, then perceptions should have thresholds. But they do not, and the reason for this, I believe, is that perceptions are not based on sensations. There are magnitudes for applied stimuli above which sensations occur and below which they do not. But there is no magnitude of information above which perceiving occurs and below which it does not.

When stimulus energy is transformed into nervous impulses, they are said to be *transmitted* to the brain. But stimulus information is not anything that could possibly be sent up a nerve bundle and delivered to the brain, inasmuch as it has

to be isolated and extracted from the ambient energy. Information as here conceived is not transmitted or conveyed, does not consist of signals or messages, and does not entail a sender and a receiver. This will be elaborated later.

When a small packet of stimulus energy is absorbed by a receptor, what is lost to the environment is gained by the living cells. The amount of energy may be as low as a few quanta, but nevertheless energy is conserved. In contrast to this fact, stimulus information is not lost from the environment when it is gained by the observer. There is no such thing as conservation of information. It is not limited in amount. The available information in ambient light, vibration, contact, and chemical action is inexhaustible.

A stimulus, then, carries some of the meaning that the word had in Latin, a goad stuck into the skin of an ox. It is a brief and discrete application of energy to a sensitive surface. As such, it specifies little beyond itself; it contains no information. But a flowing array of stimulation is a different matter entirely.

### **Ambient Energy as Available Stimulation**

The environment of an observer was said to consist of substances, the medium, and surfaces. Gravity, heat, light, sound, and volatile substances fill the medium. Chemical and mechanical contacts and vibrations impinge on the observer's body. The observer is immersed as it were in a sea of physical energy. It is a flowing sea, for it changes and undergoes cycles of change, especially of temperature and illumination. The observer, being an organism, exchanges energy with the environment by respiration, food consumption, and behavior. A very small fraction of this ambient sea of energy constitutes stimulation and provides information. The fraction is small, for only the ambient odor entering the nose is effective for smelling, only the train of air vibrations impinging on the eardrums is effective for hearing, and only the ambient light at the entrance pupil of an eye is effective for vision. But this tiny portion of the sea of energy is crucial for survival, because it contains information for things at a distance.

It should be obvious by now that this minute inflow of stimulus energy does not consist of discrete inputs—that stimulation does not consist of stimuli. The flow is continuous. There are, of course, episodes in the flow, but these are nested within one another and cannot be cut up into elementary units. Stimulation is not momentary.

Radiant energy of all wavelengths falls on an individual, that is, impinges on the skin. The infrared radiation will give warmth, and the ultraviolet will cause sunburn, but the narrow band of radiation in between, light, is the only kind that will excite the photoreceptors in the eye after entering the pupil. An eye, or at least a vertebrate chambered eye as distinguished from the faceted eye of an insect, usually takes in something less than a hemisphere of the ambient light, according to G. L. Walls (1942). A pair of eyes like those of a rabbit, pointing in opposite directions, takes in nearly the whole of the ambient



light at the same time. Ambient light is structured, as we have seen. And the purpose of a dual ocular system is to register this structure or, more exactly, the invariants of its changing structure. Ambient light is usually very rich in what we call pattern and change. The retinal images register both. And a retinal image involves stimulation of its receptive surface but not, as often supposed, a set or a sequence of stimuli.

### The Orthodox Theory of the Retinal Image

The generally accepted theory of the eye does not acknowledge that it registers the invariant structure of ambient light but asserts that it forms *an image of an object* on the back of the eye. The object, of course, is in the outer world, and the back of the eye is a photoreceptive surface attached to a nerve bundle. What is the difference between these theories?

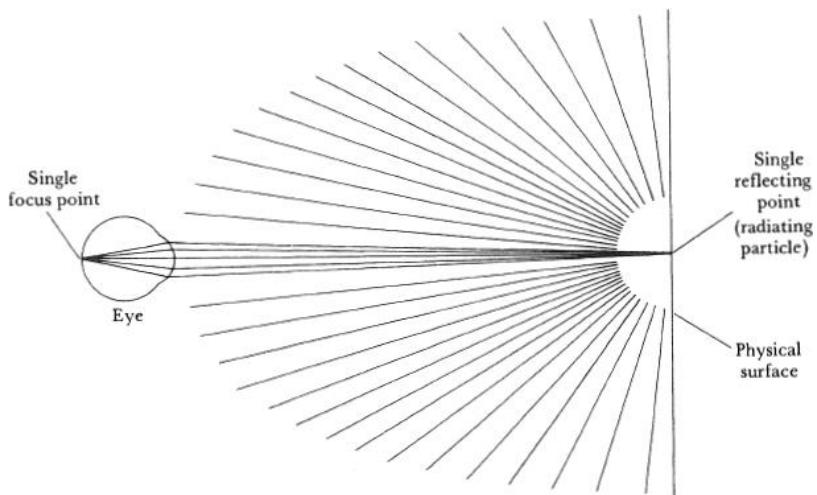
The theory of image formation in a dark chamber like the eye goes back more than 350 years to Johannes Kepler. The germ of the theory as stated by him was that everything visible radiates, more particularly that every point on a body can emit rays in all directions. An opaque reflecting surface, to be sure, receives radiation from a source and then re-emits it, but in effect it becomes a collection of radiating point sources. If an eye is present, a small cone of diverging rays enters the pupil from each point source and is caused by the lens to converge to another point on the retina. The diverging and converging rays make what is called a *focused pencil* of rays. The dense set of focus points on the retina constitutes the retinal image. There is a one-to-one projective correspondence between radiating points and focus points.

A focused pencil of rays consists of two parts, the diverging cone of radiant light and the converging cone of rays refracted by the lens, one cone with its vertex on the object and the other with its vertex in the image. This pencil is then repeated for every point on the object. Thus, there is a limitless set of rays in each pencil and a limitless set of pencils for each object. The history of optics suggests that Kepler was mainly responsible for this extraordinary intellectual invention. It involved difficult ideas, but it was and still is the unchallenged foundation of the theory of image formation. The notion of an object composed of points has proved over the centuries to be sympathetic to physicists, because most of them assume that an object really consists of its atoms. And later, in the nineteenth century, the notion of a retinal image consisting of sharp points of focused light did not seem strange to physiologists because they were familiar with punctate stimuli, for example, on the skin.

This theory of point-to-point correspondence between an object and its image lends itself to mathematical analysis. It can be abstracted to the concepts of projective geometry and can be applied with great success to the design of cameras and projectors, that is, to the making of pictures with light, photography. The theory permits lenses to be made with smaller “aberrations,” that

is, with finer points in the point-to-point correspondence. It works beautifully, in short, for the images that fall on screens or surfaces and that are intended to be looked at. But this success makes it tempting to believe that the image on the retina falls on a kind of screen and is itself something intended to be looked at, that is, a picture. It leads to one of the most seductive fallacies in the history of psychology—that the retinal image is something to be seen. I call this the “little man in the brain” theory of the retinal image (Gibson, 1966*b*, p. 226), which conceives the eye as a camera at the end of a nerve cable that transmits the image to the brain. Then there has to be a little man, a homunculus, seated in the brain who looks at this physiological image. The little man would have to have an eye to see it with, of course, a *little* eye with a *little* retinal image connected to a *little* brain, and so we have explained nothing by this theory. We are in fact worse off than before, since we are confronted with the paradox of an infinite series of little men, each within the other and each looking at the brain of the next bigger man.

If the retinal image is not transmitted to the brain as a whole, the only alternative has seemed to be that it is transmitted to the brain element by element, that is, by signals in the fibers of the optic nerve. There would then be an element-to-element correspondence between image and brain analogous to the



**FIGURE 4.3** A focused pencil of rays connecting a radiating point on a surface with a focus point in the retinal image.

The rays in the pencil are supposed to be infinitely dense. Note that only the rays that enter the pupil are effective for vision. (From *The Perception of the Visual World* by James Jerome Gibson and used with the agreement of the reprint publisher, Greenwood Press, Inc.)

point-to-point correspondence between object and image. This seems to avoid the fallacy of the little man in the brain who looks at an image, but it entails all the difficulties of what I have called the sensation-based theories of perception. The correspondence between the spots of light on the retina and the spots of sensation in the brain can only be a correspondence of intensity to brightness and of wavelength to color. If so, the brain is faced with the tremendous task of constructing a phenomenal environment out of spots differing in brightness and color. If these are what is seen directly, what is given for perception, if these are the data of sense, then the fact of perception is almost miraculous.

### **JAMES MILL ON VISUAL SENSATION, 1829**

“When I lift my eyes from the paper on which I am writing, I see from my window trees and meadows, and horses and oxen, and distant hills. I see each of its proper size, of its proper form, and at its proper distance; and these particulars appear as *immediate* informations of the eye as the colors which I see by means of it. Yet philosophy has ascertained that we derive nothing from the eye whatever but sensations of color . . . . How then, is it that we receive accurate information by the eye of size and shape and distance? By association merely” (Mill, *Analysis of the Phenomena of the Human Mind*, 1829).

How is it indeed! Mill answered, *by association*. But others answered, *by innate ideas of space or by rational inference from the sensations or by interpretation of the data*. Still others have said, *by spontaneous organization of sensory inputs to the brain*. The current fashionable answer is, *by computerlike activities of the brain on neural signals*. We have empiricism, nativism, rationalism, Gestalt theory, and now information-processing theory. Their adherents would go on debating forever if we did not make a fresh start. Has philosophy ascertained that “we derive nothing from the eye whatever but sensations of color”? No. “Sensations of color” meant *dabs* or *spots* of color, as if in a painting. Perception does not begin that way.

Even the more sophisticated theory that the retinal image is transmitted as signals in the fibers of the optic nerve has the lurking implication of a little man in the brain. For these signals must be in code and therefore have to be decoded; signals are messages, and messages have to be interpreted. In both theories the eye sends, the nerve transmits, and a mind or spirit receives. Both theories carry the implication of a mind that is separate from a body.

It is not necessary to assume that *anything whatever* is transmitted along the optic nerve in the activity of perception. We need not believe that *either* an inverted picture or a set of messages is delivered to the brain. We can think of

vision as a perceptual system, the brain being simply part of the system. The eye is also part of the system, since retinal inputs lead to ocular adjustments and then to altered retinal inputs, and so on. The process is circular, not a one-way transmission. The eye-head-brain-body system registers the invariants in the structure of ambient light. The eye is not a camera that forms and delivers an image, nor is the retina simply a keyboard that can be struck by fingers of light.

### **A Demonstration that the Retinal Image is not Necessary for Vision**

We are apt to forget that an eye is not necessarily a dark chamber, on the back surface of which an inverted image is formed by a lens in the manner described by Kepler. Although the eyes of vertebrates and mollusks are of this sort, the eyes of arthropods are not. They have what is called a *compound eye*, with no chamber, no lens, and no sensory surface but with a closely packed set of receptive tubes called *ommatidia*. Each tube points in a different direction from every other tube, and presumably the organ can thus register differences of intensity in different directions. It is therefore part of a system that registers the structure of ambient light.

In a chapter on the evolutionary development of visual systems (Gibson, 1966b, Ch. 9), I described the chambered eye and the compound eye as two different ways of accepting an array of light coming from an environment (pp. 163 ff.). The camera eye has a concave mosaic of photoreceptors, a retina. The compound eye has a convex packet of photoreceptive light tubes. The former accepts an infinite number of pencils of light, each focused to a point and combining to make a continuous image. The latter accepts a finite number of samples of ambient light, without focusing them and without forming an optical image. But if several thousand tubes are packed together, as in the eye of a dragonfly, visual perception is quite good. There is nothing behind a dragonfly's eye that could possibly be seen by you, no image on a surface, no picture. But nevertheless the dragonfly sees its environment.

Zoologists who study insect vision are so respectful of optics as taught in physics textbooks that they are constrained to think of a sort of *upright* image as being formed in the insect eye. But this notion is both vague and self-contradictory. There is no screen on which an image could be formed. The concept of an ambient optic array, even if not recognized in optics, is a better foundation for the understanding of vision in general than the concept of the retinal image. The registering of differences of intensity in different directions is *necessary* for visual perception; the formation of a retinal image is not.

### **The Concept of Optical Information**

The concept of information with which we are most familiar is derived from our experiences of communicating with other people and being communicated

with, not from our experience of perceiving the environment directly. We tend to think of information primarily as being sent and received, and we assume that some intermediate kind of transmission has to occur, a “medium” of communication or a “channel” along which the information is said to flow. Information in this sense consists of messages, signs, and signals. In early times messages, which could be oral, written, or pictorial, had to be sent by runner or by horseman. Then the semaphore system was invented, and then the electrical telegraph, wireless telegraphy, the telephone, television, and so on at an accelerated rate of development.

### **THE FALLACY OF THE IMAGE IN THE EYE**

Ever since someone peeled off the back of the excised eye of a slaughtered ox and, holding it up in front of a scene, observed a tiny, colored, inverted image of the scene on the transparent retina, we have been tempted to draw a false conclusion. We think of the image as *something to be seen*, a picture on a screen. You can see it if you take out the ox's eye, so why shouldn't the ox see it? The fallacy ought to be evident.

The question of how we can see the world as upright when the retinal image is inverted arises because of this false conclusion. All the experiments on this famous question have come to nothing. The retinal image is not anything that can be seen. The famous experiment of G. M. Stratton (1897) on reinverting the retinal image gave unintelligible results because it was misconceived.

We also communicate with others by making a picture on a surface (clay tablet, papyrus, paper, wall, canvas, or screen) and by making a sculpture, a model, or a solid image. In the history of image-making, the chief technological revolution was brought about by the invention of photography, that is, of a photosensitive surface that could be placed at the back of a darkened chamber with a lens in front. This kind of communication, which we call graphic or plastic, does not consist of signs or signals and is not so obviously a message from one person to another. It is not so obviously transmitted or conveyed. Pictures and sculptures are apt to be displayed, and thus they *contain* information and make it available for anyone who looks. They nevertheless are, like the spoken and written words of language, *man-made*. They provide information that, like the information conveyed by words, is mediated by the perception of the first observer. They do not permit firsthand experience—only experience at second hand.

The ambient stimulus information available in the sea of energy around us is quite different. The information for perception is not transmitted, does not

consist of signals, and does not entail a sender and a receiver. The environment does not communicate with the observers who inhabit it. Why should the world speak to us? The concept of stimuli as signals to be interpreted implies some such nonsense as a world-soul trying to get through to us. The world is *specified* in the structure of the light that reaches us, but it is entirely up to us to perceive it. The secrets of nature are not to be understood by the breaking of its code.

*Optical information*, the information that can be extracted from a flowing optic array, is a concept with which we are not at all familiar. Being intellectually lazy, we try to understand perception in the same way we understand communication, in terms of the familiar. There is a vast literature nowadays of speculation about the media of communication. Much of it is undisciplined and vague. The concept of information most of us have comes from that literature. But this is not the concept that will be adopted in this book. For we cannot explain perception in terms of communication; it is quite the other way around. We cannot convey information about the world to others unless we have perceived the world. And the available information for our perception is radically different from the information we convey.

## Summary

Ecological optics is concerned with many-times-reflected light in the medium, that is, *illumination*. Physical optics is concerned with electromagnetic energy, that is, *radiation*.

Ambient light coming to a point in the air is profoundly different from radiant light leaving a point source. The ambient light has structure, whereas the radiant light does not. Hence, ambient light makes available information about reflecting surfaces, whereas radiant light can at most transmit information about the atoms from which it comes.

If the ambient light were unstructured or undifferentiated, it would provide no information about an environment, although it would stimulate the photoreceptors of an eye. Thus, there is a clear distinction between stimulus information and stimulation. We do not have sensations of light triggered by stimuli under normal conditions. The doctrine of discrete stimuli does not apply to ordinary vision.

The orthodox theory of the formation of an image on a screen, based on the correspondence between radiating points and focus points, is rejected as the basis for an explanation of ecological vision. This theory applies to the design of optical instruments and cameras, but it is a seductive fallacy to conceive the ocular system in this way. One of the worst results of the fallacy is the inference that the retinal image is transmitted to the brain.

The information that can be extracted from ambient light is not the kind of information that is transmitted over a channel. There is no sender outside the head and no receiver inside the head.

# 5

## THE AMBIENT OPTIC ARRAY

The central concept of ecological optics is the ambient optic array at a point of observation. To be an *array* means to have an arrangement, and to be *ambient at a point* means to surround a position in the environment that could be occupied by an observer. The position may or may not be occupied; for the present, let us treat it as if it were not.

What is implied more specifically by an *arrangement*? So far I have suggested only that it has *structure*, which is not very explicit. The *absence of structure* is easier to describe. This would be a homogeneous field with no differences of intensity in different parts. An array cannot be homogeneous; it must be heterogeneous. That is, it cannot be undifferentiated, it must be differentiated; it cannot be empty, it must be filled; it cannot be formless, it must be formed. These contrasting terms are still unsatisfactory, however. It is difficult to define the notion of structure. In the effort to clarify it, a radical proposal will be made having to do with *invariant* structure.

What is implied by *ambient at a point*? The answer to this question is not so difficult. To be ambient, an array must surround the point completely. It must be environing. The field must be closed, in the geometrical sense of that term, the sense in which the surface of a sphere returns upon itself. More precisely, the field is unbounded. Note that the field provided by a picture on a plane surface does not satisfy this criterion. No picture can be ambient, and even a picture said to be panoramic is never a completely closed sphere. Note also that the temporary field of view of an observer does not satisfy the criterion, for it also has boundaries. This fact is obviously of the greatest importance, and we shall return to it in Chapter 7 and again in Chapter 12.

Finally, what is implied by the term *point* in the phrase *point of observation*? Instead of a geometrical point in abstract space, I mean a position in ecological

space, in a medium instead of in a void. It is a place where an observer *might* be and from which an act of observation *could* be made. Whereas abstract space consists of points, ecological space consists of places—locations or positions.

A sharp distinction will be made between the ambient array at an unoccupied point of observation and the array at a point that is occupied by an observer, human or other. When the position becomes occupied, something very interesting happens to the ambient array: it contains information about the body of the observer. This modification of the array will be given due consideration later.

The point of observation in ecological optics might seem to be the equivalent of the station point in perspective geometry, the kind of perspective used in the making of a representative painting. The station point is the point of projection for the picture plane on which the scene is projected. But the terms are not at all equivalent and should not be confused, as we shall see. A station point has to be stationary. It cannot move relative to the world, and it must not move relative to the picture plane. But a point of observation is never stationary, except as a limiting case. Observers move about in the environment, and observation is typically made from a moving position.

## How is Ambient Light Structured? Preliminary Considerations

If we reject the assumption that the environment consists of atoms in space and that, hence, the light coming to a point in space consists of rays from these atoms, what do we accept? It is tempting to assume that the environment consists of *objects* in space and that, hence, the ambient array consists of *closed-contour forms* in an otherwise empty field, or “figures on a ground.” For each object in space, there would correspond a form in the optic array. But this assumption is not close to being good enough and must also be rejected. A form in the array could not correspond to each object in space, because some objects are hidden behind others. And in any case, to put it radically, the environment does not consist of objects. The environment consists of the earth and the sky with objects *on* the earth and *in* the sky, of mountains and clouds, fires and sunsets, pebbles and stars. Not all of these are segregated objects, and some of them are nested within one another, and some move, and some are animate. But the environment is all these various things—places, surfaces, layouts, motions, events, animals, people, and artifacts that structure the light at points of observation. The array at a point does not consist of forms in a field. The figure-ground phenomenon does not apply to the world in general. The notion of a closed contour, an outline, comes from the art of drawing an object, and the phenomenon comes from the experiment of presenting an observer with a drawing to find out what she perceives. But this is not the only way, or even the best way, to investigate perception.

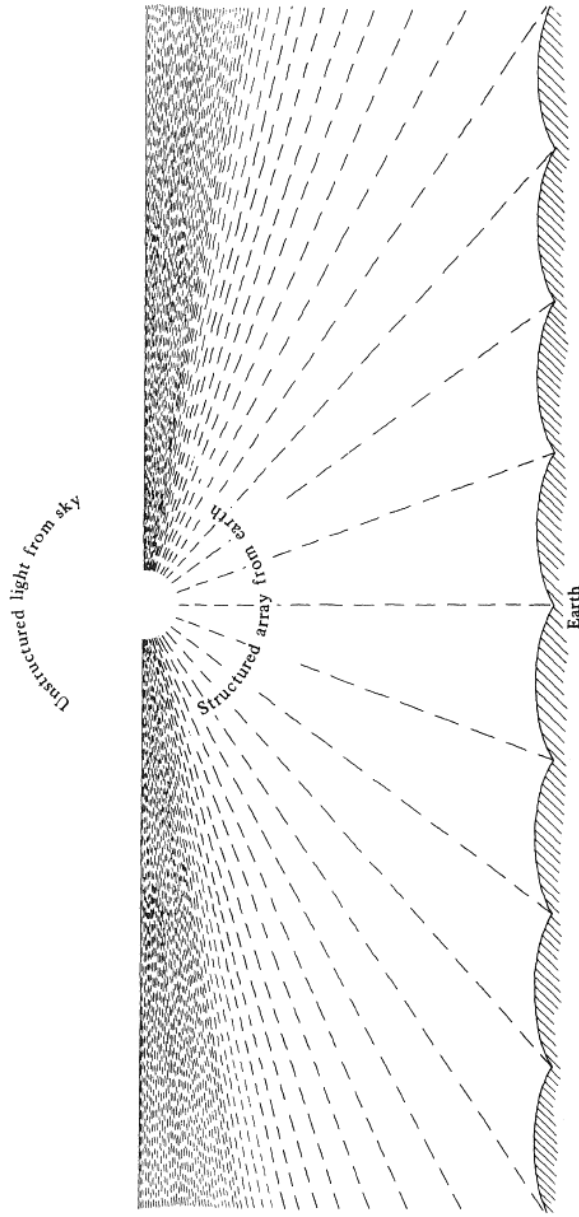


We obtain a better notion of the structure of ambient light when we think of it as divided and subdivided into component parts. For the terrestrial environment, the sky-earth contrast divides the unbounded spherical field into two hemispheres, the upper being brighter than the lower. Then both are further subdivided, the lower much more elaborately than the upper and in quite a different way. The components of the earth, as I suggested in Chapter 1, are nested at different levels of size—for example, mountains, canyons, trees, leaves, and cells. The components of the *array* from the earth also fall into a hierarchy of subordinate levels of size, but the components of the array are quite different, of course, from the components of the earth. The components of the array are the *visual angles* from the mountains, canyons, trees, and leaves (actually, what are called *solid angles* in geometry), and they are conventionally measured in degrees, minutes, and seconds instead of kilometers, meters, and millimeters. They are *intercept angles*, as we shall see. All these optical components of the array, whatever their size, become vanishingly small at the margin between earth and sky, the horizon; moreover, they change in size whenever the point of observation moves. The substantial components of the earth, on the other hand, do not change in size.

There are several advantages in conceiving the optic array in this way, as a nested hierarchy of solid angles all having a common apex instead of as a set of rays intersecting at a point. Every solid angle, no matter how small, has form in the sense that its cross-section has a form, and a solid angle is quite unlike a ray in this respect. Each solid angle is unique, whereas a ray is not unique and can only be identified arbitrarily, by a pair of coordinates. Solid angles can fill up a sphere in the way that sectors fill up a circle, but it must be remembered that there are angles within angles, so that their sum does not *add* up to a sphere. The surface of the sphere whose center is the common apex of all the solid angles can be thought of as a kind of transparent film or shell, but it should not be thought of as a picture.

The structure of an optic array, so conceived, is without gaps. It does not consist of points or spots that are discrete. It is completely filled. Every component is found to consist of smaller components. Within the boundaries of any form, however small, there are always other forms. This means that the array is more like a hierarchy than like a matrix and that it should not be analyzed into a set of spots of light, each with a locus and each with a determinate intensity and frequency. In an ambient hierarchical structure, loci are not defined by pairs of coordinates, for the relation of location is not given by degrees of azimuth and elevation (for example) but by the relation of inclusion.

The difference between the relation of *metric location* and the relation of *inclusion* can be illustrated by the following fact. The stars in the sky can be located conveniently by degrees to the right of north and degrees up from the horizon. But each star can also be located by its inclusion in one of the



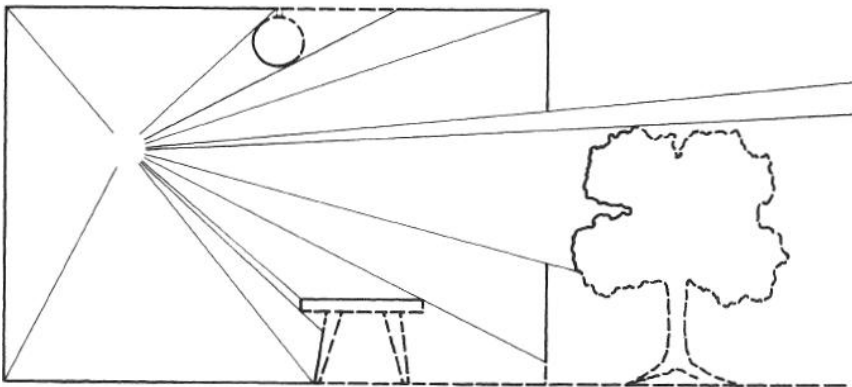
**FIGURE 5.1** The ambient optic array from a wrinkled earth outdoors under the sky.

In this illustration it is assumed that illumination has reached a steady state. The earth is shown as wrinkled or humped, but not as cluttered. The dashed lines in this drawing depict the envelopes of visual solid angles, not rays of light. The nesting of these solid angles has not been shown. The contrasts in this diagram are caused by differential illumination of the humps of the earth. Compare this with the photograph of hills and valleys in Figure 5.9. This is an optic array at a single fixed point of observation. It illustrates the main invariants of natural perspective: the separation of the two hemispheres of the ambient array at the horizon, and the increasing density of the optical texture toward its maximum at the horizon. These are invariant even when the array flows, as it does when the point of observation moves.

constellations and by the superordinate pattern of the whole sky. Similarly, the optical structures that correspond to the leaves and trees and hills of the earth are each included in the next larger structure. The texture of the earth, of course, is dense compared to the constellations of discrete stars and thus even less dependent than they are on a coordinate system. If this is so, the perception of the direction of some particular item on the earth, its direction-from-here, is not a problem in its own right. The perceiving of the environment does not consist of perceptions of the differing directions of the items of the environment.

### The Laws of Natural Perspective: The Intercept Angle

The notion of a visual angle with its apex at the eye and its base at an object in the world is very old. It goes back to Euclid who postulated what he called a “visual cone” for each object in space. The term is not exact, for the object need not be circular and the figure does not have to be a cone. Ptolemy spoke of the “visual pyramid,” which implied that the object was rectangular. Actually, we should refer to the *face* of an object, which can have any shape whatever, and to a corresponding *solid angle*, having an envelope. A cross-section of this envelope is what we call the *outline of the object*. We can now note that the solid angle shrinks as the distance of the object from the apex increases, and it is laterally squeezed as the face of the object is slanted or turned. These are the two main laws of perspective for objects. Euclid and Ptolemy and their successors for



**FIGURE 5.2** The ambient optic array from a room with a window.

This drawing shows a cluttered environment where some surfaces are projected at the point of observation and the remainder are not, that is, where some are unhidden and the others are hidden. The hidden surfaces are indicated by dotted lines. Only the faces of the layout of surfaces are shown, not the facets of their surfaces, that is, their textures.

many centuries never doubted that objects were seen by means of these solid angles, whether conical, pyramidal, or otherwise. They were the basis of ancient optics. Nothing was then known of inverted retinal images, and the comparison of the eye with a camera would not be made for a thousand years. The ancients did not understand the eye, they were puzzled by light, they had no conception of the modern doctrine that nothing gets into the eye but light, but they were clear about visual angles.

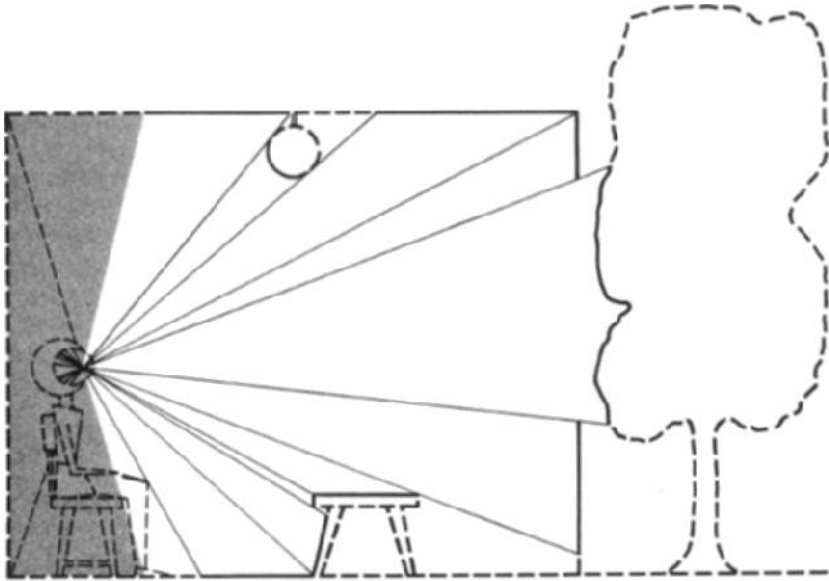
The conception of the ambient optic array as a set of solid angles corresponding to objects is thus a continuation of ancient and medieval optics. Instead of only freestanding objects present to an eye, however, I postulate an environment of illuminated surfaces. And instead of a group of solid angles, I postulate a nested complex of them. The large solid angles in the array come from the *faces* of this layout, from the facades of detached objects, and from the interspaces or holes that we call background or sky (which Euclid and Ptolemy seem never to have thought of). The small solid angles in the array come from what might be called the *facets* of the layout as distinguished from the faces, the textures of the surfaces as distinguished from their forms. As already has been emphasized, however, the distinction between these size-levels is arbitrary.

Natural perspective, as I conceive it, is the study of an ambient array of solid angles that correspond to certain distinct geometrical parts of a terrestrial environment, those that are separated by edges and corners. There are elegant trigonometric relations between the angles and the environmental parts. There are gradients of size and density of the angles along meridians of the lower half of the array, the earth, with sizes vanishing and density becoming infinite at the horizon. These relations contain a great amount of information about the parts of the earth. No one who understood them would think of questioning their validity. It is a perfectly clear and straightforward discipline, although neglected and undeveloped. But the environment does not *wholly* consist of sharply differentiated geometrical parts or forms. Natural perspective does not apply to shadows with penumbras and patches of light. It does not apply to sunlit surfaces with varying degrees of illumination. It geometrizes the environment and thus oversimplifies it. The most serious limitation, however, is that natural perspective omits motion from consideration. The ambient optic array is treated as if its structure were frozen in time and as if the point of observation were motionless.

Although I have called this discipline *natural perspective*, the ancients called it *perspectiva*, the Latin word for what we now call *optics*. In modern times, the term *perspective* has come to mean a technique—the technique of picture-making. A picture is a surface, whether it be painted by hand or processed by photography, and perspective is the art of “representing” the geometrical relationships of natural objects on that surface. When the Renaissance painters discovered the procedures for perspective representation, they very

properly called the method *artificial perspective*. They understood that this had to be distinguished from the natural perspective that governed the ordinary perception of the environment. Since that time we have become so picture-minded, so dominated by pictorial thinking, that we have ceased to make the distinction. But to confuse pictorial perspective with natural perspective is to misconceive the problem of visual perception at the outset. The so-called cues for depth in a picture are not at all the same as the information for surface layout in a frozen ambient array, although pictorial thinking about perception tempts us to assume that they are the same. Pictures are artificial displays of information frozen in time, and this fact will be evident when the special kind of visual perception that is mediated by such displays is treated in detail in Part IV.

Natural perspective, as well as artificial perspective, is restricted in scope, being concerned only with a frozen optical structure. This restriction will be removed in what follows.



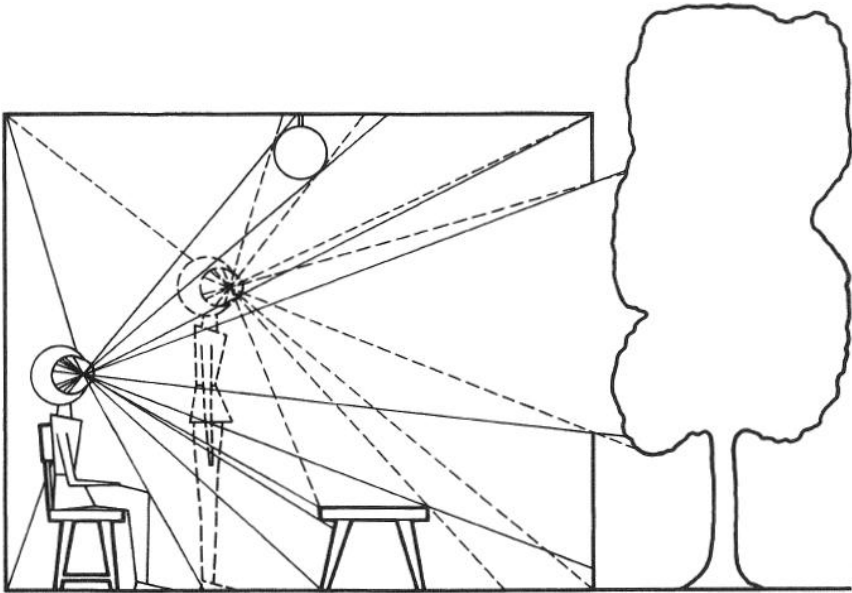
**FIGURE 5.3** The same ambient array with the point of observation occupied by a person.

When an observer is present at a point of observation, the visual system begins to function.

## Optical Structure with a Moving Point of Observation

A point of observation at rest is only the limiting case of a point of observation in motion, the null case. Observation implies movement, that is, locomotion with reference to the rigid environment, because all observers are animals and all animals are mobile. Plants do not observe but animals do, and plants do not move about but animals do. Hence, the structure of an optic array at a stationary point of observation is only a special case of the structure of an optic array at a moving point of observation. The point of observation normally proceeds along a path of locomotion, and the “forms” of the array change as locomotion proceeds. More particularly, every solid angle included within the array, large or small, is enlarged or reduced or compressed or, in some cases, wiped out. It is wiped out, of course, when its surface goes out of sight.

The optic array *changes*, of course, as the point of observation moves. But it also does *not* change, not completely. Some features of the array do not persist



**FIGURE 5.4** The change of the optic array brought about by a locomotor movement of the observer.

The thin solid lines indicate the ambient optic array for the seated observer, and the thin dashed lines the altered optic array after standing up and moving forward. The difference between the two arrays is specific to the difference between the points of observation, that is, to the path of locomotion. Note that the whole ambient array is changed, including the portion behind the head. And note that what was previously hidden becomes unhidden.

and some do. The changes come from the locomotion, and the nonchanges come from the rigid layout of the environmental surfaces. Hence, the nonchanges specify the layout and count as information about it; the changes specify locomotion and count as another kind of information, about the locomotion itself. We have to distinguish between two kinds of structure in a normal ambient array, and I shall call them the *perspective structure* and the *invariant structure*.

### ***Perspective Structure and Invariant Structure***

The term *structure* is vague, as we have seen. Let us suppose that a kind of essential structure underlies the superficial structure of an array when the point of observation moves. This essential structure consists of what is invariant despite the change. What is invariant does not emerge unequivocally except with a flux. The essentials become evident in the context of changing nonessentials.

Consider the paradox in the following piece of folk wisdom: "The more it changes, the more it is the same thing." Wherein is it true and wherein false? If *change* means *to become different but not to be converted into something else*, the assertion is true, and the saying emphasizes the fact that whatever is invariant is more evident with change than it would be without change. If *change* means *to become different by being converted into something else*, the assertion is self-contradictory, and the paradox arises. But this is not what the word ordinarily means. And assuredly it is not what change in the ambient array means. One arrangement does not become a wholly different arrangement by a displacement of viewpoint. There is no jump from one to another, only a variation of structure that serves to reveal the nonvariation of structure. The pattern of the array does not ordinarily scintillate; the forms of the array do not go from triangular to quadrangular, for example.

There are many invariants of structure, and some of them persist for long paths of locomotion while some persist only for short paths. But what I am calling the *perspective structure* changes with every displacement of the point of observation—the shorter the displacement the smaller the change, and the longer the displacement the greater the change. Assuming that the environment is never reduplicated from place to place, the arrested perspective is unique at each stationary point of observation, that is, for each point of observation there is one and only one arrested perspective. On the other hand, invariants of structure are common to all points of observation—some for all points in the whole terrestrial environment, some only for points within the boundaries of certain locales, and some only for points of observation within (say) a single room. But to repeat, the invariant structure separates off best when the frozen perspective structure begins to flow.

Consider, for example, the age-old question of how a rectangular surface like a tabletop can be given to sight when presumably all that an eye can see is a large number of forms that are trapezoids and only one form that is rectangular, that one being seen only when the eye is positioned on a line perpendicular to the center of the surface. The question has never been answered, but it

can be reformulated to ask, What are the invariants underlying the transforming perspectives in the array from the tabletop? What specifies the shape of this rigid surface as projected to a moving point of observation? Although the changing angles and proportions of the set of trapezoidal projections are a fact, the unchanging relations among the four angles and the invariant proportions over the set are another fact, equally important, and they uniquely specify the rectangular surface. There will be experimental evidence about optical transformations as information in Chapter 9.

We tend to think of each member of the set of trapezoidal projections from a rectangular object as being a form in space. A change is then a transition from one form to another, a transformation. But this habit of thought is misleading. Optical change is not a transition from one form to another but a reversible process. The superficial form becomes different, but the underlying form remains the same. The structure changes in some respects and does not change in others. More exactly, it is variant in some respects and invariant in others.

The geometrical habit of separating space from time and imagining sets of frozen forms in space is very strong. One can think of each point of observation in the medium as stationary and distinct. To each such point there would correspond a unique optic array. The set of all points is the space of the medium, and the corresponding set of all optic arrays is the whole of the available information about layout. The set of all line segments in the space specifies all the possible displacements of points of observation in the medium, and the corresponding set of transformation families gives the information that specifies all the possible paths. This is an elegant and abstract way of thinking, modeled on projective geometry. But it does not allow for the complexities of optical change and does not do justice to the fact that the optic array *flows in time* instead of going from one structure to another. What we need for the formulation of ecological optics are not the traditional notions of space and time but the concepts of variance and invariance considered as reciprocal to one another. The notion of a *set* of stationary points of observation in the medium is appropriate for the problem of a whole crowd of observers standing in different positions, each of them perceiving the environment from his own point of view. But even so, the fact that all observers can perceive the same environment depends on the fact that each point of view can move to any other point of view.

### **REDUPLICATION**

It is easy to make copies or duplicates of a picture but the world is never exactly the same in one place as it is in another. Nor is one organism ever exactly the same as another. One cubic yard of empty abstract space is exactly the same as another, but that is a different matter.



### *The Significance of Changing Perspective in the Ambient Array*

When the moving point of observation is understood as the general case, the stationary point of observation is more intelligible. It no longer is conceived as a single geometrical point in space but as a pause in locomotion, as a temporarily fixed position relative to the environment. Accordingly, an arrested perspective structure in the ambient array specifies to an observer such a fixed position, that is, rest; and a flowing perspective structure specifies an unfixed position, that is, locomotion. The optical information for distinguishing locomotion from nonlocomotion is available, and this is extremely valuable for all observers, human or animal. In physics the motion of an observer in space is "relative," inasmuch, as what we call motion with reference to one chosen frame of reference may be nonmotion with reference to another frame of reference. In ecology this does not hold, and the locomotion of an observer in the environment is absolute. The environment is simply that with respect to which either locomotion or a state of rest occurs, and the problem of relativity does not arise.

Locomotion and rest go with flowing and frozen perspective structure in the ambient array; they are what the flow and the nonflow *mean*. They contain information about the potential observer, not information about the environment, as the invariants do. But note that information about a world that surrounds a point of observation implies information about the point of observation that is surrounded by a world. Each kind of information implies the other. Later, in discussing the occupied point of observation, I shall call the former *exterospecific information* and the latter *propriospecific information*.

Not only does flowing perspective structure specify locomotion, but the particular instance of flow specifies the particular path of locomotion. That is, the difference of perspective between the beginning and the end of the optical change is specific to the difference of position between the beginning and the end of the locomotor displacement. But more than that, the *course* of the optical flow is specific to the *route* the path of locomotion takes through the environment. Between one place and another there are many different routes. The two places are specified by their different arrested perspectives, but the different routes between them are in correspondence with different optical sequences between the two perspectives. There will be more of this later. It is enough now to point out that the visual control of locomotion by an observer, purposive locomotion such as homing, migrating, finding one's way, getting from place to place, and being oriented, depends on just the kind of sequential optical information described.

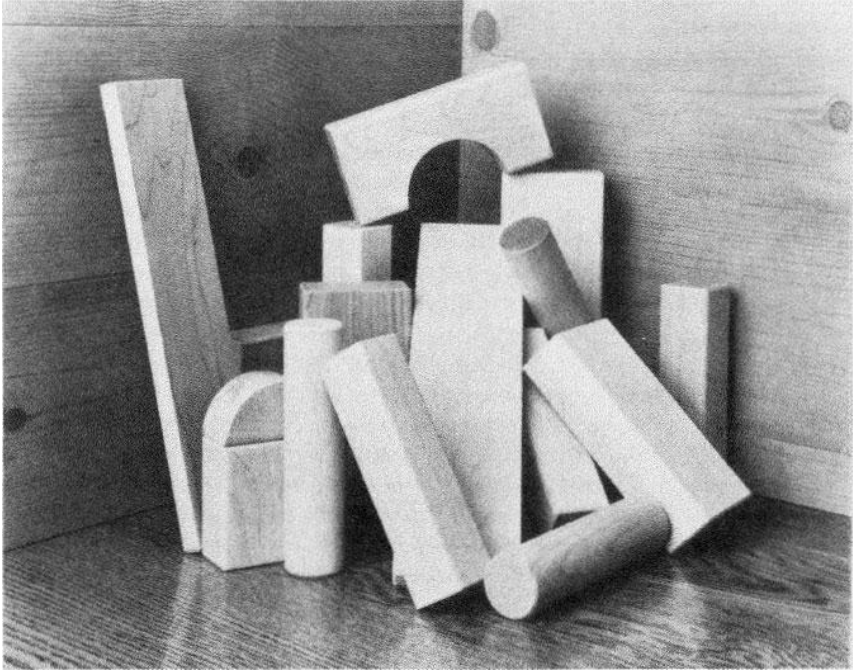
It is important to realize that the flowing perspective structure and the underlying invariant structure are concurrent. They exist at the same time. Although they specify different things, locomotion through a rigid world in

the first instance and the layout of that rigid world in the second instance, they are like the two sides of a coin, for each implies the other. This hypothesis, that optical change can seemingly specify two things at the same time, sounds very strange, as if one cause were having two effects or as if one stimulus were arousing two sensations. But there is nothing illogical about the idea of concurrent specification of two reciprocal things. Such an idea is much needed in psychology.

### The Change between Hidden and Unhidden Surfaces: Covering Edges

We are now prepared to face a fact that has seemed deeply puzzling, a fact that poses the greatest difficulty for all theories of visual perception based on sensations. The layout of the environment includes unprojected (hidden) surfaces at a point of observation as well as projected surfaces, but observers perceive the layout, not just the projected surfaces. Things are seen *in the round* and one thing is seen *in front* of another. How can this be? Information must be available for the whole layout, not just for its façades, for the covered surfaces as well as the covering surfaces. What is this information? Presumably it becomes evident over time, with changes of the array. I will argue that the information is implicit in the *edges that separate* the surfaces or, rather, in the optical specification of these edges. I am suggesting that if covering edges are specified, both the covered and the covering surfaces are also specified.

To suggest that an observer can see surfaces that are unseen is, of course, a paradox. I do not mean that. I am not saying that one can see the unseen, and I am suspicious of visionaries who claim that they can. A vast amount of mystification in the history of human thought has arisen from this paradox. The suggestion is that one can perceive surfaces that are temporarily *out of sight*, and what it is to be out of sight will be carefully defined. The important fact is that they come into sight and go out of sight as the observer moves, first in one direction and then in the opposite direction. If locomotion is reversible, as it is, whatever goes out of sight as the observer travels comes into sight as the observer returns and conversely. The generality of this principle has never been realized; it applies to the shortest locomotions, in centimeters, as well as to the longest, in kilometers. But it has not been elaborated. I will call it the *principle of reversible occlusion*. The theory of the cues for depth perception includes one cue called “movement parallax” and another called “superposition,” both related to the above principle, but these terms are vague and do not even begin to explain what needs to be explained. What we see is not depth as such but *one thing behind another*. The new principle can be made explicit. I will attempt to do so, at some length.



**FIGURE 5.5** Objects seen in the round and behind other objects.

Do you perceive covered surfaces as well as covering surfaces in this photograph? (Photo by Jim Scherer.)

### ***Projected and Unprojected Surfaces***

There are many commonsense words that refer to the fact of covered and uncovered things. Objects and surfaces are said to be hidden or unhidden, screened or unscreened, concealed or revealed, undisclosed or disclosed. We might borrow a technical word in astronomy, *occultation*, but it means primarily the shutting off of the light from a celestial source, as in an eclipse. We need a word for the cutting off of a visual solid angle, not of light rays. I have chosen the word *occlusion* for it. An occluded surface is one that is out of sight or hidden from view. An occluding edge is the edge of an occluding surface. The term was first introduced in a paper by J. J. Gibson, G. A. Kaplan, H. N. Reynolds, and K. Wheeler (1969) on the various ways in which a thing can pass between the state of being visible and the state of being invisible. The experiment will be described in Chapter 11.

Occlusion arises because of two facts about the environment, both described in Chapter 2. First, surfaces are generally opaque; and second, the basic environment, the earth, is generally cluttered. As to the first, if surfaces were as trans-

parent as air, they would not reflect light at all and there would be no use for vision. Most substances are nontransmitting (they reflect and absorb instead), and therefore light is reflected back from the surface. A few substances are partially transmitting or “translucent,” and hence a sheet of such a substance will transmit part of the radiant light but will not transmit the structure of the ambient array; it will let through photons but not visual solid angles. There can be an obstructing of the *view* without obstructing of the *light*, although an obstructing of the light will of course also obstruct the view. If we add the fact that surfaces are also generally textured, the facts of opaque surfaces as contrasted with the surfaces of semitransparent and translucent substances become intelligible.

The second fact is that the environment is generally cluttered. What I called an open environment is seldom or never realized, although it is the only case in which all surfaces are projected and none are unprojected. An open environment has what we call an unobstructed view. But the flat and level earth receding unbroken to a pure linear horizon in a great circle, with a cloudless sky, would be a desolate environment indeed. Perhaps it would not be quite as lifeless as geometrical space, but almost. The *furniture* of the earth, like the furnishings of a room, is what makes it livable. The earth as such affords only standing and walking; the furniture of the earth affords all the rest of behavior. The main items of the clutter (following the terminology adopted in Chapter 3) are *objects*, both attached and detached, *enclosures*, *convexities* such as hills, *concavities* such as holes, and *apertures* such as windows. These features of surface layout give rise to occluding surfaces or, more exactly, to the *separation* of occluding and occluded surfaces.

A surface is *projected* at a point of observation if it has a visual solid angle in the ambient optic array; it is *unprojected* if it does not. A projected surface may become unprotected in at least three ways—if its solid angle is diminished to a point, if the solid angle is compressed to a line, or if the solid angle is wiped out. In the first case we say that the surface is too far away, in the second that it is turned so as not to face the point of observation, in the third that the view is obstructed. The second case, that of facing toward or away, is instructive. A wall or a sheet of paper has two “faces” but only one can *face* a fixed point. The relation between the occluding and occluded surfaces is given by the relation of each to the point; the relation is not merely geometrical but also optical. The relation is designated when we distinguish between the *near* side and the *far* side of an object. (It is not, however, well expressed by the terms *front* and *back*, since they are ambiguous. They can refer to such surfaces as the front and the back of a house or the front and the back of a head. Terms can be borrowed from ordinary language only with discretion!)

### ***Going Out of and Coming Into Sight***

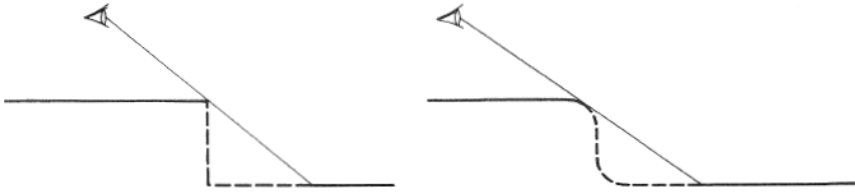
A point of observation is to be thought of as moving through the medium to and fro, back and forth, often along old paths but sometimes along new ones.

Displacements of this position are reversible and are reversed as its occupier comes and goes, even as she slightly shifts her posture. Any face or facet, any surface of the layout, that is progressively hidden during a displacement is progressively unhidden during its reversal. Going out of sight is the inverse of coming into sight. Hence, occluding and occluded surfaces interchange. The occluding ones *change into* the occluded ones and vice versa, not by changing from one entity to another but by a special transition.

The terms *disappearance* and its opposite, *appearance*, should not be used for this transition. They have slippery meanings, like *visible* and *invisible*. For a surface may disappear by going out of existence as well as by going out of sight, and the two cases are profoundly different. A surface that disappears because it is no longer projected to *any* point of observation, because it has evaporated, for example, should not be confused with a surface that disappears because it is no longer projected to a fixed point of observation. The latter can be seen from another position; the former cannot be seen from any position. Failure to distinguish these meanings of *disappear* is common; it encourages careless observation and vague beliefs in ghosts, or in the reality of the “unseen.” To *disappear* can also refer to a surface that continues to exist but is no longer projected to any point of observation because of darkness. Or we might speak of something disappearing “in the distance,” referring to a surface barely projected to a point of observation because its visual solid angle has diminished to a limit. These modes of so-called disappearance are quite radically different. The differences between (1) a surface that ceases to exist, (2) a surface that is no longer illuminated, (3) a surface that lies on the horizon, and (4) a surface that is occluded are described in a paper by Gibson, Kaplan, Reynolds, and Wheeler (1969) and are illustrated in a motion picture film (Gibson, 1968a). An experimental study of the perception of occlusion using motion picture displays has been reported by Kaplan (1969).

### ***The Loci of Occlusion: Occluding Edges***

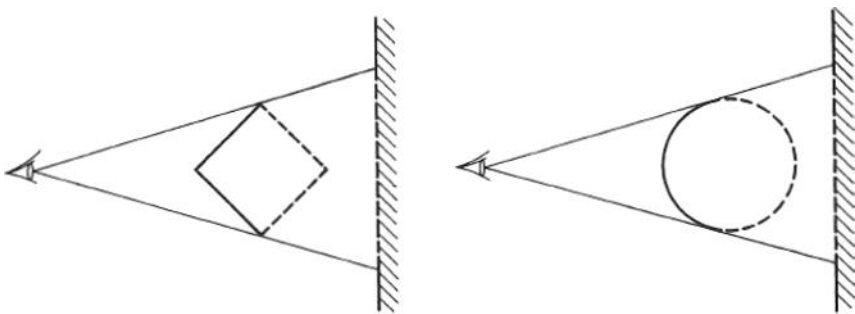
We must now distinguish an edge that is simply the junction of two surfaces from an edge that causes one surface to hide another, an *occluding edge*. In the proposed terminology of layout in Chapter 3, I defined an *edge* as the apex of a convex dihedral (as distinguished from a *corner*, which is the apex of a concave dihedral). But an occluding edge is a dihedral where only one of the surfaces is projected to the point of observation—an *apical* occluding edge. I also defined a *curved convexity* (as distinguished from a *curved concavity*), and another kind of occluding edge is the *brow* of this convexity, that is, the line of tangency of the envelope of the visual solid angle—a *curved* occluding edge. The apical occluding edge is “sharp,” and the curved occluding edge is “rounded.” The two are illustrated in Figure 5.6. The latter slides along the surface as the point of observation moves, but the former does not. Note that an occluding edge always requires a convexity of some sort, a protrusion of the substance into the medium.



**FIGURE 5.6** The sharp occluding edge and the rounded occluding edge at a fixed point of observation.

The hidden portions of the surface layout are indicated by dotted lines.

These two kinds of occluding edges are found in the eels of corridors, the brinks of cliffs, the brows of hills, and the near sides of holes in the ground. One face or facet or part of the layout hides another to which it may be connected and which it may adjoin. This is different from what I called a detached object, by which I mean the movable or moving object having a topologically closed surface with substance inside and medium outside. The detached object produces a visual solid angle in the optic array, as noted by Euclid and Ptolemy, and yields a closed-contour figure in the visual field, as described by Edgar Rubin and celebrated by the gestalt psychologists under the name of the “figure-ground phenomenon.” Occluding edges are a special case, because not only does the near side of the object hide the far side but the object covers a sector of the surface behind it, the ground, for example. The occluding edges may be apical, as when the object is a polyhedron, or the locus of the tangent of the envelope of the solid angle to the surface, as when the object is curved. These are illustrated in Figure 5.7, where both the hiding of the far side and the covering of the background are shown. The object is itself rounded or



**FIGURE 5.7** Both the far side of an object and the background of the object are hidden by its occluding edges.

Two detached objects are shown, one with sharp occluding edges and the other with rounded occluding edges.

solid, and it is superposed on the ground, which is also continuous behind the object. These two kinds of occlusion may be treated separately.

### ***Self-occlusion and Superposition***

An object, in the present terminology, is both voluminous and superposed. It exists in volume and it may lie in front of another surface, or another object. In short, an object always occludes itself and generally also occludes something else. The effect of a moving point of observation is different in the two cases.

Projected and unprojected surfaces interchange as the point of observation moves, but the interchange between parts of the object is not like that between parts of the background. There is an interchange between *opposite faces* of the object but an interchange of *adjacent areas* of the surface behind the object. For the object, the near side turns into the far side and vice versa, whereas for the background an uncovered area becomes covered and vice versa. The change of optical structure in the former case is by way of perspective transformation, whereas the disturbance of optical structure in the latter case is more radical, a “kinetic disruption” being involved.

In Figure 5.7, as the point of observation moves each face of the facade of the polyhedron undergoes transformation, for example, from trapezoid to square to trapezoid. Ultimately, when the face is maximally foreshortened, it is what we call “edge on,” that is, it becomes an occluding edge. The near face turns into a far face by way of the edge. While this is happening at one edge, the other edge is revealing a previously hidden face. A far face turns into a near face. The two occluding edges in the diagram are perfectly reciprocal; while one is converting near into far, the other is converting far into near. The width of the polyhedron goes into depth, and the depth comes back into width. Width and depth are thus interchangeable.

Similarly, one could describe the transformation of each facet of the textured surface of the curved object. If the object is a sphere, the circular occluding edge (the *outline*, in pictorial terminology) does not transform, but the optical structure within it does. At one edge the texture is progressively turning from projected into unprojected, from near into far, while at the other edge the texture is progressively turning from unprojected into projected, from far into near. The transition occurs at the limit of the slant transformation, the ultimate of perspective foreshortening, but actually the optical texture reaches and goes beyond this purported limit. It has to go beyond it because it comes from beyond that limit at the other occluding edge.

### ***Superposition***

Now consider the separated background behind the objects in Figure 5.7, the fact of superposition as distinguished from the fact of solidity. As the point of



observation moves, the envelope of the visual solid angle sweeps across the surface. The leading edge progressively covers the texture of the surface, while the trailing edge progressively uncovers it. I have suggested metaphorically that the texture is “wiped out” and “unwiped” at the lateral borders of the figure (Gibson, 1966*b*, pp. 199 ff.). This was inspired by the metaphors used by A. Michotte in describing experiments on what he called the “tunnel effect” (Michotte, Thinès, and Crabbé, 1964). A somewhat more exact description of this optical change will be given below. But note that if the texture that is progressively covered has the same structure as the texture that is progressively uncovered the unity of the surface is well specified.

The metaphor of “wiping” is inexact. A better description of the optical transition was given by Gibson, Kaplan, Reynolds, and Wheeler (1969), and it was also described by Kaplan (1969) as a “kinetic disruption.” There is a disturbance of the structure of the array that is not a transformation, not even a transformation that passes through its vanishing limit, but a breaking of its adjacent order. More exactly, there is either a progressive decrementing of components of structure, called *deletion*, or its opposite, a progressive incrementing of components of structure, called *accretion*. An edge that is covering the background deletes from the array; an edge that is uncovering the background accretes to it. There is no such disruption for the surface that is covering or uncovering, only for the surface that is being covered or uncovered. And nondisruption, I suggest, is a kind of invariance.

### ***The Information to Specify the Continuation of Surfaces***

A surface always “bends under” an occluding edge, and another surface generally “extends behind” it. These surfaces are connected or continuous. Is there information in a changing optic array to specify the connectedness or continuity?

Here is a tentative hypothesis for the continuous object surface:

Whenever a perspective transformation of form or texture in the optic array goes to its limit and when a series of forms or textures are progressively foreshortened to this limit, a continuation of the surface of an object is specified at an occluding edge. This is the formula for going out of sight; the formula is reversed for coming into sight.

Here is a tentative hypothesis for the continuous background surface:

Whenever there occurs a regular disturbance of the persistence of forms and textures in the optic array such that they are progressively deleted at a contour, the continuation of the surface of a ground is specified at an occluding edge. This is for going out of sight; substituting accretion for deletion gives the formula for coming into sight.



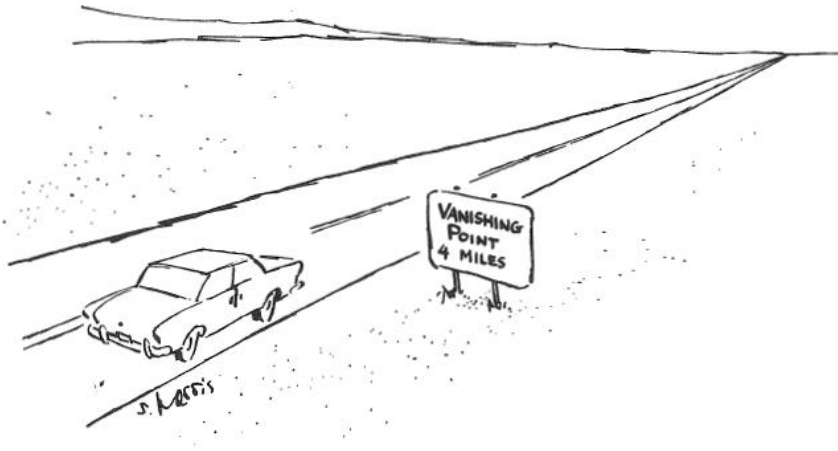
These two hypotheses make no assertions about perception, only about the information that is normally available for perception. They do not refer to *space*, or to the *third dimension*, or to *depth*, or to *distance*. Nothing is said about forms or patterns in two dimensions. But they suggest a radically new basis for explaining the perception of solid superposed objects, a new theory based not on cues or clues or signs but on the direct pickup of solidity and superposition. An object is in fact voluminous; a background is in fact continuous. A picture or an image of an object is irrelevant to the question of how it is perceived.

The assumption for centuries has been that the sensory basis for the perception of an object is the outline form of its image on the retina. Object perception can only be based on form perception. First the silhouette is detected and then the depth is added, presumably because of past experiences with the cues for depth. But the fact is that the progressive foreshortening of the face of an object is perceived as the turning of the object, which is precisely what the transformation specifies, and is never perceived as a change of form, which ought to be seen if the traditional assumption is correct—that the silhouette is detected and then the depth is added.

The two hypotheses stated above depend on a changing optic array, and so far the only cause of such change that has been considered is the moving point of observation. The reader will have noted that a moving *object* will also bring about the same kinds of disturbance in the structure of the array that have been described above. A moving object in the world is an event, however, not a form of locomotion, and the information for the perception of events will be treated in Chapter 6.

### ***The Case of Very Distant Surfaces***

It is interesting to compare the occluding edges of objects and other convexities on the surface of the earth with the *horizon* of the earth, the great circle dividing the ambient array into two hemispheres. It is the limit of perspective minification for terrestrial surfaces, just as the edge-on line is the limit of perspective compression (foreshortening) for a terrestrial surface. Objects such as railroad trains on the Great Plains and ships on the ocean are said to vanish in the distance as they move away from a fixed point of observation. The line of the horizon in the technology of pictorial perspective is said to be the locus of vanishing points for the size of earth-forms and for the convergence of parallel edges on the earth. The railroad train “vanishes” at the same optical point where the railroad tracks “meet” in the distance. The horizon is therefore analogous to an occluding edge *in* being one of the loci at which things go out of and come into sight. But going out of sight in the distance is very different from going out of sight at a sharp or a rounded edge nearby. The horizon of the earth, therefore, is not an occluding edge for any terrestrial object or earth-form. It does not in fact *look* like an occluding edge. It could only be visualized as an occluding edge for the lands



**FIGURE 5.8** Cartoon. (Drawing by S. Harris; © 1975 The New Yorker Magazine, Inc.)

and seas beyond the horizon if the seemingly flat earth were conceived as curved and if the environment were thought of as a globe too vast to see.

It has long been a puzzle to human observers, however, that the horizon is in fact visibly an occluding edge for *celestial* objects such as the sun and the moon. Such objects undergo progressive deletion at a contour, as at sunset, and undergo progressive accretion at the same contour, as at moonrise. This is in accordance with the second hypothesis above. The object is obviously beyond the horizon, more distant than the visible limit of earthly distance, and yet there is some information for its being a solid surface. This conflicting information explains, I think, the apparently enormous size of the sun and the moon at the horizon. It also explains many of the ideas of pre-Copernican astronomy about heavenly bodies. We should realize that the terrestrial environment was the only environment that people could be *sure* of before Copernicus—the only environment that could be perceived directly. Terrestrial objects and surfaces had affordances for behavior, but celestial objects did not. More will be said about the perception of objects on earth as distinguished from objects in the sky in Part III.

### **Summary: The Optics of Occlusion**

1. In the ideal case of a terrestrial earth without clutter, all parts of the surface are projected to all points of observation. But such an open environment would hardly afford life.
2. In the case of an earth with furniture, with a layout of opaque surfaces on a substratum, some parts of the layout are projected to any given fixed point of observation and the remaining parts are unprojected to that point.

3. The optically uncovered surface of an object is always *separated* from the optically covered surface at the occluding edge. At the same time, it is always *connected* with the optically covered surface at the occluding edge.

4. The continuation of the far side with the near side is specified by the *reversibility* of occlusion.

5. Any surface of the layout that is hidden at a given fixed point of observation will be unhidden at some other fixed point.

6. Hidden and unhidden surfaces *interchange*. Whatever is revealed by a given movement is concealed by the reverse of that movement. This principle of reversible occlusion holds true for both movements of the point of observation and motions of detached objects.

7. We can now observe that the *separation* between hidden and unhidden surfaces at occluding edges is best specified by the *perspective* structure of an array, whereas the *connection* between hidden and unhidden surfaces at edges is specified by the underlying *invariant* structure. Hence, probably, a pause in locomotion calls attention to the difference between the hidden and the unhidden, whereas locomotion makes evident the continuousness between the hidden and the unhidden.

The seeming paradox of the perceiving or apprehending of hidden surfaces will be treated further in Chapter 11.

## How is Ambient Light Structured? A Theory

Let us return to the question of how ambient light is given its invariant structure, the question asked at the beginning of this chapter but not answered except in a preliminary way. Ambient light can only be structured by something that surrounds the point of observation, that is, by an environment. It is not structured by an empty medium of air or by a fog-filled medium. There have to be surfaces—both those that emit light and those that reflect light. Only because ambient light is structured by the substantial environment can it contain information about it.

So far it has been emphasized that ambient light is made to constitute an array by a single feature of these surfaces, their layout. But just *how* does the layout structure the light? The answer is not simple. It involves the puzzling complexities of light and shade. Moreover, the layout of surfaces is not the only cause of the structuring of light; the *conglomeration* of surfaces makes a contribution, that is, the fact that the environment is multicolored. The different surfaces of the layout are made of different substances with different reflectances. Both lighted or shaded surfaces and black or white surfaces make their separate contributions to the invariant structure of ambient light. And how light-or-shade can be perceived separately from black-or-white has long been a puzzling problem for any theory of visual sense perception.

I tried to formulate a theory of the structuring of ambient light in my last book (Gibson, 1966*b*), asserting that three causes existed, the layout of surfaces, the pigmentation of surfaces, and the shadowing of surfaces (pp. 208–216). But the third of these causes is not cognate with the other two, and the interaction between them was not clearly explained. The theory was static. Here, I shall formulate a theory of the sources of *invariant* optical structure in relation to the sources of *variation* in optical structure. What is clear to me now that was not clear before is that structure as such, frozen structure, is a myth, or at least a limiting case. Invariants of structure do not exist except in relation to variants.

### ***The Sources of Invariant Optical Structure***

The main invariants of the terrestrial environment, its persisting features, are the layout of its surfaces and the reflectances of these surfaces. The layout tends to persist because most of the substances are sufficiently solid that their surfaces are rigid and resist deformation. The reflectances tend to persist because most of the substances are chemically inert at their interfaces with the air, and their surfaces keep the same composition, that is, the same colors, both achromatic and chromatic. Actually, at the level of microlayout (texture) and microcomposition (conglomeration), layout and reflectances merge. Or, to put it differently, the layout texture and the pigment texture become inseparable.

Note once more that an emphasis on the geometry of surfaces is abstract and oversimplified. The faces of the world are *not* made of some amorphous, colorless, ghostly substance, as geometry would lead us to believe, but are made of mud or sand, wood or metal, fur or feathers, skin or fabric. The faces of the world are colorful as well as geometrical. And what they afford depends on their substance as well as their shape.

### ***The Sources of Variant Optical Structure***

There are two regular and recurrent sources of changing structure in the ambient light (apart from local events, which will be considered in the next chapter). First, there are the changes caused by a moving point of observation, and second, there are the changes caused by a moving source of illumination, usually the sun. Many pages have been devoted to the former, and we must now consider the latter. The motion of the sun across the sky from sunrise to sunset has been for countless millions of years a basic regularity of nature. It is a fact of ecological optics and a condition of the evolution of eyes in terrestrial animals. But its importance for the theory of vision has not been fully recognized.

The puzzling complexities of light and shade cannot be understood without taking into account the fact of a *moving* source of illumination. For whenever the source of light moves, the direction of the light falling on the surfaces of the

world is altered and the shadows themselves move. The layout and coloration of surfaces persist, but the lightedness and shadedness of these surfaces do not. It is not just that the optic array is different at noon with high illumination from what it is at twilight with low illumination; it is that the optic array has a different structure in the afternoon than it has in the morning.

### Variants and Invariants with a Moving Source of Illumination

Just how does pure layout structure the ambient light? It is easy to understand how a mosaic of black and white substances would structure the ambient light but not how a pure layout would do so. For in this case the structuring would have to be achieved wholly by differential illumination, by light and shade. There are two principles of light and shade under natural conditions that seem to be clear: the direction of the prevailing illumination and the progressive weakening of illumination with multiple reflection.

The illumination on a surface comes from the sun, the sky, and other surfaces that face the surface in question. A surface that faces the sun is illuminated “directly,” a surface that faces away from the sun but still faces the sky is illuminated less directly, and a surface within a semienclosure that faces only other surfaces is illuminated still less directly. The more the light has reverberated, the more of it is absorbed and the dimmer it becomes. Hence it is that surfaces far from the mouth of a cave are more weakly illuminated than those near the mouth. But within any airspace, any concavity of the terrain or any semienclosure, there is a direction of the *prevailing* illumination, that is, a direction from which more light comes than from any other.

The illumination of any face of the layout relative to adjacent faces depends on its inclination to the prevailing illumination. Crudely speaking, the surface that “faces the light” gets more than its neighbor. More exactly, a surface perpendicular to the prevailing illumination gets the most, a surface inclined to it gets less, a surface parallel to it gets still less, and a surface inclined *away* from it gets the least. The pairs of terms *lighted* and *shadowed* or *in light* and *in shadow* should not be taken as dichotomies, for there are all gradations of relative light and shade. These two principles of the direction and the amount of illumination are an attempt to distill a certain ecological simplicity from the enormous complexities of analytical physical optics and the muddled practice of illumination engineering.

A wrinkled surface of the same substance evidently structures the ambient light by virtue of two facts: there is always a prevailing direction of illumination, and consequently the slopes facing in this direction throw back more energy than the slopes not facing in this direction. A *flat* surface of *different* substances structures the ambient light by virtue of the simple fact that the parts of high reflectance throw back more energy than the parts of low reflectance.

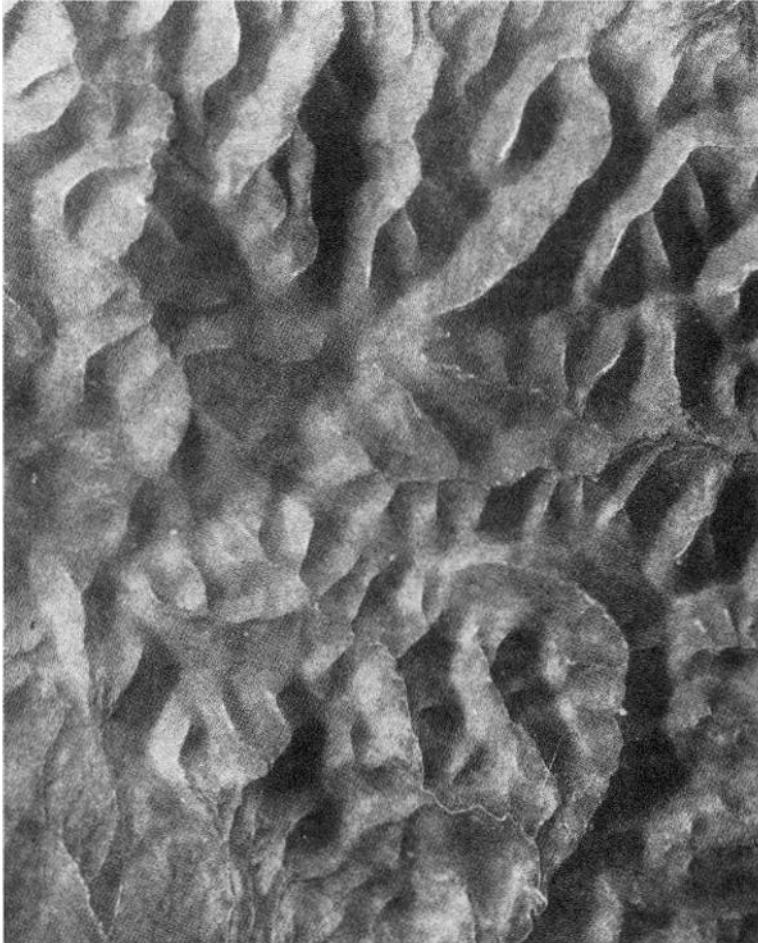
Figure 5.9 shows an array from a wrinkled layout of terrestrial surfaces, actually an aerial photograph of barren hills and valleys. The bare earth of this desert has everywhere the same reflectance. The top of the photograph is to the north of the terrain. The picture was taken in the morning, and the sun is in the east. Some of the slopes face east, and some face west; the former are lighted and the latter shaded. It can be observed that various inclinations of these surfaces to the direction of the prevailing illumination determine various relative intensities in the array; the more a surface departs from the perpendicular to this direction, the darker is the corresponding patch in the optic array.

Now consider what happens as the sun moves across the sky. All those surfaces that were lighted in the morning will be shaded in the afternoon, and all those that were shaded in the morning will be lighted in the afternoon. There is a continual, if slow, process of change from lighted to shaded on certain slopes of the layout and the reverse change on certain other slopes. These slopes are related by orientation. Two faces of any convexity are related in this way, as are two faces of any concavity. A ridge can be said to consist of two opposite slopes, and so can a valley. The reciprocity of light and shade on such surfaces might be described by saying that the lightness and the shadedness *exchange places*. The underlying surfaces do not interchange of course, and their colors, if any, do not interchange. They are persistent, but the illumination is variable in this special reciprocal way.

In the optic array, presumably, there is an underlying invariant structure to specify the edges and corners of the layout and the colors of the surfaces, and at the same time there is a changing structure to specify the temporary direction of the prevailing illumination. Some components of the array never exchange places—that is, they are never permuted—whereas other components of the array do. The former specify a solid surface; the latter specify insubstantial shadows only. The surface and its color are described as opaque; the shadow is described as transparent.

The decreasing of illumination on one slope and the increasing of illumination on an adjacent slope as the sun moves are analogous to the foreshortening of one slope along with the inverse foreshortening of an adjacent slope as the point of observation moves. I suggest that the true relative colors of the adjacent surfaces emerge as the lighting changes, just as the true relative shapes of the adjacent surfaces emerge as the perspective changes. The perspectives of the convexities and concavities of Figure 5.9 are variant with locomotion; the shadows of these convexities and concavities are variant with time of day; the constant properties of these surfaces underlie the changing perspectives and the changing shadows and are specified by invariants in the optic array.

It is true that the travel of the sun across the sky is very slow and that the correlated interchange of the light and the shade on surfaces is a very gradual fluctuation. Neither is as obvious as the motion perspective caused by loco-



**FIGURE 5.9** Hills and valleys on the surface of the barren earth.

The hills in this aerial photograph, the convexities or protuberances, can be compared to the “humps” shown in Figure 5.1

motion. But the fact is that shifting shadows and a moving sun are regularities of ecological optics whether or not they are ever noticed by any animal. They have set the conditions for the perception of the terrain by terrestrial animals since life emerged from the sea. They make certain optical information available. And, although shifting shadows and a moving sun are too slow to be noticed in daylight, a moving source of illumination and the resultant shadows become more obvious at night. One has only to carry a light from place to place in a cluttered environment in order to notice the radical shifts in the pattern of the optic array caused by visibly moving shadows. And yet, of course, the layout



of surfaces and their relative coloration is visible underneath the moving shadows.

How the differential colors of surfaces are specified in the optic array separately from the differential illumination of surfaces is, of course, a great puzzle. The difference between black and white is never confused with the difference between lighted and shadowed, at least not in a natural environment as distinguished from a controlled laboratory display. There are many theories of this so-called constancy of colors in perception, but none of them is convincing. A new approach to the problem is suggested by the above considerations.

From an ecological point of view, the color of a surface is relative to the colors of adjacent surfaces; it is not an absolute color. Its reflectance ratio is specified only in relation to other reflectance ratios of the layout. For the natural environment is an *aggregate* of substances. Even a surface is sometimes a *conglomerate* of substances. This means that a range of black, gray, and white surfaces and a range of chromatically colored surfaces will be projected as solid angles in a normal optic array. The colors are not seen separately, as stimuli, but together, as an arrangement. And this range of colors provides an invariant structure that underlies *both* the changing shadow structure with a moving sun and the changing perspective structure with a moving observer. The edges and corners, the convexities and concavities, are thus specified as multicolored surfaces, not as mere slopes; as speckled or grained or piebald or whatever, not as ghostly gray shapes.

The experimental discoveries of E. H. Land (1959) concerning color perception with what he calls a “complete image” as distinguished from color perception with controlled patches of radiation in a laboratory are to be understood in the above way, I believe.

### Ripples and Waves on Water: A Special Case

It is interesting and revealing to compare the optical information for a solid wrinkled surface as shown in Figure 5.9 and the information for a liquid wavy surface, which the reader will have to visualize. Both consist of convexities and concavities, but they are motionless on the solid surface and moving on the liquid surface. In both cases the convexities are lighted on one slope and shadowed on the other. In both cases the surface is all of the same color or reflectance. The difference between the two arrays is to be found chiefly in the two *forms of fluctuation* of light and shade. In the terrestrial array, light and shade exchange places slowly in one direction; they do not oscillate. In the aquatic array, light and shade interchange rapidly in both directions; they oscillate. In fact, when the sun is out and the ripples act as mirrors, the reflection of the sun can be said to flicker or flash on and off. This specific form of fluctuation is characteristic of a water surface.



## Summary

When ambient light at a point of observation is structured it is an ambient optic array. The point of observation may be stationary or moving, relative to the persisting environment. The point of observation may be unoccupied or occupied by an observer.

The structure of an ambient array can be described in terms of visual solid angles with a common apex at the point of observation. They are angles of intercept, that is, they are determined by the persisting environment. And they are nested, like the components of the environment itself.

The concept of the visual solid angle comes from natural perspective, which is the same as ancient optics. No two such visual angles are identical. The solid angles of an array change as the point of observation moves, that is, the perspective structure changes. Underlying the perspective structure, however, is an invariant structure that does *not* change. Similarly, the solid angles of an array change as the sun in the sky moves, that is, the shadow structure changes. But there are also invariants that underlie the changing shadows.

The moving observer and the moving sun are conditions under which terrestrial vision has evolved for millions of years. But the invariant principle of reversible occlusion holds for the moving observer, and a similar principle of reversible illumination holds for the moving sun. Whatever goes out of sight will come into sight, and whatever is lighted will be shaded.

# 8

## THE THEORY OF AFFORDANCES

I have described the environment as the surfaces that separate substances from the medium in which the animals live. But I have also described what the environment *affords* animals, mentioning the terrain, shelters, water, fire, objects, tools, other animals, and human displays. How do we go from surfaces to affordances? And if there is information in light for the perception of surfaces, is there information for the perception of what they afford? Perhaps the composition and layout of surfaces *constitute* what they afford. If so, to perceive them is to perceive what they afford. This is a radical hypothesis, for it implies that the “values” and “meanings” of things in the environment can be directly perceived. Moreover, it would explain the sense in which values and meanings are external to the perceiver.

The *affordances* of the environment are what it *offers* the animal, what it *provides* or *furnishes*, either for good or ill. The verb *to afford* is found in the dictionary, but the noun *affordance* is not. I have made it up. I mean by it something that refers to both the environment and the animal in a way that no existing term does. It implies the complementarity of the animal and the environment. The antecedents of the term and the history of the concept will be treated later; for the present, let us consider examples of an affordance.

If a terrestrial surface is nearly horizontal (instead of slanted), nearly flat (instead of convex or concave), and sufficiently extended (relative to the size of the animal) and if its substance is rigid (relative to the weight of the animal), then the surface *affords support*. It is a surface of support, and we call it a substratum, ground, or floor. It is stand-on-able, permitting an upright posture for quadrupeds and bipeds. It is therefore walk-on-able and run-over-able. It is not sink-into-able like a surface of water or a swamp, that is, not for heavy terrestrial animals. Support for water bugs is different.

Note that the four properties listed—horizontal, flat, extended, and rigid—would be *physical* properties of a surface if they were measured with the scales and standard units used in physics. As an affordance of support for a species of animal, however, they have to be measured *relative to the animal*. They are unique for that animal. They are not just abstract physical properties. They have unity relative to the posture and behavior of the animal being considered. So an affordance cannot be measured as we measure in physics.

Terrestrial surfaces, of course, are also climb-on-able or fall-off-able or get-underneath-able or bump-into-able relative to the animal. Different layouts afford different behaviors for different animals, and different mechanical encounters. The human species in some cultures has the habit of sitting as distinguished from kneeling or squatting. If a surface of support with the four properties is also knee-high above the ground, it affords sitting on. We call it a *seat* in general, or a stool, bench, chair, and so on, in particular. It may be natural like a ledge or artificial like a couch. It may have various shapes, as long as its functional layout is that of a seat. The color and texture of the surface are irrelevant. Knee-high for a child is not the same as knee-high for an adult, so the affordance is relative to the size of the individual. But if a surface is horizontal, flat, extended, rigid, and knee-high relative to a perceiver, it can in fact be sat upon. If it can be discriminated as having just these properties, it should *look* sit-on-able. If it does, the affordance is perceived visually. If the surface properties are seen relative to the body surfaces, the self, they constitute a seat and have meaning.

There could be other examples. The different substances of the environment have different affordances for nutrition and for manufacture. The different objects of the environment have different affordances for manipulation. The other animals afford, above all, a rich and complex set of interactions, sexual, predatory, nurturing, fighting, playing, cooperating, and communicating. What other persons afford, comprises the whole realm of social significance for human beings. We pay the closest attention to the optical and acoustic information that specifies what the other person is, invites, threatens, and does.

## The Niches of the Environment

Ecologists have the concept of a *niche*. A species of animal is said to utilize or occupy a certain niche in the environment. This is not quite the same as the *habitat* of the species; a niche refers more to *how* an animal lives than to *where* it lives. I suggest that a niche is a set of affordances.

The natural environment offers many ways of life, and different animals have different ways of life. The niche implies a kind of animal, and the animal implies a kind of niche. Note the complementarity of the two. But note also that the environment as a whole with its unlimited possibilities existed prior to animals. The physical, chemical, meteorological, and geological conditions of

the surface of the earth and the pre-existence of plant life are what make animal life possible. They had to be invariant for animals to evolve.

There are all kinds of nutrients in the world and all sorts of ways of getting food; all sorts of shelters or hiding places, such as holes, crevices, and caves; all sorts of materials for *making* shelters, nests, mounds, huts; all kinds of locomotion that the environment makes possible, such as swimming, crawling, walking, climbing, flying. These offerings have been taken advantage of; the niches have been occupied. But, for all we know, there may be many offerings of the environment that have *not* been taken advantage of, that is, niches not yet occupied.

In architecture a niche is a place that is suitable for a piece of statuary, a place into which the object fits. In ecology a niche is a setting of environmental features that are suitable for an animal, into which it fits metaphorically.

An important fact about the affordances of the environment is that they are in a sense objective, real, and physical, unlike values and meanings, which are often supposed to be subjective, phenomenal, and mental. But, actually, an affordance is neither an objective property nor a subjective property; or it is both if you like. 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.

The niche for a certain species should not be confused with what some animal psychologists have called the *phenomenal environment* of the species. This can be taken erroneously to be the “private world” in which the species is supposed to live, the “subjective world,” or the world of “consciousness.” The behavior of observers depends on their perception of the environment, surely enough, but this does not mean that their behavior depends on a so-called private or subjective or conscious environment. The organism depends on its environment for its life, but the environment does not depend on the organism for its existence.

## Man’s Alteration of the Natural Environment

In the last few thousand years, as everybody now realizes, the very face of the earth has been modified by man. The layout of surfaces has been changed, by cutting, clearing, leveling, paving, and building. Natural deserts and mountains, swamps and rivers, forests and plains still exist, but they are being encroached upon and reshaped by man-made layouts. Moreover, the *substances* of the environment have been partly converted from the natural materials of the earth into various kinds of artificial materials such as bronze, iron, concrete, and bread. Even the *medium* of the environment—the air for us and the water for fish—is becoming slowly altered despite the restorative cycles that yielded a steady state for millions of years prior to man.

Why has man changed the shapes and substances of his environment? To change what it affords him. He has made more available what benefits him and less pressing what injures him. In making life easier for himself, of course, he has made life harder for most of the other animals. Over the millennia, he has made it easier for himself to get food, easier to keep warm, easier to see at night, easier to get about, and easier to train his offspring.

This is not a *new* environment—an artificial environment distinct from the natural environment—but the same old environment modified by man. It is a mistake to separate the natural from the artificial as if there were two environments; artifacts have to be manufactured from natural substances. It is also a mistake to separate the cultural environment from the natural environment, as if there were a world of mental products distinct from the world of material products. There is only one world, however diverse, and all animals live in it, although we human animals have altered it to suit ourselves. We have done so wastefully, thoughtlessly, and, if we do not mend our ways, fatally.

The fundamentals of the environment—the substances, the medium, and the surfaces—are the same for all animals. No matter how powerful men become they are not going to alter the fact of earth, air, and water—the lithosphere, the atmosphere, and the hydrosphere, together with the interfaces that separate them. For terrestrial animals like us, the earth and the sky are a basic structure on which all lesser structures depend. We cannot change it. We all fit into the substructures of the environment in our various ways, for we were all, in fact, formed by them. We were created by the world we live in.

## Some Affordances of the Terrestrial Environment

Let us consider the affordances of the medium, of substances, of surfaces and their layout, of objects, of animals and persons, and finally a case of special interest for ecological optics, the affording of concealment by the occluding edges of the environment (Chapter 5).

### *The Medium*

Air affords breathing, more exactly, respiration. It also affords unimpeded locomotion relative to the ground, which affords support. When illuminated and fog-free, it affords visual perception. It also affords the perception of vibratory events by means of sound fields and the perception of volatile sources by means of odor fields. The airspaces between obstacles and objects are the paths and the places where behavior occurs.

The optical information to specify air when it is clear and transparent is not obvious. The problem came up in Chapter 4, and the experimental evidence about the seeing of “nothing” will be described in the next chapter.

## *The Substances*

Water is more substantial than air and always has a surface with air. It does not afford respiration for us. It affords drinking. Being fluid, it affords pouring from a container. Being a solvent, it affords washing and bathing. Its surface does not afford support for large animals with dense tissues. The optical information for water is well specified by the characteristics of its surface, especially the unique fluctuations caused by rippling (Chapter 5).

Solid substances, more substantial than water, have characteristic surfaces (Chapter 2). Depending on the animal species, some afford nutrition and some do not. A few are toxic. Fruits and berries, for example, have more food value when they are ripe, and this is specified by the color of the surface. But the food values of substances are often misperceived.

Solids also afford various kinds of manufacture, depending on the kind of solid state. Some, such as flint, can be chipped; others, such as clay, can be molded; still others recover their original shape after deformation; and some resist deformation strongly. Note that manufacture, as the term implies, was originally a form of manual behavior like manipulation. Things were fabricated *by hand*. To identify the substance in such cases is to perceive what can be done with it, what it is good for, its utility; and the hands are involved.

## *The Surfaces and their Layouts*

I have already said that a horizontal, flat, extended, rigid surface affords support. It permits equilibrium and the maintaining of a posture with respect to gravity, this being a force perpendicular to the surface. The animal does not fall or slide as it would on a steep hillside. Equilibrium and posture are prerequisite to other behaviors, such as locomotion and manipulation. There will be more about this in Chapter 12, and more evidence about the perception of the ground in Chapter 9. The ground is quite literally the *basis* of the behavior of land animals. And it is also the basis of their visual perception, their so-called space perception. Geometry began with the study of the earth as abstracted by Euclid, not with the study of the axes of empty space as abstracted by Descartes. The affording of support and the geometry of a horizontal plane are therefore not in different realms of discourse; they are not as separate as we have supposed.

The flat earth, of course, lies *beneath* the attached and detached objects on it. The earth has “furniture,” or as I have said, it is cluttered. The solid, level, flat surface extends behind the clutter and, in fact, extends all the way out to the horizon. This is not, of course, the earth of Copernicus; it is the earth at the scale of the human animal, and on that scale it is flat, not round. Wherever one goes, the earth is separated from the sky by a horizon that, although it may be hidden by the clutter, is always there. There will be evidence to show that

the horizon can always be seen, in the sense that it can be visualized, and that it can always be felt, in the sense that any surface one touches is experienced in relation to the horizontal plane.

Of course, a horizontal, flat, extended surface that is *nonrigid*, a stream or lake, does not afford support for standing, or for walking and running. There is no footing, as we say. It may afford floating or swimming, but you have to be equipped for that, by nature or by learning.

A *vertical*, flat, extended, and rigid surface such as a wall or a cliff face is a barrier to pedestrian locomotion. Slopes between vertical and horizontal afford walking, if easy, but only climbing, if steep, and in the latter case the surface cannot be flat; there must be “holds” for the hands and feet. Similarly, a slope downward affords falling if steep; the brink of a cliff is a falling-off place. It is dangerous and looks dangerous. The affordance of a certain layout is perceived if the layout is perceived.

Civilized people have altered the steep slopes of their habitat by building stairways so as to afford ascent and descent. What we call the steps afford stepping, up or down, relative to the size of the person’s legs. We are still capable of getting around in an arboreal layout of surfaces, tree branches, and we have ladders that afford this kind of locomotion, but most of us leave that to our children.

A cliff face, a wall, a chasm, and a stream are barriers; they do not afford pedestrian locomotion unless there is a door, a gate, or a bridge. A tree or a rock is an obstacle. Ordinarily, there are paths between obstacles, and these openings are visible. The progress of locomotion is guided by the perception of barriers and obstacles, that is, by the act of steering into the openings and away from the surfaces that afford injury. I have tried to describe the optical information for the control of locomotion (Gibson, 1958), and it will be further elaborated in Chapter 13. The *imminence* of collision with a surface during locomotion is specified in a particularly simple way, by an explosive rate of magnification of the optical texture. This has been called *looming* (e.g., Schiff, 1965). It should not be confused, however, with the magnification of an opening between obstacles, the opening up of a *vista* such as occurs in the approach to a doorway.

## ***The Objects***

The affordances of what we loosely call *objects* are extremely various. It will be recalled that my use of the terms is restricted and that I distinguish between *attached* objects and *detached* objects. We are not dealing with Newtonian objects in space, all of which are detached, but with the furniture of the earth, some items of which are attached to it and cannot be moved without breakage.

Detached objects must be comparable in size to the animal under consideration if they are to afford behavior. But those that are comparable afford an

astonishing variety of behaviors, especially to animals with hands. Objects can be manufactured and manipulated. Some are portable in that they afford lifting and carrying, while others are not. Some are graspable and other not. To be graspable, an object must have opposite surfaces separated by a distance less than the span of the hand. A five-inch cube can be grasped, but a ten-inch cube cannot (Gibson, 1966*b*, p. 119). A large object needs a “handle” to afford grasping. Note that the size of an object that constitutes a graspable size is specified in the optic array. If this is true, it is *not* true that a tactual sensation of size has to become associated with the visual sensation of size in order for the affordance to be perceived.

Sheets, sticks, fibers, containers, clothing, and tools are detached objects that afford manipulation (Chapter 3). Additional examples are given below.

1. An elongated object of moderate size and weight affords wielding. If used to hit or strike, it is a *club* or *hammer*. If used by a chimpanzee behind bars to pull in a banana beyond its reach, it is a sort of *rake*. In either case, it is an extension of the arm. A rigid staff also affords leverage and in that use is a *lever*. A pointed elongated object affords piercing—if large it is a *spear*, if small a *needle* or *awl*.

2. A rigid object with a sharp dihedral angle, an edge, affords cutting and scraping; it is a *knife*. It may be designed for both striking and cutting, and then it is an *axe*.

3. A graspable rigid object of moderate size and weight affords throwing. It may be a *missile* or only an object for play, a *ball*. The launching of missiles by supplementary tools other than the hands alone—the sling, the bow, the catapult, the gun, and so on—is one of the behaviors that makes the human animal a nasty, dangerous species.

4. An elongated elastic object, such as a *fiber*, *thread*, *thong*, or *rope*, affords knotting, binding, lashing, knitting, and weaving. These are kinds of behavior where manipulation leads to manufacture.

5. A hand-held tool of enormous importance is one that, when applied to a surface, leaves traces and thus affords *trace-making*. The tool may be a *stylus*, *brush*, *crayon*, *pen*, or *pencil*, but if it marks the surface it can be used to depict and to write, to represent scenes and to specify words.

We have thousands of names for such objects, and we classify them in many ways: pliers and wrenches are tools; pots and pans are utensils; swords and pistols are weapons. They can all be said to have properties or qualities: color, texture, composition, size, shape and features of shape, mass, elasticity, rigidity, and mobility. Orthodox psychology asserts that *we perceive these objects insofar as we discriminate their properties or qualities*. Psychologists carry out elegant experiments in the laboratory to find out how and how well these qualities are discriminated. The psychologists assume that objects are *composed* of their qualities. But



I now suggest that what we perceive when we look at objects are their affordances, not their qualities. We can discriminate the dimensions of difference if required to do so in an experiment, but what the object affords us is what we normally pay attention to. The special combination of qualities into which an object can be analyzed is ordinarily not noticed.

If this is true for the adult, what about the young child? There is much evidence to show that the infant does not begin by first discriminating the qualities of objects and then learning the combinations of qualities that specify them. Phenomenal objects are *not* built up of qualities; it is the other way around. The affordance of an object is what the infant begins by noticing. The meaning is observed before the substance and surface, the color and form, are seen as such. An affordance is an invariant combination of variables, and one might guess that it is easier to perceive such an invariant unit than it is to perceive all the variables separately. It is never necessary to distinguish *all* the features of an object and, in fact, it would be impossible to do so. Perception is economical. “Those features of a thing are noticed which distinguish it from other things that it is not—but not *all* the features that distinguish it from *everything* that it is not” (Gibson, 1966*b*, p. 286).

### **TO PERCEIVE AN AFFORDANCE IS NOT TO CLASSIFY AN OBJECT**

The fact that a stone is a missile does not imply that it cannot be other things as well. It can be a paperweight, a bookend, a hammer, or a pendulum bob. It can be piled on another rock to make a cairn or a stone wall. These affordances are all consistent with one another. The differences between them are not clear-cut, and the arbitrary names by which they are called do not count for perception. If you know what can be done with a graspable detached object, what it can be used for, you can call it whatever you please.

The theory of affordances rescues us from the philosophical muddle of assuming fixed classes of objects, each defined by its common features and then given a name. As Ludwig Wittgenstein knew, you *cannot* specify the necessary and sufficient features of the class of things to which a name is given. They have only a “family resemblance.” But this does not mean you cannot learn how to use things and perceive their uses. You do not have to classify and label things in order to perceive what they afford.

### ***Other Persons and Animals***

The richest and most elaborate affordances of the environment are provided by other animals and, for us, other people. These are, of course, detached objects

with topologically closed surfaces, but they change the shape of their surfaces while yet retaining the same fundamental shape. They move from place to place, changing the postures of their bodies, ingesting and emitting certain substances, and doing all this spontaneously, initiating their own movements, which is to say that their movements are *animate*. These bodies are subject to the laws of mechanics and yet *not* subject to the laws of mechanics, for they are not *governed* by these laws. They are so different from ordinary objects that infants learn almost immediately to distinguish them from plants and nonliving things. When touched they touch back, when struck they strike back; in short, they *interact* with the observer and with one another. Behavior affords behavior, and the whole subject matter of psychology and of the social sciences can be thought of as an elaboration of this basic fact. Sexual behavior, nurturing behavior, fighting behavior, cooperative behavior, economic behavior, political behavior—all depend on the perceiving of what another person or other persons afford, or sometimes on the misperceiving of it.

What the male affords the female is reciprocal to what the female affords the male; what the infant affords the mother is reciprocal to what the mother affords the infant; what the prey affords the predator goes along with what the predator affords the prey; what the buyer affords the seller cannot be separated from what the seller affords the buyer, and so on. The perceiving of these mutual affordances is enormously complex, but it is nonetheless lawful, and it is based on the pickup of the information in touch, sound, odor, taste, and ambient light. It is just as much based on stimulus information as is the simpler perception of the support that is offered by the ground under one's feet. For other animals and other persons can only give off information about themselves insofar as they are tangible, audible, odorous, tastable, or visible.

The other person, the generalized *other*, the *alter* as opposed to the *ego*, is an ecological object with a skin, even if clothed. It is an object, although it is not *merely* an object, and we do right to speak *of* *he* or *she* instead of *it*. But the other person has a surface that reflects light, and the information to specify what he or she is, invites, promises, threatens, or does can be found in the light.

### ***Places and Hiding Places***

The habitat of a given animal contains *places*. A place is not an object with definite boundaries but a region (Chapter 3). The different places of a habitat may have different affordances. Some are places where food is usually found and others where it is not. There are places of danger, such as the brink of a cliff and the regions where predators lurk. There are places of refuge from predators. Among these is the place where mate and young are, the home, which is usually a partial enclosure. Animals are skilled at what the psychologist calls place-learning. They can find their way to significant places.

An important kind of place, made intelligible by the ecological approach to visual perception, is a place that affords concealment, a *hiding place*. Note that it involves social perception and raises questions of epistemology. The concealing of oneself from other observers and the hiding of a detached object from other observers have different kinds of motivation. As every child discovers, a good hiding place for one's body is not necessarily a good hiding place for a treasure. A detached object can be concealed both from other observers and from the observer himself. The observer's body can be concealed from other observers but *not* from himself, as the last chapter emphasized. Animals as well as children hide themselves and also hide objects such as food.

One of the laws of the ambient optic array (Chapter 5) is that at any fixed point of observation some parts of the environment are revealed and the remaining parts are concealed. The reciprocal of this law is that the observer himself, his body considered as part of the environment, is revealed at some fixed points of observation and concealed at the remaining points. An observer can perceive not only that other observers are unhidden or hidden from him but also that he is hidden or unhidden from other observers. Surely, babies playing peek-a-boo and children playing hide-and-seek are practicing this kind of apprehension. To *hide* is to position one's body at a place that is concealed at the points of observation of other observers. A "good" hiding place is one that is concealed at nearly all points of observation.

All of these facts and many more depend on the principle of occluding edges at a point of observation, the law of reversible occlusion, and the facts of opaque and nonopaque substances. What we call privacy in the design of housing, for example, is the providing of opaque enclosures. A high degree of concealment is afforded by an enclosure, and complete concealment is afforded by a complete enclosure. But note that there are peepholes and screens that permit seeing without being seen. A transparent sheet of glass in a window transmits both illumination and information, whereas a *translucent* sheet transmits illumination but not information. There will be more of this in Chapter 11.

Note also that a glass wall affords seeing through but not walking through, whereas a cloth curtain affords going through but not seeing through. Architects and designers know such facts, but they lack a theory of affordances to encompass them in a system.

### ***Summary: Positive and Negative Affordances***

The foregoing examples of the affordances of the environment are enough to show how general and powerful the concept is. Substances have biochemical offerings and afford manufacture. Surfaces afford posture, locomotion, collision, manipulation, and in general behavior. Special forms of layout afford shelter and concealment. Fires afford warming and burning. Detached objects—tools, utensils, weapons—afford special types of behavior to primates and

humans. The other animal and the other person provide mutual and reciprocal affordances at extremely high levels of behavioral complexity. At the highest level, when vocalization becomes speech and manufactured displays become images, pictures, and writing, the affordances of human behavior are staggering. No more of that will be considered at this stage except to point out that speech, pictures, and writing still have to be perceived.

At all these levels, we can now observe that some offerings of the environment are beneficial and some are injurious. These are slippery terms that should only be used with great care, but if their meanings are pinned down to biological and behavioral facts the danger of confusion can be minimized. First, consider substances that afford ingestion. Some afford nutrition for a given animal, some afford poisoning, and some are neutral. As I pointed out before, these facts are quite distinct from the affording of pleasure and displeasure in eating, for the experiences do not necessarily correlate with the biological effects. Second, consider the brink of a cliff. On the one side it affords walking along, locomotion, whereas on the other it affords falling off, injury. Third, consider a detached object with a sharp edge, a knife. It affords cutting if manipulated in one manner, but it affords being cut if manipulated in another manner. Similarly, but at a different level of complexity, a middle-sized metallic object affords grasping, but if charged with current it affords electric shock. And fourth, consider the other person. The animate object can give caresses or blows, contact comfort or contact injury, reward or punishment, and it is not always easy to perceive which will be provided. Note that all these benefits and injuries, these safeties and dangers, these positive and negative affordances are properties of things *taken with reference to an observer* but not properties of the *experiences of the observer*. They are not subjective values; they are not feelings of pleasure or pain added to neutral perceptions.

There has been endless debate among philosophers and psychologists as to whether values are physical or phenomenal, in the world of matter or only in the world of mind. For affordances as distinguished from values, the debate does not apply. Affordances are neither in the one world or the other inasmuch as the theory of two worlds is rejected. There is only one environment, although it contains many observers with limitless opportunities for them to live in it.

### The Origin of the Concept of Affordances: A Recent History

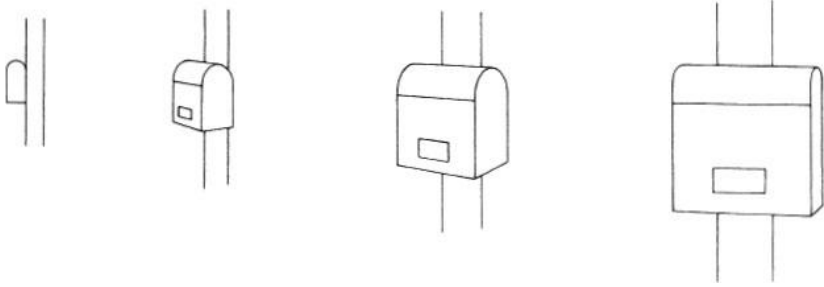
The gestalt psychologists recognized that the meaning or the value of a thing seems to be perceived just as immediately as its color. The value is clear *on the face of it*, as we say, and thus it has a *physiognomic* quality in the way that the emotions of a man appear *on his face*. To quote from the *Principles of Gestalt Psychology* (Koffka, 1935), “Each thing says what it is . . . a fruit says ‘Eat me’; water says ‘Drink me’; thunder says ‘Fear me’; and woman says ‘Love me’” (p. 7). These values are vivid and essential features of the experience itself.

Koffka did not believe that a meaning of this sort could be explained as a pale context of memory images or an unconscious set of response tendencies. The postbox “invites” the mailing of a letter, the handle “wants to be grasped,” and things “tell us what to do with them” (p. 353). Hence, they have what Koffka called “demand character.”

Kurt Lewin coined the term *Aufforderungscharakter*, which has been translated as *invitation character* (by J. F. Brown in 1929) and as *valence* (by D. K. Adams in 1931; cf. Marrow, 1969, p. 56, for the history of these translations). The latter term came into general use. *Valences* for Lewin had corresponding *vectors*, which could be represented as arrows pushing the observer toward or away from the object. What explanation could be given for these valences, the characters of objects that invited or demanded behavior? No one, not even the gestalt theorists, could think of them as physical and, indeed, they do not fall within the province of ordinary physics. They must therefore be phenomenal, given the assumption of dualism. If there were *two* objects, and if the valence could not belong to the physical object, it must belong to the phenomenal object—to what Koffka called the “behavioral” object but not to the “geographical” object. The valence of an object was bestowed upon it in experience, and bestowed by a need of the observer. Thus, Koffka argued that the postbox has a demand character only when the observer needs to mail a letter. He is attracted to it when he has a letter to post, not otherwise. The value of something was assumed to change as the need of the observer changed.

The concept of affordance is derived from these concepts of valence, invitation, and demand but with a crucial difference. The affordance of something does *not change* as the need of the observer changes. The observer may or may not perceive or attend to the affordance, according to his needs, but the affordance, being invariant, is always there to be perceived. An affordance is not bestowed upon an object by a need of an observer and his act of perceiving it. The object offers what it does because it is what it is. To be sure, we define *what it is* in terms of ecological physics instead of physical physics, and it therefore possesses meaning and value to begin with. But this is meaning and value of a new sort.

For Koffka it was the *phenomenal* postbox that invited letter-mailing, not the *physical* postbox. But this duality is pernicious. I prefer to say that the real postbox (the *only* one) affords letter-mailing to a letter-writing human in a community with a postal system. This fact is perceived when the postbox is identified as such, and it is apprehended whether the postbox is in sight or out of sight. To feel a special attraction to it when one has a letter to mail is not surprising, but the main fact is that it is perceived as part of the environment—as an item of the neighborhood in which we live. Everyone above the age of six knows what it is for and where the nearest one is. The perception of its affordance should therefore not be confused with the temporary special attraction it may have.



**FIGURE 8.1** The changing perspective structure of a postbox during approach by an observer.

As one reduces the distance to the object to one-third, the visual solid angle of the object increases three times. Actually this is only a detail near the center of an outflowing optic array. (From *The Perception of the Visual World* by James Jerome Gibson and used with the agreement of the reprint publisher, Greenwood Press, Inc.)

The gestalt psychologists explained the directness and immediacy of the experience of valences by postulating that the ego is an object in experience and that a “tension” may arise between a phenomenal object and the phenomenal ego. When the object is in “a dynamic relation with the ego” said Koffka, it has a demand character. Note that the “tension,” the “relation,” or the “vector” must arise in the “field,” that is, in the field of phenomenal experience. Although many psychologists find this theory intelligible, I do not. There is an easier way of explaining why the values of things seem to be perceived immediately and directly. It is because the affordances of things for an observer are specified in stimulus information. They *seem* to be perceived directly because they *are* perceived directly.

The accepted theories of perception, to which the gestalt theorists were objecting, implied that *no* experiences were direct except sensations and that sensations mediated all other kinds of experience. Bare sensations had to be clothed with meaning. The seeming directness of meaningful perception was therefore an embarrassment to the orthodox theories, and the Gestaltists did right to emphasize it. They began to undermine the sensation-based theories. But their own explanations of why it is that a fruit says “Eat me” and a woman says “Love me” are strained. The gestalt psychologists objected to the accepted theories of perception, but they never managed to go beyond them.

### The Optical Information for Perceiving Affordances

The theory of affordances is a radical departure from existing theories of value and meaning. It begins with a new definition of what value and meaning *are*. The perceiving of an affordance is not a process of perceiving a value-free phys-

ical object to which meaning is somehow added in a way that no one has been able to agree upon; it is a process of perceiving a value-rich ecological object. Any substance, any surface, any layout has some affordance for benefit or injury to someone. Physics may be value-free, but ecology is not.

The central question for the theory of affordances is not whether they exist and are real but whether information is available in ambient light for perceiving them. The skeptic may now be convinced that there is information in light for some properties of a surface but not for such a property as being good to eat. The taste of a thing, he will say, is not specified in light; you can see its form and color and texture but not its palatability; you have to *taste* it for that. The skeptic understands the stimulus variables that specify the dimensions of visual sensation; he knows from psychophysics that brightness corresponds to intensity and color to wavelength of light. He may concede the invariants of structured stimulation that specify surfaces and how they are laid out and what they are made of. But he may boggle at invariant combinations of invariants that specify the affordances of the environment for an observer. The skeptic familiar with the experimental control of stimulus variables has enough trouble understanding the invariant variables I have been proposing without being asked to accept invariants of invariants.

Nevertheless, a unique combination of invariants, a *compound* invariant, is just another invariant. It is a unit, and the components do not *have* to be combined or associated. Only if percepts were combinations of sensations would they have to be associated. Even in the classical terminology, it could be argued that when a number of stimuli are completely covariant, when they *always* go together, they constitute a single “stimulus.” If the visual system is capable of extracting invariants from a changing optic array, there is no reason why it should not extract invariants that seem to us highly complex.

The trouble with the assumption that high-order optical invariants specify high-order affordances is that experimenters, accustomed to working in the laboratory with low-order stimulus variables, cannot think of a way to *measure* them. How can they hope to isolate and control an invariant of optical structure so as to apply it to an observer if they cannot quantify it? The answer comes in two parts, I think. First, they should not hope to *apply* an invariant to an observer, only to make it available, for it is not a stimulus. And, second, they do not have to quantify an invariant, to apply numbers to it, but only to give it an exact mathematical description so that other experimenters can make it available to *their* observers. The virtue of the psychophysical experiment is simply that it is disciplined, not that it relates the psychical to the physical by a metric formula.

An affordance, as I said, points two ways, to the environment and to the observer. So does the information to specify an affordance. But this does not in the least imply separate realms of consciousness and matter, a psychophysical dualism. It says only that the information to specify the utilities of the environ-

ment is accompanied by information to specify the observer himself, his body, legs, hands, and mouth. This is only to reemphasize that exteroception is accompanied by proprioception—that to perceive the world is to coperceive oneself. This is wholly inconsistent with dualism in any form, either mind-matter dualism or mind-body dualism. The awareness of the world and of one's complementary relations to the world are not separable.

The child begins, no doubt, by perceiving the affordances of things for her, for her own personal behavior. She walks and sits and grasps relative to her own legs and body and hands. But she must learn to perceive the affordances of things for other observers as well as for herself. An affordance is often valid for all the animals of a species, as when it is part of a niche. I have described the invariants that enable a child to perceive the same solid shape at different points of observation and that likewise enable two or more children to perceive the same shape at different points of observation. These are the invariants that enable two children to perceive the common *affordance* of the solid shape despite the different perspectives, the affordance of a toy, for example. Only when each child perceives the values of things for others as well as for herself does she begin to be socialized.

### Misinformation for Affordances

If there is information in the ambient light for the affordances of things, can there also be misinformation? According to the theory being developed, if information is picked up perception results; if misinformation is picked up misperception results.

The brink of a cliff affords falling off; it is in fact dangerous and it looks dangerous to us. It seems to look dangerous to many other terrestrial animals besides ourselves, including infant animals. Experimental studies have been made of this fact. If a sturdy sheet of plate glass is extended out over the edge it no longer affords falling and in fact is not dangerous, but it may still *look* dangerous. The optical information to specify depth-downward-at-an-edge is still present in the ambient light; for this reason the device was called a *visual cliff* by E. J. Gibson and R. D. Walk (1960). Haptic information was available to specify an adequate surface of support, but this was contradictory to the optical information. When human infants at the crawling stage of locomotion were tested with this apparatus, many of them would pat the glass with their hands but would not venture out on the surface. The babies misperceived the affordance of a transparent surface for support, and this result is not surprising.

Similarly, an adult can misperceive the affordance of a sheet of glass by mistaking a closed glass door for an open doorway and attempting to walk through it. He then crashes into the barrier and is injured. The affordance of collision was not specified by the outflow of optical texture in the array, or it was insufficiently specified. He mistook glass for air. The occluding edges of



the doorway were specified and the empty visual solid angle opened up symmetrically in the normal manner as he approached, so his behavior was properly controlled, but the imminence of collision was not noticed. A little dirt on the surface, or highlights, would have saved him.

These two cases are instructive. In the first a surface of support was mistaken for air because the optic array specified air. In the second case a *barrier* was mistaken for air for the same reason. Air downward affords falling and is dangerous. Air forward affords passage and is safe. The mistaken perceptions led to inappropriate actions.

Errors in the perception of the surface of support are serious for a terrestrial animal. If quicksand is mistaken for sand, the perceiver is in deep trouble. If a covered pitfall is taken for solid ground, the animal is trapped. A danger is sometimes hidden—the shark under the calm water and the electric shock in the radio cabinet. In the natural environment, poison ivy is frequently mistaken for ivy. In the artificial environment, acid can be mistaken for water.

### **THINGS THAT LOOK LIKE WHAT THEY ARE**

If the affordances of a thing are perceived correctly, we say that it looks like what it *is*. But we must, of course, *learn* to see what things really are—for example, that the innocent-looking leaf is really a nettle or that the helpful-sounding politician is really a demagogue. And this can be very difficult.

A wildcat may be hard to distinguish from a cat, and a thief may look like an honest person. When Koffka asserted that “each thing says what it is,” he failed to mention that it may lie. More exactly, a thing may not look like what it is.

Nevertheless, however true all this may be, the basic affordances of the environment are perceivable and are usually perceivable directly, without an excessive amount of learning. The basic properties of the environment that make an affordance are specified in the structure of ambient light, and hence the affordance itself is specified in ambient light. Moreover, an invariant variable *that is commensurate with the body of the observer himself* is more easily picked up than one not commensurate with his body.

### **Summary**

The medium, substances, surfaces, objects, places, and other animals have affordances for a given animal. They offer benefit or injury, life or death. This is why they need to be perceived.

The possibilities of the environment and the way of life of the animal go together inseparably. The environment constrains what the animal can do, and the concept of a niche in ecology reflects this fact. Within limits, the human animal can alter the affordances of the environment but is still the creature of his or her situation.

There is information in stimulation for the physical properties of things, and presumably there is information for the environmental properties. The doctrine that says we must distinguish among the variables of things before we can learn their meanings is questionable. Affordances are properties taken with reference to the observer. They are neither physical nor phenomenal.

The hypothesis of information in ambient light to specify affordances is the culmination of ecological optics. The notion of invariants that are related at one extreme to the motives and needs of an observer and at the other extreme to the substances and surfaces of a world provides a new approach to psychology.

# 9

## EXPERIMENTAL EVIDENCE FOR DIRECT PERCEPTION

### Persisting Layout

Direct perception is what one gets from seeing Niagara Falls, say, as distinguished from seeing a picture of it. The latter kind of perception is *mediated*. So when I assert that perception of the environment is direct, I mean that it is not mediated by *retinal* pictures, *neural* pictures, or *mental* pictures. *Direct perception* is the activity of getting information from the ambient array of light. I call this a process of *information pickup* that involves the exploratory activity of looking around, getting around, and looking at things. This is quite different from the supposed activity of getting information from the inputs of the optic nerves, whatever they may prove to be.

The evidence for direct visual perception has accumulated slowly, over many years. The very idea had to be developed, the results of old experiments had to be reinterpreted, and new experiments had to be carried out. The next two chapters are devoted to the experimental evidence.

The experiments will be considered under three main headings: first, the direct perception of surface layout; second, the direct perception of *changing* surface layout; and third, the direct perception of the movements of the self. This chapter is devoted to the direct perception of surface layout.

### **Evidence for the Direct Perception of Surface Layout**

Some thirty years ago, during World War II, psychologists were trying to apply the theory of depth perception to the problems of aviation, especially the problem of how a flier lands an airplane. Pilots were given tests for depth perception, and there was controversy as to whether depth perception was learned or innate. The same tests are still being given, and the same disagreement continues.

The theory of depth perception assumes that the third dimension of space is lost in the two-dimensional retinal image. Perception must begin with form perception, the flat patchwork of colors in the visual field. But there are supposedly *cues* for depth, which, if they are utilized, will add a third dimension to the flat visual field. A list of the cues for depth is given in most psychology textbooks: linear perspective, apparent size, superposition, light and shade, relative motion, aerial perspective, accommodation (the monocular cues), along with binocular disparity and convergence (the binocular cues). You might suppose that adequate tests could be made of a prospective flier's ability to use these cues and that experiments could be devised to find out whether or not they were learned.

The trouble was that none of the tests based on the cues for depth predicted the success or failure of a student pilot, and none of the proposals for improving depth perception by training made it any easier to learn to fly. I was deeply puzzled by this fact. The accepted theory of depth perception did not work. It did not apply to problems where one might expect it to apply. I began to suspect that the traditional list of cues for depth was inadequate. And in the end I came to believe that the whole theory of depth perception was false.

I suggested a new theory in a book on what I called the *visual world* (Gibson, 1950b). I considered "the possibility that there is literally no such thing as a perception of space without the perception of a continuous background surface" (p. 6). I called this a *ground theory* of space perception to distinguish it from the *air theory* that seemed to underlie the old approach. The idea was that the world consisted of a basic surface with adjoining surfaces, not of bodies in empty air. The character of the visual world was given not by objects but by the background of the objects. Even the space of the airplane pilot, I said, was determined by the ground and the horizon of the earth, not by the air through which he flies. The notion of space of three dimensions with three axes for Cartesian coordinates was a great convenience for mathematics, I suggested, but an abstraction that had very little to do with actual perception.

I would now describe the ground theory as a theory of the *layout* of surfaces. By *layout*, I mean the relations of surfaces to the ground and to one another, their arrangement. The layout includes both places and objects, together with other features. The theory asserts that the perception of surface layout is direct. This means that perception does not begin with two-dimensional form perception. Hence, there is no special kind of perception called depth perception, and the third dimension is not lost in the retinal image since it was never in the environment to begin with. It is a loose term. If *depth* means the dimension of an object that goes with height and width, there is nothing special about it. Height becomes depth when the object is seen from the top, and width becomes depth when the object is seen from the side. If depth means distance from *here*, then it involves self-perception and is continually changing as the observer moves about. The theory of depth perception is based on confusion and perpetuated by the fallacy of the retinal picture.

I now say that there is information in ambient light for the perception of the layout of surfaces but not that there are cues or clues for the perception of depth. The traditional list of cues is worthless if perception does not begin with a flat picture. I tried to reformulate the list in 1950 as “gradients and steps of retinal stimulation” (Gibson, 1950*b*, pp. 137 ff.). The hypothesis of gradients was a good beginning, but the reformulation failed. It had the great handicap of being based on physiological optics and the retinal image instead of ecological optics and the ambient array.

Such is the hypothesis of the direct perception of surface layout. What is the evidence to support it? Some experiments had been carried out even before 1950, outdoor experiments in the open air instead of laboratory experiments with spots of light in a darkroom, but they were only a beginning (Gibson, 1947). Much more experimental evidence has accumulated in the last twenty-five years.

### ***The Psychophysics of Space and Form Perception***

The studies to be described were thought of as psychophysical experiments at the time they were performed. There was to be a new psychophysics of perception as well as the old psychophysics of sensation. For I thought I had discovered that there were stimuli for perceptions in much the same way that there were known to be stimuli for sensations. This now seems to me a mistake. I failed to distinguish between stimulation proper and stimulus information, between what happens at passive receptors and what is available to active perceptual systems. Traditional psychophysics is a laboratory discipline in which physical stimuli are applied to an observer. He is prodded with controlled and systematically varied bits of energy so as to discover how his experience varies correspondingly. This procedure makes it difficult or impossible for the observer to extract invariants over time. Stimulus prods do not ordinarily carry information about the environment.

What I had in mind by a psychophysics of perception was simply the emphasis on perception as direct instead of indirect. I wanted to exclude an extra process of inference or construction. I meant (or should have meant) that animals and people *sense* the environment, not in the meaning of having sensations but in the meaning of *detecting*. When I asserted that a gradient in the retinal image was a *stimulus* for perception, I meant only that it was sensed as a unit; it was not a collection of points whose separate sensations had to be put together in the brain. But the concept of the stimulus was not clear to me. I should have asserted that a gradient is stimulus *information*. For it is first of all an invariant property of an optic array. I should not have implied that a percept was an automatic response to a stimulus, as a sense impression is supposed to be. For even then I realized that perceiving is an act, not a response, an act of attention, not a triggered impression, an achievement, not a reflex.

So what I should have meant by a “psychophysical” theory of perception in 1950 and by perception as a “function of stimulation” in the essay I wrote in 1959 (Gibson, 1959) was the hypothesis of a one-stage process for the perception of surface layout instead of a two-stage process of first perceiving flat forms and then interpreting the cues for depth.

I now believe that there is no such thing as flat-form perception, just as there is no such thing as depth perception. (There are drawings and pictures, to be sure, but these are not “forms,” as I will explain in Part IV. The theory of form perception in psychology is no less confused than the theory of depth perception.) But this was not clear when I wrote my book in 1950, where I promised not only a psychophysics of space perception in Chapter 5 but also a psychophysical approach to form perception in Chapter 10. This sounded promising and progressive. Visual outline forms, I suggested, are not unique entities. “They could be arranged in a systematic way such that each form would differ only gradually and continuously from all others” (Gibson, 1950*b*, p. 193). What counts is not the form as such but the dimensions of variation of form. And psychophysical experiments could be carried out if these dimensions were isolated.

Here was the germ of the modern hypothesis of the distinctive features of graphic symbols. It also carries the faint suggestion of a much more radical hypothesis, that what the eye picks up is a sequential transformation, not a form. The study of form discrimination by psychophysical methods has flourished in the last thirty years. W. R. Garner, Julian Hochberg, Fred Attneave, and others have achieved the systematic variation of outline forms and patterns in elegant ways (e.g., Garner, 1974). My objection to this research is that it tells us nothing about perceiving the environment. It still assumes that vision is simplest when there is a form on the retina that copies a form on a surface facing the retina. It perpetuates the fallacy that form perception is basic. It holds back the study of invariants in a changing array. But the hypothesis that forms are directly perceived does not upset the orthodoxies of visual theory as does the hypothesis that invariants are directly perceived, and hence it is widely accepted.

The psychophysical approach to surface perception is much more radical than the psychophysical approach to form perception, and it has *not* been widely accepted over the last twenty-five years. Has its promise been fulfilled? Some experiments can be summarized, and the evidence should be pulled together.

### ***Experiments on the Perception of a Surface as Distinguished from Nothing***

#### ***Metzger's Experiment***

Is tridimensional space perception based on bidimensional sensations to which the third dimension is added, or is it based on surface perception? The first experiment bearing on this issue is that of W. Metzger in 1930. He faced the

eyes of his observer with a large, dimly lighted plaster wall, which rendered the light coming to the visual system unfocusable. Neither eye could accommodate, and probably the eyes could not converge. The total field (*Ganzfeld*) was, as he put it, *homogeneous*. Under high illumination, the observer simply perceived the wall, and the outcome was so obvious as to be uninteresting. But under low illumination, the fine-grained texture of the surface was no longer registered by a human eye, and the observer reported seeing what he called a fog or haze or mist of light. He certainly did not see a surface in two dimensions, and therefore Metzger was tempted to conclude that he saw something in three dimensions; that is, he was perceiving “space.”

But I did not see depth in the “mist of light.” Another way to get a homogeneous field is to confront the eyes with a hemisphere of diffusing glass highly illuminated from the outside (Gibson and Dibble, 1952). A better way is to cover each eye with a fitted cap of strongly diffusing translucent material worn like a pair of goggles (Gibson and Waddell, 1952). The structure of the entering light, the optical texture, can thus be eliminated at any level of intensity. What my observers and I saw under these conditions could better be described as “nothing” in the sense of “no thing.” It was like looking at the sky. There was no surface and no object at any distance. Depth was not present in the experience but missing from it. What the observer saw, as I would now put it, was an empty *medium*.

The essence of Metzger’s experiment and its subsequent repetitions is not the plaster wall or the panoramic surface or the diffusing glass globe or the eyecaps. The experiment *provides* discontinuities in the light to an eye at one extreme and *eliminates* them at the other. The purpose of the experiment is to control and vary the projective capacity of light. This must be isolated from the stimulating capacity of light. Metzger’s experiment points to the distinction between an optic array with structure and a nonarray without structure. To the extent that the array has structure it specifies an environment.

A number of experiments using a panoramic surface under low illumination have been carried out, although the experimenters did not always realize what they were doing. But all the experiments involved more or less faint discontinuities in the light to the eye. What the observers said they saw is complex and hard to describe. One attempt was made by W. Cohen in 1957, and the other experiments have been surveyed by L. L. Avant (1965). It is fair to say that there are intermediate perceptions between seeing *nothing* and seeing *something* as the discontinuities become stronger. These are the polar opposites of perception that are implied by Metzger’s experiment, not the false opposites of seeing in two dimensions and seeing in three dimensions.

The confusion over whether there is or is not “depth” in Metzger’s luminous fog is what led me to think that the whole theory of depth, distance, the third dimension, and space is misconceived. The important result is the neglected one that a surface is seen when the array has structure, that is, differences in

different directions. A perfectly flat surface in front of the eyes is still a layout, that is, a wall. And that is all that “seeing in two dimensions” can possibly mean.

### *The Experiment with Translucent Eye-Caps*

Eliminating optical texture from the light entering the eye by means of translucent diffusing goggles is an experiment that has been repeated many times. The observer is blind, not to light, for the photoreceptors are still stimulated, but to the environment, for the ocular system is inactivated; its adjustments are frustrated. The observer cannot *look at* or *look around*, and I shall devote a chapter to this activity later. The eye-caps have also been adapted for experiments on the development of vision in young animals. It was known that when diurnal animals such as primates were reared from birth in complete darkness they were blind by certain criteria when brought into an illuminated environment (although this was not true of nocturnal animals whose ancestors were used to getting around in the dark). Now it was discovered that animals deprived of optical structure but not of optical stimulation were also partly blind when the eye-caps were removed. Crudely speaking, they could not *use* their eyes properly. Anatomical degeneration of the photoreceptors had not occurred, as with the animals reared in the dark, but the exploratory adjustments of the visual system had not developed normally. The experiments are described in Chapter 12 of *Perceptual Learning and Development* by Eleanor J. Gibson (1969).

### *Experiments with a Sheet of Glass*

It is fairly well known that a clean sheet of plate glass that projects no reflections or highlights to the observer's eye is, as we say, invisible. This fact is not self-explanatory, but it is very interesting. It means that one perceives air where a material surface exists, because air is specified by the optic array. I have seen people try to walk through plate-glass doors to their great discomfiture and deer try to jump through plate-glass windows with fatal results.

A perfectly clear sheet of glass transmits both light considered as energy and an *array* of light considered as information. A frosted or pebbled sheet of glass transmits optical energy but *not* optical information. The clear sheet can be seen through, as we say, but the frosted sheet cannot. The latter can be seen, but the former cannot. An imperceptible sheet of glass can be made increasingly perceptible by letting dust or powder fall on it or by spattering it. Even the faintest specks can specify the surface. In this intermediate case, the sheet transmits both the array from the layout behind the glass and the array from the glass itself. We say that we see the farther surface *through* the glass surface. The optical structure of one is *mixed* or *interspersed* with the optical structure of the other.



The transparency of the near surface, more properly its semitransparency, is then perceived (Gibson, 1976). One sees two surfaces, separated in depth, in the same direction from here or, better, within the same visual solid angle of the ambient array. At least one sees them separated if the interspersed structures are different, or if the elements of one move relative to the elements of the other (E. J. Gibson, Gibson, Smith, and Flock, 1959).

Many of the above assertions are based on informal experiments that have not been published. But the reader can check them for himself with little trouble. I conclude that a surface is experienced when the structural information to specify it is picked up.

### *Experiments with a Pseudotunnel*

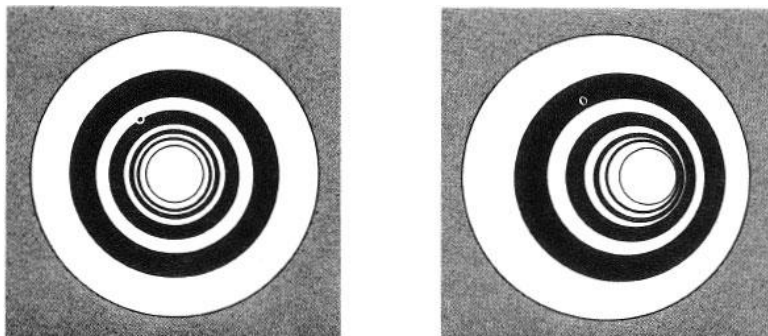
In the case of a sheet of glass, a surface may exist and go unperceived if it is not specified. In the next experiment, a surface may be nonexistent but may be perceived if it is specified. The pseudosurface in this case was not flat and frontal but was a semienclosure, a cylindrical tunnel viewed from one end. I called it an *optical tunnel* to suggest that the surface was not material or substantial but was produced by the light to the eye. Another way of describing it would be to say that it was a *virtual* but not a *real* tunnel.

The purpose of the experiment was to provide information for the perception of the inside surface of a cylinder without the ordinary source of this information, the inside surface of a cylinder. I would now call this a *display*. The fact that the perception was illusory is incidental. I wanted to elicit a synthetic perception, and I, therefore, had to synthesize the information. It was an experiment in perceptual psychophysics, more exactly, psycho-optics. The observers were fooled, to be sure, but that was irrelevant. There was no information in the array to specify that it *was* a display. This situation, I shall argue, is very rare.

My collaborators and I (Gibson, Purdy, and Lawrence, 1955) generated a visual solid angle of about  $30^\circ$  at the point of observation. This array consisted of alternating dark and light rings nested within one another, separated by abrupt circular contours. The number of rings and contours from the periphery to the center of the array could be varied. At one extreme there were thirty-six contours, and at the other seven.

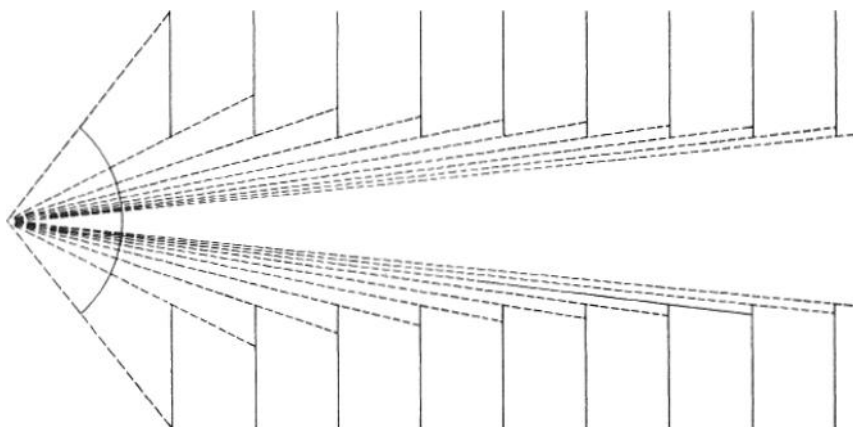
Thus the *mean density* of the contrasts in the array was varied from fine to coarse. The *gradient* of this density could also be varied; normally the density increased from the periphery toward the center.

The source of this display, the apparatus, was a set of large, very thin, plastic sheets, each hiding the next, with a one-foot hole cut in the center of each. They were indirectly illuminated from above or below. The contours in the array were caused by the edges of the sheets. The texture of the plastic was so fine as to be invisible. Black and white sheets could be hung in alternation one behind another, or, as a control, all-black or all-white surfaces could be



**FIGURE 9.1** The optic array coming to the eye from the optical tunnel.

There are nine contrasts in this cross-section of the array, that is, nine transitions of luminous intensity. The next figure shows a longitudinal section. The point of observation for the figure on the left is centered with the tunnel, whereas the point of observation for the figure on the right is to the right of center. (From J. J. Gibson, J. Purdy, and L. Lawrence: "A Method of Controlling Stimulation for the Study of Space Perception: The Optical Tunnel," *Journal of Experimental Psychology*, 1955, 50, 1–14. Copyright 1955 by the American Psychological Association. Reprinted by permission.)



**FIGURE 9.2** A longitudinal section of the optical tunnel shown in Figure 9.1.

Nine plastic sheets are shown, black and white alternating, with the cut edges of the nine holes aligned. The increase in the density of the contrasts from the periphery to the center of the array is evident. (From J. J. Gibson, J. Purdy, and L. Lawrence: "A Method of Controlling Stimulation for the Study of Space Perception: The Optical Tunnel," *Journal of Experimental Psychology*, 1955, 50, 1–14. Copyright 1955 by the American Psychological Association. Reprinted by permission.)

displayed. The observers looked into these holes from a booth, and extreme precautions were taken to prevent them from having any preconception of what they would see.

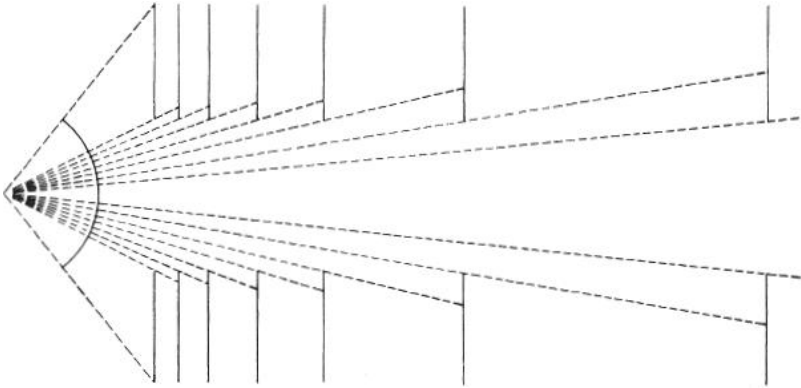
The principal result was as follows. When all-black or all-white surfaces were used, the observers saw nothing; the area within the first hole was described as a hazy or misty fog, a dark or light film, without obvious depth. At the other extreme, when thirty-six dark and light rings were displayed, all observers saw a continuous striped cylindrical surface, a solid tunnel. No edges were seen, and “a ball could be rolled from the far end to the entrance.”

When nineteen contrasts were displayed, two-thirds of the observers described a solid tunnel. When thirteen contrasts were displayed, half did so; and when seven contrasts were displayed, only one-third did so. In each case, the remainder said they saw either segments of surface with air in between or a series of circular edges (which was, of course, correct). With fewer contrasts, the experience became progressively less continuous and substantial. The proximity of these contours had proved to be crucial. *Surfaciness* depended on their mean density in the array.

What about the cylindrical shape of the surface, the receding layout of the tunnel? This could be altered in a striking way and the tunnel converted into a flat surface like an archery target with rings around a bull’s-eye simply by rearranging the sheets in the way illustrated. The *gradient* of increasing proximity toward the center of the array gives way to an equal proximity. But the target surface instead of the tunnel surface appeared only if the observer’s head was fixed and one eye was covered, that is, if the array was frozen and single. If the head was moved or the other eye used, the tunnel shape was again seen. The frozen array specified a flat target, but the dual or transforming array specified a receding tunnel. This is only one of many experiments in which perception with monocular fixed vision is exceptional.

### Conclusion

These experiments with a dimly lighted wall, with translucent eye-caps, with a sheet of glass, and with a pseudotunnel seem to show that the perception of *surfaciness* depends on the proximity to one another of discontinuities in the optic array. A surface is the interface between matter in the gaseous state and matter in the liquid or solid state. A surface comes to exist as the matter on one side of the interface becomes more *substantial* (Chapter 2). The medium is insubstantial. Mists, clouds, water, and solids are increasingly substantial. These substances are also increasingly *opaque*, except for a substance like glass, which is rare in nature. What these experiments have done is to vary systematically the optical information for the perception of substantiality and opacity. (But see the next chapter on the perception of *coherence*.)



**FIGURE 9.3** An arrangement that provides an array with a constant density of contrasts from periphery to center.

Only the first seven apertures are shown. The observer does not see a tunnel with this display but a flat surface with concentric rings, something like an archery target, so long as the head is immobile and one eye is covered. (From J. J. Gibson, J. Purdy, and L. Lawrence: "A Method of Controlling Stimulation for the Study of Space Perception: The Optical Tunnel," *Journal of Experimental Psychology*, 1955, 50, 1–14. Copyright 1955 by the American Psychological Association. Reprinted by permission.)

The experiment with the pseudotunnel also seems to show that the perception of a surface as such entails the perception of its layout, such as the front-facing layout of a wall or the slanting layout of a tunnel. Both are kinds of layout, and the traditional distinction between two-dimensional and three-dimensional vision is a myth.

### ***Experiments on the Perception of the Surface of Support***

The ground outdoors or the floor indoors is the main surface of support. Animals have to be supported against gravity. If the layout of surfaces is to be substituted for space in the theory of perception, this *fundamental* surface should get first consideration. How is it perceived? Animals like us can always *feel* the surface of support except when falling freely. But we can also *see* the surface of support under our feet if we are, in fact, supported. The ground is always specified in the lower portion of the ambient array. The standing infant can always see it and can always see her feet hiding parts of it. This is a law of ecological optics.

#### ***The Glass Floor***

A floor can be experimentally modified. When the "visual cliff" was being constructed for experiments with young animals by E. J. Gibson and R. D.

Walk (1960), observations were made with a large sheet of glass that was horizontal instead of vertical, a glass floor instead of a glass wall. The animal or child can be put down on this surface under two conditions: when it is visible, by virtue of textured paper placed just under the glass, and when it is invisible, with the paper placed far below the glass. The glass affords support under both conditions but provides *optical information* for support only under the first. There is mechanical contact with the feet in both cases but optical information for contact with the feet only in the first.

The animals or babies tested in this experiment would walk or crawl normally when they could both see and feel the surface but would not do so when they could only feel the surface; in the latter case, they froze, crouched, and showed signs of discomfort. Some animals even adopted the posture they would have when falling (E. J. Gibson and Walk, 1960, pp. 65–66). The conclusion seems to be that some animals require optical information for support along with the inertial and tactual information in order to walk normally. For my part, I should feel very uncomfortable if I had to stand on a large observation platform with a transparent floor through which the ground was seen far below.

The optical information in this experiment, I believe, is contradictory to the haptic information. One sees oneself as being up in the air, but one feels oneself in contact with a surface of support and, of course, one feels the normal pull of gravity in the vestibular organ. In such cases of contradictory or conflicting information, the psychologist cannot predict which will be picked up. The perceptual outcome is uncertain.

Note that the perception of the ground and the copercption of the self are inseparable in this situation. One's body *in relation to* the ground is what gets attention. Perception and proprioception are complementary. But the commonly accepted theories of space perception do not bring out this fact.

### *The Visual Cliff*

The visual cliff experiments of E. J. Gibson, R. D. Walk, and subsequently others are very well known. They represented a new approach to the ancient puzzle of depth perception, and the results obtained with newborn or dark-reared animals were surprising because they suggested that depth perception was innate. But the sight of a cliff is *not* a case of perceiving the third dimension. One perceives the affordance of its edge. A cliff is a feature of the terrain, a highly significant, special kind of dihedral angle in ecological geometry, a falling-off place. The edge at the top of a cliff is dangerous. It is an occluding edge. But it has the special character of being an edge of the surface of support, unlike the edge of a wall. One can safely walk around the edge of a wall but not off the edge of a cliff. To perceive a cliff is to detect a layout but, more than that, it is to detect an *affordance*, a negative affordance for locomotion, a place where the surface of support ends.

An affordance is for a species of animal, a layout *relative to* the animal and commensurate with its body. A cliff is a drop-off that is large relative to the size of the animal, and a step is a drop-off that is small relative to its size. A falling-off edge is dangerous, but a stepping-down edge is not. What animals need to perceive is not layout as such but the affordances of the layout, as emphasized in the last chapter. Consider the difference between the edge of a horizontal surface and the edge of a vertical surface, the edge of a floor and the edge of a wall. You go *over* the former whereas you can go *around* the latter. Both are dihedral angles, and both are occluding edges. But the meanings of the two kinds of “depth” are entirely different.

Gibson and Walk (1960; Walk and Gibson, 1961) constructed a virtual cliff with the glass-floor apparatus. They tested animals and babies to determine whether or not they would go forward over the virtual cliff. Actually, they provided two edges on either side of a narrow platform, one a falling-off edge and the other a stepping-down edge appropriate to the species of animal being tested. The animals’ choices were recorded. Nearly all terrestrial animals chose the shallow edge instead of the deep one.

The results have usually been discussed in terms of depth perception and the traditional cues for depth. But they are more intelligible in terms of the perception of layout and affordances. The separation in depth at an edge of the surface of support is not at all the same thing as the depth dimension of abstract space.

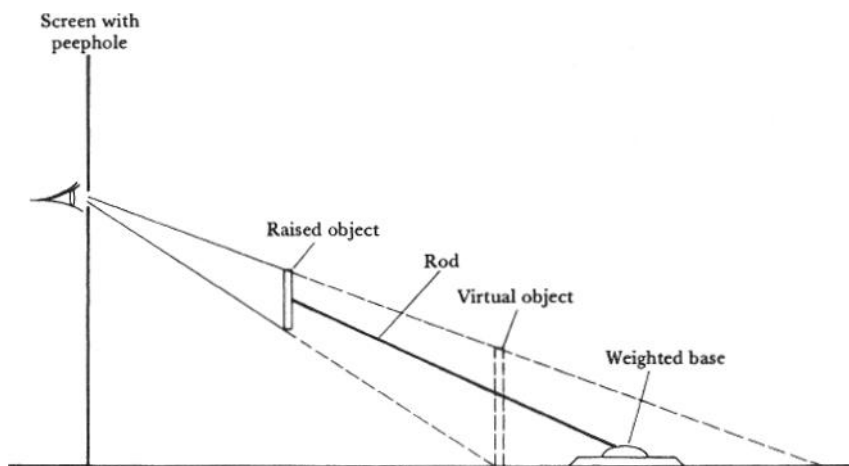


FIGURE 9.4 The invisibly supported object.

The real object is held up in the air by a hidden rod attached to a heavy base. The virtual object appears to be resting on the ground where the bottom edge of the real object hides the ground, so long as vision is monocular and frozen. One sees a concave corner, not an occluding edge. Because the virtual object is at twice the distance of the real object, it is seen as twice the size.

As for innate versus learned perception, it is much more sensible to assume an innate capacity to notice falling-off places in terrestrial animals than it is to assume that they have innate ideas or mental concepts of geometry.

### *An Object Resting on the Ground*

I suggested that one sees the contact of his feet with the ground. This is equally true for other objects than feet. We see whether an object is on the ground or up in the air. How is this contact with or separation from the ground perceived? The answer is suggested by an informal experiment described in my book on the visual world (Gibson, 1950*b*, Fig. 72, pp. 178 ff.), which might be called the *invisibly-supported-object experiment*. I did not clearly understand it at the time, but the optics of occluding edges now makes it more intelligible.

A detached object can be attached to a long rod that is hidden to the observer. The rod can be lowered by the experimenter so that the object rests on the ground or raised so that it stands up in the air. The object can be a cardboard rectangle or trapezoid or a ball, but it must be large enough to hide the rod and its base. An observer who stands at the proper position and looks with two eyes, or with one eye and a normally moving head, perceives a resting object as resting on the surface of support and a raised object as raised above the surface of support. The size and distance of the object are seen correctly. But an observer who looks with one eye and a fixed head, through a peephole or with a biting board, gets an entirely different perception. A resting object is seen correctly, but a raised object is also seen to be resting on the surface. It is seen *at the place where its edge hides the texture of the surface*. It appears farther away and larger than it really is.

This illusion is very interesting. It appears only with monocular arrested vision—a rare and unnatural kind of vision. The increments and decrements of the texture of the ground at the edges of the object have been eliminated, both those of one eye relative to the other and those that are progressive in time at each eye. In traditional theory, the cues of binocular and motion parallax are absent. But it is just these increments and decrements of the ground texture that *specify* the separation of object from ground. The absence of this accretion/deletion specifies contact of the object with the ground. A surface is perceived to “stand up” or “stand out” from the surface that extends behind it only to the extent that the gap is specified. And this depends on seeing from different points of observation, either two points of observation at the same time or different points of observation at different times.

A flat surface that “goes back to” or “lies flat on” the ground will seem to have a different size, shape, and even reflectance than it has when it stands forth in the air. This feature of the illusion is also very interesting, and I have demonstrated it many times. The first published study of it is that of J. E. Hochberg and J. Beck (1954).

### ***Experiments with the Ground as Background***

Investigators in the tradition of space perception and the cues for depth have usually done experiments with a background in the frontal plane, that is, a surface facing the observer, a wall, a screen, or a sheet of paper. A form in this plane is most similar to a form on the retina, and extension in this plane might be seen as a simple sensation. This follows from retinal image optics. But investigators of environment perception do experiments with the *ground* as background, studying surfaces instead of forms, and using ecological optics. Instead of studying distance in the air, they study recession along the ground. Distance as such cannot be seen directly but can only be inferred or computed. Recession along the ground can be seen directly.

### ***Distance and Size Perception on the Ground***

Although the linear perspective of a street in a painting had been known since the Renaissance, and the converging appearance of a parallel alley of trees in a landscape had been discussed since the eighteenth century, no one had ever studied the perception of a naturally textured ground. Linear perspective was an obvious cue for distance, but the gradient of density or proximity of the texture of the ground was not so obvious. E. G. Boring has described the old experiments with artificial alleys (1942, pp. 290–296), but the first experiment with an ordinary textured field outdoors, I believe, was published at the end of World War II (Gibson, 1947). A plowed field without furrows receding almost to the horizon was used. No straight edges were visible. This original experiment required the judgment of the height of a stake planted in the field at some distance up to half a mile. At such a distance the optical size of the elements of texture and the optical size of the stake itself were extremely small.

Up until that time the unanimous conclusion of observers had been that parallel lines were seen to converge and that objects were seen to be smaller “in the distance.” There was a tendency toward “size constancy” of objects, to be sure, but it was usually incomplete. The assumption had always been that size constancy must “break down.” It was supposed that an object will cease to be even *visible* at some eventual distance and that presumably it ceases to be visible by way of becoming smaller. (See Gibson, 1950*b*, p. 183, for a statement of this line of reasoning.) With the naive observers in the open field experiment, however, the judgments of the size of the stake did *not* decrease, even when it was a ten-minute walk away and becoming hard to make out. The judgments became more *variable* with distance but not smaller. Size constancy did not break down. The size of the object only became less *definite* with distance, not smaller.

The implication of this result, I now believe, is that certain invariant ratios were picked up unawares by the observers and that the size of the retinal image went unnoticed. No matter how far away the object was, it intercepted or occluded the



same number of texture elements of the ground. This is an invariant ratio. For any distance the proportion of the stake extending above the horizon to that extending below the horizon was invariant. This is another invariant ratio. These invariants are not cues but information for direct size perception. The observers in this experiment were aviation trainees and were not interested in the perspective appearance of the terrain and the objects. They could not care less for the patchwork of colors in the visual field that had long fascinated painters and psychologists. They were set to pick up information that would permit a size-match between the distant stake and one of a set of nearby stakes. The perception of the size and distance of an object on the ground had proved to be unlike the perception of the size and distance of an object in the sky. The invariants are missing in the latter case. The silhouette of an airplane might be a fifty-foot fighter at a one-mile altitude or a hundred-foot bomber at a two-mile altitude. Airplane spotters could be trained to estimate altitude, but only by the method of recognizing the shape, knowing the size by having memorized the wingspan, and inferring the distance from the angular size. Errors were considerable at best. This kind of inferential knowledge is not characteristic of ordinary perception. Baron von Helmholtz called it “unconscious” inference even in the ordinary case, but I am skeptical.

### *Comparison of Stretches of Distance Along the Ground*

The size of an object on the ground is not entirely separable from the sizes of the objects that compose the ground. The terrain is made of clods and particles of earth, or rocks and pebbles, or grass clumps and grass blades. These nested objects might have size constancy just as much as orthodox objects. In the next set of experiments on ground perception, the very distinction between size and distance breaks down. What had to be compared were not stakes or objects but *stretches* of the ground itself, distances between markers placed by the experimenter. In this case distances between *here* and *there* could be compared with distances between *there* and *there*. These open-field experiments were conducted by Eleanor J. Gibson (Gibson and Bergman, 1954; Gibson, Bergman, and Purdy, 1955; Purdy and Gibson, 1955).

Markers could be set down and moved anywhere in a level field of grass up to 350 yards away. The most interesting experiment of the series required the observer to *bisect* a stretch of distance, which could extend either from his feet to a marker or from one marker to another (Purdy and Gibson, 1955). A mobile marker on wheels had to be stopped by the observer at the halfway point. The ability to bisect a length had been tested in the laboratory with an adjustable stick called a Galton bar but not with a piece of ground on which the observer stood.

All observers could bisect a stretch of distance without difficulty and with some accuracy. The farther stretch could be matched to the nearer one, although the visual angles did not match. The farther visual angle was compressed

relative to the nearer, and its surface was, to use a vague term, foreshortened. But no constant error was evident. A stretch from *here* to *there* could be equated with a stretch from *there* to *there*. The conclusion must be that observers were not paying attention to the visual angles; they must have been noticing information. They might have been detecting, without knowing it, the *amount of texture* in a visual angle. The number of grass clumps projected in the farther half of a stretch of distance is exactly the same as the number projected in the nearer half. It is true that the optical texture of the grass becomes denser and more vertically compressed as the ground recedes from the observer, but the rule of *equal amounts of texture for equal amounts of terrain* remains invariant.

This is a powerful invariant. It holds for either dimension of the terrain, for width as well as for depth. In fact, it holds for any regularly textured surface whatever, that is, any surface of the same substance. And it holds for walls and ceilings as well as for floors. To say that a surface is regularly textured is only to assume that bits of the substance tend to be evenly spaced. They do not have to be perfectly regular like crystals in a lattice but only “stochastically” regular.

The implications of this experiment on fractionating a stretch of the ground are radical and far-reaching. The world consists not only of distances from *here*, my world, but also of distances from *there*, the world of another person. These intervals seem to be strikingly equivalent.

The rule of equal amounts of texture for equal amounts of terrain suggests that both size and distance are perceived directly. The old theory that the perceiver *allows for* the distance in perceiving the size of something is unnecessary. The assumption that the cues for distance *compensate for* the sensed smallness of the retinal image is no longer persuasive. Note that the pickup of the *amount of texture* in a visual solid angle of the optic array is not a matter of counting units, that is, of measuring with an arbitrary unit. The other experiments of this open-field series required the observers to make *absolute* judgments, so-called, of distances in terms of yards. They could *learn* to do so readily enough (E. J. Gibson and Bergman, 1954; E. J. Gibson, Bergman, and Purdy, 1955), but it was clear that one had to *see* the distance before one could apply a number to it.

### *Observations of the Ground and the Horizon*

When the terrain is flat and open, the horizon is in the ambient optic array. It is a great circle between the upper and the lower hemisphere separating the sky and the earth. But this is a limiting case. The farther stretches of the ground are usually hidden by frontal surfaces such as hills, trees, and walls. Even in an enclosure, however, there has to be a surface of support, a textured floor. The maximum coarseness of its optical texture is straight down, where the feet are, and the density increases outward from this center. These radial gradients projected from the surface of support increase with increasing size of the floor.

The densities of texture do not become infinite except when there is an infinitely distant horizon. Only at this limit is the optical structure of the array wholly compressed. But the gradients of density specify where the outdoors horizon would be, even in an enclosure. That is, there exists an implicit horizon even when the earth-sky horizon is hidden.

### EVEN SPACING

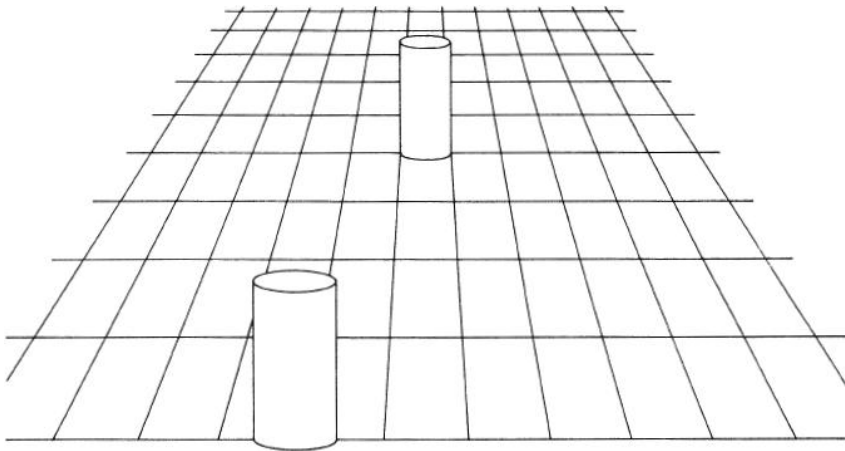
The fact that the parts of the terrestrial environment tend to be “evenly spaced” was noted in my early book on the visual world (Gibson, 1950*b*, pp. 77–78). This is equivalent to the rule of equal amounts of texture for equal amounts of terrain. The fact can be stated in various ways. However stated, it seems to be a fact that can be seen, not necessarily an intellectual concept of abstract space including numbers and magnitudes. Ecological geometry does not have to be learned from textbooks.

The concept of a *vanishing point* comes from artificial perspective, converging parallels, and the theory of the picture plane. The *vanishing limit* of optical structure at the horizon comes from natural perspective, ecological optics, and the theory of the ambient optic array. The two kinds of perspective should not be confused, although they have many principles in common (Chapter 5).

The terrestrial horizon is thus an invariant feature of terrestrial vision, an invariant of any and all ambient arrays, at any and all points of observation. The horizon never moves, even when every other structure in the light is changing. This stationary great circle is, in fact, that to which all optical motions have reference. It is neither subjective nor objective; it expresses the *reciprocity* of observer and environment; it is an invariant of *ecological optics*.

The horizon is the same as the skyline only in the case of the open ground or the open ocean. The earth-sky contrast may differ from the true horizon because of hills or mountains. The horizon is perpendicular to the pull of gravity and to the two poles of the ambient array at the centers of the two hemispheres; in short, the horizon is horizontal. With reference to this invariant, all other objects, edges, and layouts in the environment are judged to be either *upright* or *tilted*. In fact, the observer perceives *himself* to be in an upright or tilted posture relative to this invariant. (For an early and more complex discussion of visual uprightness and tilt in terms of the retinal image, see Gibson, 1952, on the “phenomenal vertical.”)

The facts about the terrestrial horizon are scarcely mentioned in traditional optics. The only empirical study of it is one by H. A. Sedgwick (1973) based on ecological optics. He shows how the horizon is an important source of invariant information for the perception of all kinds of objects. All terrestrial objects, for



**FIGURE 9.5** The base of each pillar covers the same amount of the texture of the ground.

The width of each pillar is that of one paving stone. The pillars will be seen to have the same width if this information is picked up. The height of each pillar is specified by a similar invariant, the “horizon-ratio” relation, described later.

example, of the same height are cut by the horizon in the same ratio, no matter what the angular size of the object may be. This is the “horizon ratio relation” in its simplest form. Any two trees or poles bisected by the horizon are the same height, and they are also precisely twice my eye-height. More complex ratios specify more complex layouts. Sedgwick showed that judgments of the sizes of objects represented in pictures were actually determined by these ratios.

The perceiving of what might be called *eye level* on the walls, windows, trees, poles, and buildings of the environment is another case of the complementarity between seeing the layout of the environment and seeing oneself in the environment. The horizon is at eye level relative to the furniture of the earth. But this is my eye level, and it goes up and down as I stand and sit. If I want my eye level, the horizon, to rise above all the clutter of the environment, I must climb up to a high place. The perception of *here* and the perception of *infinitely distant from here* are linked.

### ***Experiments on the Perception of Slant***

Experiments on the direct perception of layout began in 1950. From the beginning, the crucial importance of the *density of optical texture* was evident. How could it be varied systematically in an experiment? Along with the outdoor experiments, I wanted to try indoor experiments in the laboratory. I did not

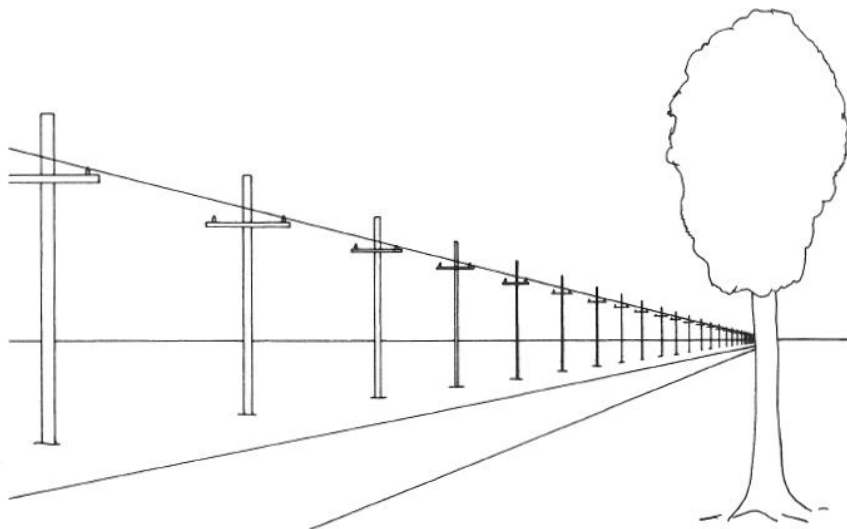
then understand ambient light but only the retinal image, and this led me to experiment with texture density in a *window* or *picture*. The density could be increased upward in the display (or downward or rightward or leftward), and the virtual surface would then be expected to *slant* upward (or downward or whatever). The surface should slant *away* in the direction of increasing texture density; it should be inclined from the frontal plane at a certain angle that corresponded to the rate of change of density, the *gradient* of density. Every piece of surface in the world, I thought, had this quality of slant (Gibson, 1950a). The slant of the apparent surface behind the apparent window could be judged by putting the palm of the hand at the same inclination from the frontal plane and recording it with an adjustable "palm board." This appeared to be a neat psychophysical experiment, for it isolated a variable, the gradient of density.

The first experiment (Gibson 1950a) showed that with a uniform density over the display the phenomenal slant is zero and that with increases of density in a given direction one perceives increasing slant in that direction. But the apparent slant was not proportional to the geometrically predicted slant. It was less than it should be theoretically. The experiment has been repeated with modifications by Gibson and J. Cornsweet (1952), J. Beck and J. J. Gibson (1955), R. Bergman and J. J. Gibson (1959), and many other investigators. It is *not* a neat psychophysical experiment. Phenomenal slant does not simply correspond to the gradient. The complexities of the results are described by H. R. Flock (1964, 1965) and by R. B. Freeman (1965).

What was wrong with these experiments? In consideration of the theory of layout, we can now understand it. The kind of slant studied was *optical*, not *geographical*, as noted by Gibson and Cornsweet (1952). It was relative to the frontal plane perpendicular to the line of sight, not relative to the surface of the earth, and was thus merely a new kind of depth, a quality added to each of the flat forms in the patchwork of the visual field. I had made the mistake of thinking that the experience of the layout of the environment could be *compounded* of all the optical slants of each piece of surface. I was thinking of slant as an absolute quality, whereas it is always relative. Convexities and concavities are not made up of elementary impressions of slant but are instead unitary features of the layout.

The impression of slant cannot be isolated by displaying a texture inside a window, for the perception of the occluding edge of the window will affect it; the surface is slanted relative to the surface that has the window in it. The separation of these surfaces is underestimated, as the experimental results showed.

The supposedly absolute judgment of the slant of a surface behind a window becomes more accurate when a graded decrease of *velocity* of the texture across the display is substituted for a graded increase of *density* of the texture, as demonstrated by Flock (1964). The virtual surface "stands back" from the virtual window. It slants away in the direction of decreasing flow of the texture but is



**FIGURE 9.6** The invariant horizon ratio for terrestrial objects.

The telephone poles in this display are all cut by the horizon in the same ratio. The proportion differs for objects of different heights. The line where the horizon cuts the tree is just as high above the ground as the point of observation, that is, the height of the observer's eye. Hence everyone can see his own *eye-height* on the standing objects of the terrain.

perceived to be a rigidly moving surface if the flow gradient is mathematically appropriate. But this experiment belongs not with experiments on surface layout but with those on *changing* surface layout, and these experiments will be described later.

### Is There Evidence Against the Direct Perception of Surface Layout?

There are experiments, of course, that seem to go against the theory of a direct perception of layout and to support the opposite theory of a *mediated* perception of layout. The latter theory is more familiar. It asserts that perception is mediated by assumptions, preconceptions, expectations, mental images, or any of a dozen other hypothetical mediators. The demonstrations of Adelbert Ames, once very popular, are well known for being interpreted in this way, especially the Distorted Room and the Rotating Trapezoidal Window.

These demonstrations are inspired by the *argument from equivalent configurations*. A diagram illustrating equivalent configurations is given in Figure 9.7. The argument is that many possible objects can give rise to one retinal image and that hence a retinal image cannot specify the object that gave rise to it.

But the image, according to the argument, is all one has for information. The perception of an object, therefore, requires an *assumption* about which of the many possible objects that could exist gave rise to the present image (or to the visual solid angle corresponding to it). The argument is supposed to apply to each of a collection of objects in space.

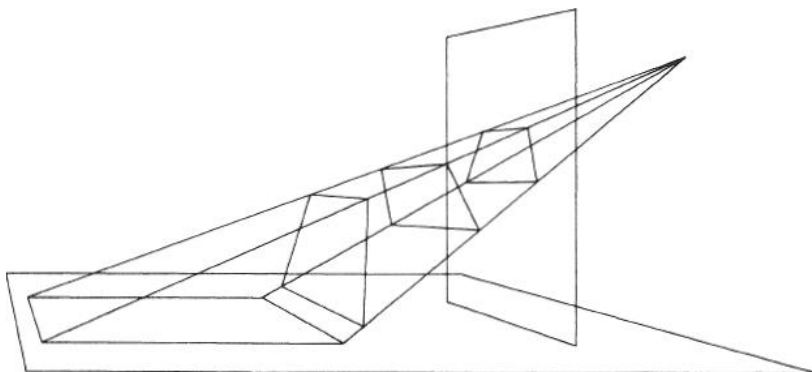
A distorted room with trapezoidal surfaces can be built so as to give rise to a visual solid angle at the point of observation identical with the solid angle from a normal rectangular room. Or a trapezoidal window with trapezoids for windowpanes can be built and made to rotate so that its changing visual solid angle is identical with the changing solid angle from a rectangular window slanted  $45^\circ$  away from the real distorted window. The window is always one-eighth of a rotation behind itself, as it were. A single and stationary point of observation is taken for granted. An observer who looks with one eye and a stationary head misperceives the trapezoidal surfaces and has the experience of a set of rectangular surfaces, a “virtual” form or window, instead of the actual plywood construction invented by the experimenter. Anomalies of perception result that are striking and curious. The eye has been fooled.

The explanation is that, in the absence of information, the observer has presupposed (assumed, expected, or whatever) the existence of rectangular surfaces causing the solid angles at the eye. That is reasonable, but it is then concluded that presuppositions are necessary for perception in general, since a visual solid angle cannot specify its object. There will always be equivalent configurations for any solid angle or any set of solid angles at a point of observation.

The main fallacy in this conclusion, as the reader will recognize, is the generalization from peephole observation to ordinary observation, the assumption that because the perspective structure of an optic array does not specify the surface layout nothing in the array can specify the layout. The hypothesis of invariant structure that underlies the perspective structure and emerges clearly when there is a shift in the point of observation goes unrecognized. The fact is that when an observer uses two eyes and certainly when one looks from various points of view the abnormal room and the abnormal window are perceived for what they are, and the anomalies cease.

The demonstrations do not prove, therefore, that the perception of layout cannot be direct and must be mediated by preconceptions, as Adelbert Ames and his followers wanted to believe (Ittelson, 1952). Neither do the many other demonstrations that, over the centuries, have purported to prove it.

The diagram of equivalent configurations illustrates one of the perplexities inherent to the retinal image theory of perception: if many different objects can give rise to the same stimulus, how do we ever perceive an object? The other half of the puzzle is this: if the same object can give rise to many different stimuli, how can we perceive the object? (Note that the second question implies a moving object but that neither question admits the fact of a moving observer.) Koffka was perplexed by this dual puzzle (1935, pp. 228 ff.) and many other experimenters have tried to



**FIGURE 9.7** Equivalent configurations within the same visual solid angle.

This perspective drawing shows a rectangle and three transparent trapezoids, all of which fit within the envelope of the same visual solid angle. Thus all four quadrangles are theoretically equivalent for a single eye at a fixed point of observation. They are, however, ghosts, not surfaces.

resolve it, but without success (e.g., Beck and Gibson, 1955). The only way out, I now believe, is to abandon the dogma that a retinal stimulus exists in the form of a picture. What specifies an object are invariants that are themselves “formless.”

## Summary

The experiment of providing either structure or no structure in the light to an eye results in the perception of a surface or no surface. The difference is not between seeing in two dimensions and seeing in three dimensions, as earlier investigators supposed.

The closer together the discontinuities in an experimentally induced optic array, the greater is the “surfacedness” of the perception. This was true, at least, for a 30° array having seven contours at one extreme and thirty-six at the other.

Optical contact of one’s body with the surface of support as well as mechanical contact seem to be necessary for some terrestrial animals if they are to stand and walk normally.

Perceiving the meaning of an edge in the surface of support, either a falling-off edge or a stepping-down edge, seems to be a capability that animals develop. This is not abstract depth perception but affordance perception.

Experiments on the perception of distance along the ground instead of distance through the air suggest that such perception is based on invariants in the array instead of cues. The rule of equal amounts of texture for equal amounts of terrain is one such invariant, and the horizon ratio relation is another. On this basis, the dimensions of things on the ground are perceived



directly, and the old puzzle of the constancy of perceived size at different distances does not arise.

The fact of the terrestrial horizon in the ambient array should not be confused with the vanishing point of linear perspective in pictorial optics.

A series of experiments on the perception of the slant of a surface relative to the line of sight did not confirm the absolute gradient hypothesis. The implication was that the slants of surfaces relative to one another and to the ground, the depth-shapes of the layout, are what get perceived.

Experiments based on the argument from equivalent configurations do not prove the need to have presuppositions in order to perceive the environment, since they leave out of account the fact that an observer normally moves about.

# 13

## LOCOMOTION AND MANIPULATION

The theory of affordances implies that to see things is to see how to get about among them and what to do or not do with them. If this is true, visual perception serves behavior, and behavior is controlled by perception. The observer who does not move but only stands and looks is not behaving at the moment, it is true, but he cannot help seeing the affordances for behavior in whatever he looks at.

Moving from place to place is supposed to be “physical” whereas perceiving is supposed to be “mental,” but this dichotomy is misleading. Locomotion is guided by visual perception. Not only does it depend on perception but perception depends on locomotion inasmuch as a moving point of observation is necessary for any adequate acquaintance with the environment. So we must perceive in order to move, but we must also move in order to perceive.

Manipulation is another kind of behavior that depends on perception and also facilitates perception. Let us consider in this chapter how vision enters into these two kinds of behavior.

### **The Evolution of Locomotion and Manipulation**

#### ***Support***

Animals, no less than other bodies, are pulled downward by the force of gravity. They fall unless supported. In water the animal is supported by the medium, which has about the same density as its body. But in air the animal must have a substantial surface below if it is not to become a Newtonian falling body.

Locomotion has evolved from swimming in the sea to crawling and walking on land to clinging and climbing on the protuberances that clutter up the land

and, finally, to flying through the air, the most rapid kind of locomotion but the most risky. Fish are supported by the medium, terrestrial animals by a substantial surface on the underside, and birds (when they are not at rest) by airflow, the aerodynamic force called *lift*. Zoologists sometimes classify animals as aquatic, terrestrial, or aerial, having in mind the different ways of getting about in water, on land, or in the air.

### ***Visual Perception of Support***

A terrestrial animal must have a surface that pushes up on its feet, or its underside. The experiments reported in Chapter 4 with the glass floor apparatus suggest that many terrestrial animals cannot maintain normal posture unless they can see their feet on the ground. With optical information to specify their feet *off* the ground, they act as if they were falling freely, crouching and showing signs of fear. But when a textured surface is brought up under the glass floor, the animals stand and walk normally (E. J. Gibson, 1969, pp. 267–270).

This result implies that contact of the feet with the surface of support as against separation of the feet from the surface is specified optically, at the occluding edges of the feet. The animal who moves its head or uses two eyes can perceive either *no* separation in depth between its feet and the floor or the kind of separation it would see if it were suspended in air. Contact is specified both optically and mechanically.

Note that a rigid surface of earth can be distinguished from a nonrigid surface of water by its color, texture, and the absence or presence of ripples. A surface of water does not afford support for chicks, but it does for ducklings. The latter take to the water immediately after hatching; the former do not.

### ***Manipulation***

Manipulation presumably evolved in primates, along with bipedal locomotion and the upright posture, by the conversion of the forelimbs from legs into arms and of the forepaws into what we call hands. Walking on two legs, it is sometimes said, leaves the hands free for other acts. The hands are specified by “five-pronged squirming protrusions” into the field of view from below (Chapter 7). They belong to the self, but they are constantly touching the objects of the outer world by reaching and grasping. The shapes and sizes of objects, in fact, are perceived in *relation* to the hands, as graspable or not graspable, in terms of their affordances for manipulation. Infant primates learn to see objects and their hands in conjunction. The perception is constrained by manipulation, and the manipulation is constrained by perception.

## The Control of Locomotion and Manipulation

Locomotion and manipulation, like the movements of the eyes described in the last chapter, are kinds of behavior that cannot be reduced to responses. The persistent effort to do so by physiologists and psychologists has come to a dead end. But the ancient Cartesian doctrine still hangs on, that animals are reflex machines and that humans are the same except for a soul that rules the body by switching impulses at the center of the brain. The doctrine will not do. Locomotion and manipulation are not triggered by stimuli from outside the body, nor are they initiated by commands from inside the brain. Even the classification of incoming impulses in nerves as *sensory* and outgoing impulses as *motor* is based on the old doctrine of mental sensations and physical movements. Neurophysiologists, most of them, are still under the influence of dualism, however much they deny philosophizing. They still assume that the brain is the seat of the mind. To say, in modern parlance, that it is a computer with a program, either inherited or acquired, that plans a voluntary action and then commands the muscles to move is only a little better than Descartes's theory, for to say this is still to remain confined within the doctrine of responses.

Locomotion and manipulation are neither triggered nor commanded but *controlled*. They are constrained, guided, or steered, and only in this sense are they ruled or governed. And they are controlled not by the brain but by information, that is, by seeing oneself in the world. Control lies in the animal-environment system. Control is by the animal *in* its world, the animal itself having subsystems for perceiving the environment and concurrently for getting about in it and manipulating it. The rules that govern behavior are not like laws enforced by an authority or decisions made by a commander; behavior is regular without being regulated. The question is how this can be.

### **WHAT HAPPENS TO INFANT PRIMATES DEPRIVED OF THE SIGHT OF THEIR HANDS?**

Monkeys reared from birth in a device that kept them from seeing the hands and body but not from feeling them move and touching things were very abnormal monkeys. When freed from the device, they acted at first as if they could not reach for and grasp an object but must grope for it. An opaque shield with a cloth bib fitted tightly around the monkey's neck had eliminated visual kinesthesia and had thus prevented the development of visual control of reaching and grasping. So I interpret the results of an experiment by R. Held and J. A. Bauer (1974). See my discussion of the optical information for hand movement in Chapter 7.

### ***The Medium Contains the Information for Control***

It should be kept in mind that animals live in a medium that, being insubstantial, permits them to move about, if supported. We are tempted to call the medium “space,” but the temptation should be resisted. For the medium, unlike space, permits a steady state of reverberating illumination to become established such that it contains information about surfaces and their substances. That is, there is an array at every point of observation and a changing array at every moving point of observation. The medium, as distinguished from space, allows compression waves from a mechanical event, sound, to reach all points of observation and also allows the diffusion field from a volatile substance, odor, to reach them (Gibson, 1966*b*, Ch. 1). The odor is specific to the volatile substance, the sound is specific to the event, and the visual solid angle is the most specific of all, containing all sorts of structured invariants for perceiving the affordance of the object. This is why to perceive something is also to perceive how to approach it and what to do about it.

Information in a medium is not propagated as signals are propagated but is *contained*. Wherever one goes, one can see, hear, and smell. Hence, perception in the medium accompanies locomotion in the medium.

### ***Visual Kinesthesia and Control***

Before getting into the problem of control, we should be clear about the difference between active and passive movement, a difference that is especially important in the case of locomotion. For animal locomotion may be uncontrolled; the animal may be simply transported. This can happen in various ways. A flow of the medium can transport the animal, as happens to the bird in a wind and the fish in a stream. Or an individual may be transported by another animal, as happens to a monkey clinging to its mother or a baby carried in a cradleboard. Or the observer may be a passenger in a vehicle. In all these cases, the animal can *see* its locomotion without initiating, governing, or steering it. The animal has the information for transportation but cannot regulate it. In my terminology, the observer has visual kinesthesia but no visual control of the movement. This distinction is essential to an understanding of the problem of control. The traditional theory of the senses is incapable of making it, however, and followers of the traditional theory become mired in the conceptual confusion arising from the slippery notion of feedback.

Visual kinesthesia specifies locomotion relative to the environment, whereas the other kinds of kinesthesia may or may not do so. The control of locomotion in the environment must therefore be visual. Walking, bicycling, and driving involve very different kinds of classical kinesthesia but the same visual kinesthesia. The muscle movements must be governed by vision. If you want to go somewhere, or to know where you are going, you can only trust your eyes. The

bird in a wind even has to fly in order to stay in the same place. To prevent being carried away, it must arrest the flow of the ambient array.

Before we can hope to understand controlled locomotion, therefore, we must answer several preliminary questions about the information in ambient light. I can think of four. What specifies locomotion or stasis? What specifies an obstacle or an opening? What specifies imminent contact with a surface? What specifies the benefit or the injury that lies ahead? These questions must be answered before we can begin to ask what the *rules* are for starting and stopping, for approaching and retreating, for going this way or that way, and so on.

## The Optical Information Necessary for Control of Locomotion

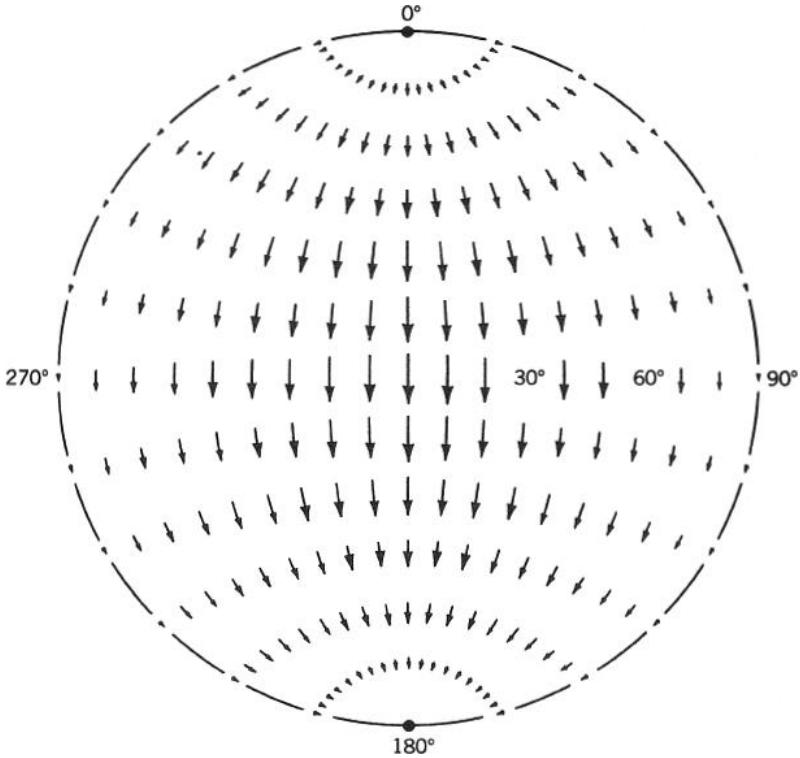
For each of the four questions above, I shall list a number of assertions about optical information. I will try to put together what the previous chapters have established.

### *What Specifies Locomotion or Stasis?*

1. *Flow of the ambient array specifies locomotion, and nonflow specifies stasis.* By *flow* is meant the change analyzed as *motion perspective* (Gibson, Olum, and Rosenblatt, 1955) for the abstract case of an uncluttered environment and a moving point of observation. A better term would be *flow perspective*, or *streaming perspective*. It yields the “melon-shaped family of curves” illustrated in Figure 13.1 and is based on rays of light from particles of the terrain, not on solid angles from features of the terrain. Thus, it has the great advantages of geometrical analysis but also has its disadvantages. Nevertheless, the flow as such specifies locomotion and the invariants specify the layout of surfaces in which locomotion occurs.

2. *Outflow specifies approach to and inflow specifies retreat from.* An invariant feature of the ambient flow is that one hemisphere is centrifugal and the other centripetal. Outflow entails magnification, and inflow entails minification. There is always both a going-to and a coming-from during locomotion. A creature with semipanoramic vision can register both the outflow and the inflow at the same time, but human creatures can sample only one or the other, by looking “ahead” or by looking “behind.” Note that a reversal of the flow pattern specifies a reversal of locomotion.

3. *The focus or center of outflow specifies the direction of locomotion in the environment.* More exactly, that visual solid angle at the center of outflow specifies the surface in the environment, or the object, or the opening, toward which the animal is moving. This statement is not analytical. Because the overall flow is radial in both hemispheres, the two foci are implicit in any sufficiently large sample of the ambient array, and even humans can thus *see* where they are going without having to *look* where they are going. The “melon-shaped family of curves” continues outside the edges of the temporary field of view.



**FIGURE 13.1** The flow velocities in the lower hemisphere of the ambient optic array with locomotion parallel to the earth.

The vectors are plotted in angular coordinates, and all vectors vanish at the horizon. This drawing should be compared with Figure 7.3 showing the motion perspective to a flying bird. (From Gibson, Olum, and Rosenblatt, 1955. © 1955 by the Board of Trustees of the University of Illinois. Reprinted by permission of the University of Illinois Press.)

4. *A shift of the center of outflow from one visual solid angle to another specifies a change in the direction of locomotion, a turn, and a remaining of the center within the same solid angle specifies no change in direction.* The ambient optic array is here supposed to consist of nested solid angles, not of a bundle of lines. The direction of locomotion is thus anchored to the layout, not to a coordinate system. The flow of the ambient array can be *transposed over the invariant structure of the array*, so that where one is going is seen relative to the surrounding layout. This unfamiliar notion of invariant structure underlying the changing perspective structure is one that I tried to make explicit in Chapter 5; here is a good example of it. The illustrations in Chapter 7 showing arrows superposed on a picture of

the terrain were supposed to suggest this invariance under change but, of course, it cannot be pictured.

5. *Flow of the textured ambient array just behind certain occluding protrusions into the field of view specifies locomotion by an animal with feet.* If you lower your head while walking, a pair of moving protrusions enters the field of view from its lower edge (Chapter 7), and these protrusions move up and down alternately. A cat sees the same thing except that what it sees are *front* feet. The extremities are in optical contact with the flowing array at the locus of maximal flow and maximally coarse texture. They occlude parts of the surface, but it is seen to extend behind them. Convexities and concavities in the surface will affect the timing of contact, and therefore you and the cat must place your feet with regard to the *footing*.

### ***What Specifies an Obstacle or an Opening?***

I distinguish two general cases for the affording of locomotion, which I will call *obstacle* and *opening*. An *obstacle* is a rigid object, detached or attached, a surface with occluding edges. An *opening* is an aperture, hole, or gap in a surface, also with occluding edges. An obstacle affords collision. An opening affords passage. Both have a closed or nearly closed contour in the optic array, but the edge of the obstacle is inside the contour, whereas the edge of the opening is outside the contour. A round object hides in one direction, and a round opening hides in the opposite direction. The way to tell the difference between an obstacle and an opening, therefore, is as follows.

#### **ON LOOKING AT THE ROAD WHILE DRIVING**

It must be admitted that when I turn around while driving our car and reply to my wife's protests that I can perfectly well see where I am going without having to look where I am going because the focus of outflow is implicit, she is not reassured.

6. *Loss (or gain) of structure outside a closed contour during approach (or retreat) specifies an obstacle. Gain (or loss) of structure inside a closed contour during approach (or retreat) specifies an opening.* This is the only absolutely trustworthy way to tell the difference between an obstacle and an opening. In both cases the visual solid angle goes to a hemisphere as you approach it, but you collide with the obstacle and enter the opening. Magnification of the form as such, the outline, does not distinguish them. But as you come up to the obstacle it hides more and more of the vista, and as you come up to the opening it reveals more and more of the vista. Deletion outside the occluding edge and accretion inside the occluding



edge will distinguish the two. Psychologists and artists alike have been confused about the difference between things and holes, surfaces and apertures. The figure-ground phenomenon that so impressed the gestalt psychologists and that is still taken to be a prototype of perception is misleading. A closed contour as such in the optic array does *not* specify an object in the environment.

What specifies the near edge of an opening in the ground, a hole or gap in the surface of support? This is very important information for a terrestrial animal.

7. *Gain of structure above a horizontal contour in the ambient array during approach specifies a brink in the surface of support. A brink is a drop-off in the ground, a step, or the edge of a perch. It is the essential feature of the experiments on the visual cliff that were described in Chapter 9 (for example, E. J. Gibson and Walk, 1960). It is depth downward at an occluding edge, and depending on the amount of depth relative to the size of the animal, it affords stepping-down or falling-off. The rat, chick, or human infant who sees its feet close to such an occluding edge needs to take care. The experimental evidence suggests that the changing occlusion at the edge, not the abrupt increase in the density of optical texture, is the effective information for the animal.*

This formula applies to a *horizontal* contour in the array coming from the ground. What about a *vertical* contour in the array coming from a wall?

8. *Gain of structure on one side of a vertical contour in the ambient array during approach specifies the occluding edge of a barrier, and the side on which gain occurs is the side of the edge that affords passage. This is the edge of a house, the end of a wall, or the vertical edge of a doorway, often loosely called a corner. On one side of the edge the vista beyond is hidden, and on the other side it is revealed; on one side there is potential collision, and on the other potential passage. The trunk of a tree has two such curved edges not far apart. To “go around the corner” is to reveal the surfaces of the new vista. Rats do it in mazes, and people do it in cities. To find one’s way in a cluttered environment is to go around a series of occluding edges, and the problem is to choose the correct edges to go around (see Figure 11.2).*

### ***What Specifies Imminent Contact with a Surface?***

In an early essay on the visual control of locomotion (Gibson, 1958), I wrote:

Approach to a solid surface is specified by a centrifugal flow of the texture of the optic array. Approach to an object is specified by a magnification of the closed contour in the array corresponding to the edges of the object. A *uniform* rate of approach is accompanied by an *accelerated* rate of magnification. At the theoretical point where the eye touches the object, the latter will intercept a visual angle of  $180^\circ$ . The magnification reaches an explosive rate in the last moments before contact. This accelerated expansion . . . specifies imminent collision.

This was true enough as far as it went. I was thinking of the problem of how a pilot lands on a field or how a bee lands on a flower. The explosive magnification, the “looming” as I called it, has to be canceled if a “soft” landing is to be achieved. I never thought of the entirely different problem of steering through an opening. The optical information provided by various kinds of magnification is evidently not as simple as I thought in 1958.

The complexities were not clarified by the empirical studies of Schiff, Caviness, and Gibson (1962) and Schiff (1965), who provided the optical information for the approach *of* an object in space instead of the information for approach *to* a surface in the environment. They displayed an expanding dark silhouette in the center of a luminous translucent screen, as described in Chapter 10. No one saw himself being transposed; everyone saw something indefinite coming toward them, as if it were in the sky. The display consisted of an expanding single form, a shadow or silhouette, not the magnifying of a nested structure of subordinate forms that characterizes approach to a real surface. The magnifying of detail *without limit* was missing from the display.

9. *The magnification of a nested structure in which progressively finer details keep emerging at the center specifies approach of an observer to a surface in the environment.* This formula emphasizes the facets within the faces of a substantial surface, such as that of an obstacle, an object, an animate object, or a surface of rest that the observer might encounter. In order to achieve contact without collision, the nested magnification must be made to cease at the appropriate level instead of continuing to its limit. There seems to be an optimal degree of magnification for contact with a surface, depending on what it affords. For food one moves up to *eating* distance; for manipulating one moves up to *reaching* distance; for print one moves up to *reading* distance.

### ***What Specifies the Benefit or Injury that Lies Ahead?***

Bishop Berkeley suggested in 1709 that the chief end of vision was for animals “to foresee the benefit or injury which is like to ensue upon the application of their own bodies to this or that body which is at a distance.” What the philosopher called foresight is what I call the *perception of the affordance*. To see at a distance what the object affords on contact is “necessary for the preservation of an animal.”

I differ from Bishop Berkeley in assuming that information is available in the light to the animal for what an encounter with the object affords. But I agree with him about the utility of vision.

10. *Affordances for the individual upon encountering an object are specified in the optic array from the object by invariants and invariant combinations. Tools, food, shelter, mates, and amiable animals are distinguished from poisons, fires, weapons, and hostile animals by their shapes, colors, textures, and deformations.* The positive and negative affordances of things in the environment are what makes locomotion through

the medium such a fundamental kind of behavior for animals. Unlike a plant, the animal can go to the beneficial and stay away from the injurious. But it must be able to perceive the affordances from afar. A rule for the visual control of locomotion might be this: so move as to obtain beneficial encounters with objects and places and to prevent injurious encounters.

## Rules for the Visual Control of Locomotion

I suggested at the beginning that behavior was controlled by information about the world and the self conjointly. The information has now been described. What about the control?

I asserted that behavior was controlled by *rules*. Surely, however, they are not rules enforced by an authority. The rules are not commands from a brain; they emerge from the animal–environment system. But the only way to describe rules is in words, and a rule expressed in words is a command. I am faced with a paradox. The rules for the control of locomotion will sound like commands, although they are not intended to. I can only suggest that the reader should interpret them as rules *not formulated in words*.

The rules that follow are for *visual control*, not muscular, articular, vestibular, or cutaneous control. The visual system normally supersedes the haptic system for locomotion and manipulation, as I tried to explain in *The Senses Considered as Perceptual Systems* (Gibson, 1966*b*). This means that the rules for locomotion will be the same for crawling on all fours, walking, running, or driving an automobile. The particular muscles involved do not matter. Any group of muscles will suffice if it brings about the relation of the animal to its environment stated in the rule.

*Standing.* The basic rule for a pedestrian animal is *stand up*; that is, keep the feet in contact with a surface of support. It is also well to keep the oval boundaries of the field of view normal with the implicit horizon of the ambient array; if the head is upright the rest of the body follows.

*Starting, stopping, going back.* *To start, make the array flow. To stop, cancel the flow. To go back, make the flow reverse.* According to the first two formulas listed in the previous pages, to cause outflow is to get closer and to cause inflow is to get farther away.

*Steering.* *To turn, shift the center of outflow from one patch in the optic array to another,* according to the the third and fourth formulas. Steering requires that openings be distinguished from barriers, obstacles, and brinks. The rule is: *To steer, keep the center of outflow outside the patches of the array that specify barriers, obstacles, and brinks and within a patch that specifies an opening* (sixth, seventh, and eighth formulas). Following this rule will avert collisions and prevent falling off.

*Approaching.* To approach is to magnify a patch in the array, but magnification is complicated (formulas two and six). There are many rules involving magnification. Here are a few. *To permit scrutiny, magnify the patch in the array to such a*

*degree that the details can be looked at. To manipulate something graspable, magnify the patch to such a degree that the object is within reach. To bite something, magnify the patch to such an angle that the mouth can grasp it. To kiss someone, magnify the face-form, if the facial expression is amiable, so as almost to fill the field of view. (It is absolutely essential for one to keep one's eyes open so as to avoid collision. It is also wise to learn to discriminate those subtle invariants that specify amiability.) To read something, magnify the patch to such a degree that the letters become distinguishable. The most general rule for approach is this: To realize the positive affordances of something, magnify its optical structure to that degree necessary for the behavioral encounter.*

*Entering enclosures.* An enclosure such as a burrow, cave, nest, or hut affords various benefits upon entry. It is a place of warmth, a shelter from rain and wind, and a place for sleep. It is often a home, the place where mate and offspring are. It is also a place of safety, a hiding place affording both concealment from enemies and a barrier to their locomotion. An enclosure must have an opening to permit entry, and the opening must be identified. The rule seems to be as follows: *to enter an enclosure, magnify the angle of its opening to 180° and open up the vista. Make sure that there is gain of structure inside the contour and not loss outside, or else you will collide with an obstacle* (formulas six and nine).

*Keeping a safe distance.* The opposite of approach is retreat. Psychologists have sometimes assumed that the *alternative* to approach is retreat. Kurt Lewin's theory of behavior, for example, was based on approach to an object with a positive "valence" and retreat from an object with a negative "valence." This fits with a theory of conflict between approach and retreat, and a compromise between opposite tendencies. But it is wrong to assume that approach and retreat are alternatives. There is no need to flee from an obstacle, a barbed-wire fence, the edge of a river, the edge of a cliff, or a fire. The only need is to maintain a safe distance, a "margin of safety," since these things do not pursue the observer. A ferocious tiger has a negative valence, but a cliff does not. The rule is this, I think: *To prevent an injurious encounter, keep the optical structure of the surface from magnifying to the degree that specifies an encounter* (formulas two and ten).

For moving predators and enemies, *flight* is an appropriate form of action since they can approach. The rule for flight is, *so move as to minify the dangerous form and to make the surrounding optic array flow inward*. If, despite flight, the form magnifies, the enemy is catching up; if it minifies, one is getting away. At the predator's point of observation, of course, the rule is opposite to that for the prey: *so move as to magnify the succulent form by making the surrounding array flow outward until it reaches the proper angular size for capturing*.

## Rules for the Visual Control of Manipulation

The rules for the visual control of the movements of the hands are more complex than those for the control of locomotion. But the human infant who

watches these squirming protuberances into his field of view is not formulating rules and, in any case, complexity does not seem to cause trouble for the nervous system. I am unable to formulate the rules in words except for a few easy cases.

Locomotor approach often terminates in reaching and grasping. *Reaching* is an elongation of the arm-shape and a minification of the five-pronged hand-shape until contact occurs. If the object is hand-size, it is graspable; if too large or too small, it is not. Children learn to see sizes in terms of prehension: they see the span of their grasp and the diameter of a ball at the same time (Gibson, 1966*b*, fig. 7.1, p. 119). Long before the child can discriminate one inch, or two, or three, he can see the fit of the object to the pincerlike action of the opposable thumb. The child learns his scale of sizes as commensurate with his body, not with a measuring stick.

The affordance of an elongated object for pounding and striking is easily learned. The skill of hammering or striking a target requires visual control, however. It involves what we vaguely call *aiming*. I will not try to state the rules for aiming except to suggest that it entails a kind of centering or symmetricalizing of a diminishing form on a fixed form.

*Throwing* as such is easy. Simply cause the visual angle of the object you have in your hand to shrink, and it will “zoom” in a highly interesting manner. You have to let go, of course, and this is a matter of haptic control, not visual control. Aimed throwing is much harder, as ballplayers know. It is a sort of reciprocal of steered locomotion.

*Tool-using* in general is rule governed. The rule for pliers is analogous to that for prehending, the tool being metaphorically an extension of the hand. The use of a stick as a rake for getting a banana outside the cage was one of the achievements of a famous chimpanzee (Köhler, 1925).

Knives, axes, and pointed objects afford the cutting and piercing of other objects and surfaces, including other animals. But the manipulation must be carefully controlled, for the observer’s own skin can be cut or pierced as well as the other surface. The tool must be grasped by the handle, not the point; that is, the rule for reaching and the rules for maintaining the margin of safety must both be followed. Visual contact with one part of the surface is beneficial but with another part is injurious, and the “sharp” part is not always easy to discriminate. The case is similar to that of walking along a cliff edge in this respect: one must steer the movement so as to skirt the danger.

The uses of the hands are almost unlimited. And manipulation subserves many other forms of behavior of which it is only a part, eating, drinking, transporting, nursing, caressing, gesturing, and the acts of trace-making, depicting, and writing, which will concern us in Part IV.

The point to remember is that the visual control of the hands is inseparably connected with the visual perception of objects. The act of throwing complements the perception of a throwable object. The transporting of things is part and parcel of seeing them as portable or not.

*Conclusion about manipulation.* One thing should be evident. The movements of the hands do not consist of responses to stimuli. Manipulation cannot be understood in those terms. Is the only alternative to think of the hands as instruments of the mind? Piaget, for example, sometimes seems to imply that the hands are tools of a child's intelligence. But this is like saying that the hand is a tool of an inner child in more or less the same way that an object is a tool for a child with hands. This is surely an error. The alternative is not a return to mentalism. We should think of the hands as neither triggered nor commanded but *controlled*.

## Manipulation and the Perceiving of Interior Surfaces

Finally, it should be noted that a great deal of manipulation occurs for the sake of perceiving hidden surfaces. I can think of three kinds of such manipulation: *opening up*, *uncovering*, and *taking apart*. Each of these has an opposite, as one would expect from the law of reversible occlusion: *closing*, *covering*, and *putting together*.

*Opening* and *closing* apply to the lids and covers of hollow objects and also to drawers, compartments, cabinets, and other enclosures. Children are fascinated by the act of opening so as to reveal the interior and closing so as to conceal it. They then come to perceive the continuity between the inner and the outer surfaces. The closed box and the covered pot are then seen to have an inside as well as an outside.

*Covering* and *uncovering* apply to a cloth, or a child's blanket, or to revealing and concealing by an opaque substance, as in a sandbox. The movement of the hand that conceals the object is not always so clearly the reverse of the movement that reveals it as it is in the case of closing-opening, however. The perceiving of hidden surfaces may well be more difficult in this case.

*Taking apart* and *putting together* apply to an object composed of smaller objects, that is, a composite that can be disassembled and assembled. There are toys of this sort. Blocks that can be fitted together make such a composite object. Taking apart is usually a simpler act of manipulation than putting together. Children need to see what is inside these compound objects, and it is only to be expected that they should take them apart, or break them apart if need be. After such visual-manual cooperation, they can perceive the interior surfaces of the object together with the cracks, joins, and apertures that separate them. This is the way children come to apprehend a mechanism such as a clock or an internal combustion engine.

## Summary

Active locomotor behavior, as contrasted with passive transportation, is under the continuous control of the observer. The dominant level of such control is visual. But this could not occur without what I have called *visual kinesthesia*, the awareness of movement or stasis, of starting or stopping, of approaching or

retreating, of going in one direction or another, and of the imminence of an encounter. Such awarenesses are necessary for control.

Also necessary is an awareness of the affordance of the encounter that will terminate the locomotor act and of the affordances of the openings and obstacles, the brinks and barriers, and the corners on the way (actually the occluding edges).

When locomotion is thus visually controlled, it is regular without being a chain of responses and is purposive without being commanded from within.

Manipulation, like active locomotion, is visually controlled. It is thus dependent on an awareness of both the hands as such and the affordances for handling. But its regularities are not so easy to formulate.

# 14

## THE THEORY OF INFORMATION PICKUP AND ITS CONSEQUENCES

In this book the traditional theories of perception have been abandoned. The perennial doctrine that two-dimensional images are restored to three-dimensional reality by a process called depth perception will not do. Neither will the doctrine that the images are transformed by the cues for distance and slant so as to yield constancy of size and shape in the perception of objects. The deep-seated notion of the retinal image as a still picture has been abandoned.

The simple assumption that perceptions of the world are caused by stimuli from the world will not do. The more sophisticated assumption that perceptions of the world are caused when sensations triggered by stimuli are supplemented by memories will not do either. Not even the assumption that a sequence of stimuli is converted into a phenomenal scene by memory will do. The very notion of stimulation as typically composed of discrete stimuli has been abandoned.

The established theory that exteroception and proprioception arise when exteroceptors and proprioceptors are stimulated will not do. The doctrine of special channels of sensation corresponding to specific nerve bundles has been abandoned.

The belief of empiricists that the perceived meanings and values of things are supplied from the past experience of the observer will not do. But even worse is the belief of nativists that meanings and values are supplied from the past experience of the race by way of innate ideas. The theory that meaning is attached to experience or imposed on it has been abandoned.

Not even the current theory that the inputs of the sensory channels are subject to “cognitive processing” will do. The inputs are described in terms of information theory, but the processes are described in terms of old-fashioned mental acts: recognition, interpretation, inference, concepts, ideas, and storage and retrieval of ideas. These are still the operations of the mind upon the deliverances



of the senses, and there are too many perplexities entailed in this theory. It will not do, and the approach should be abandoned.

What sort of theory, then, will explain perception? Nothing less than one based on the pickup of information. To this theory, even in its undeveloped state, we should now turn.

Let us remember once again that it is the perception of the environment that we wish to explain. If we were content to explain only the perception of forms or pictures on a surface, of nonsense figures to which meanings must be attached, of discrete stimuli imposed on an observer willy-nilly, in short, the items most often presented to an observer in the laboratory, the traditional theories might prove to be adequate and would not have to be abandoned. But we should not be content with that limited aim. It leaves out of account the eventful world and the perceiver's awareness of being in the world. The laboratory does not have to be limited to simple stimuli, so-called. The experiments reported in Chapters 9 and 10 showed that information can be displayed.

### **What is New About the Pickup of Information?**

The theory of information pickup differs radically from the traditional theories of perception. First, it involves a new notion of perception, not just a new theory of the process. Second, it involves a new assumption about what there is to be perceived. Third, it involves a new conception of the information for perception, with two kinds always available, one about the environment and another about the self. Fourth, it requires the new assumption of perceptual systems with overlapping functions, each having outputs to adjustable organs as well as inputs from organs. We are especially concerned with vision, but none of the systems, listening, touching, smelling, or tasting, is a channel of sense. Finally, fifth, optical information pickup entails an activity of the system not heretofore imagined by any visual scientist, the concurrent registering of both persistence and change in the flow of structured stimulation. This is the crux of the theory but the hardest part to explicate, because it can be phrased in different ways and a terminology has to be invented.

Consider these five novelties in order, ending with the problem of detecting variants and invariants or change and nonchange.

### ***A Redefinition of Perception***

Perceiving is an achievement of the individual, not an appearance in the theater of his consciousness. It is a keeping-in-touch with the world, an experiencing of things rather than a having of experiences. It involves awareness-of instead of just awareness. It may be awareness of something in the environment or something in the observer or both at once, but there is no content of awareness independent of that of which one is aware. This is close to the act psychology of

the nineteenth century except that perception is not a mental act. Neither is it a bodily act. Perceiving is a psychosomatic act, not of the mind or of the body but of a living observer.

The act of picking up information, moreover, is a continuous act, an activity that is ceaseless and unbroken. The sea of energy in which we live flows and changes without sharp breaks. Even the tiny fraction of this energy that affects the receptors in the eyes, ears, nose, mouth, and skin is a flux, not a sequence. The exploring, orienting, and adjusting of these organs sink to a minimum during sleep but do not stop dead. Hence, perceiving is a stream, and William James's description of the stream of consciousness (1890, Ch. 9) applies to it. Discrete percepts, like discrete ideas, are "as mythical as the Jack of Spades."

The continuous act of perceiving involves the coperceiving of the self. At least, that is one way to put it. The very term *perception* must be redefined to allow for this fact, and the word *proprioception* must be given a different meaning than it was given by Sherrington.

### ***A New Assertion About What is Perceived***

My description of the environment (Chapters 1–3) and of the changes that can occur in it (Chapter 6) implies that places, attached objects, objects, and substances are what are mainly perceived, together with events, which are changes of these things. To see these things is to perceive what they afford. This is very different from the accepted categories of what there is to perceive as described in the textbooks. Color, form, location, space, time, and motion—these are the chapter headings that have been handed down through the centuries, but they are not what is perceived.

### ***Places***

A *place* is one of many adjacent places that make up the habitat and, beyond that, the whole environment. But smaller places are nested within larger places. They do not have boundaries, unless artificial boundaries are imposed by surveyors (my piece of land, my town, my country, my state). A place at one level is what you can see from here or hereabouts, and locomotion consists of going from place to place in this sense (Chapter 11). A very important kind of learning for animals and children is place-learning—learning the affordances of places and learning to distinguish among them—and way-finding, which culminate in the state of being oriented to the whole habitat and knowing where one is in the environment.

A place persists in some respects and changes in others. In one respect, it cannot be changed at all—in its location relative to other places. A place cannot be *displaced* like an object. That is, the adjacent order of places cannot be permuted; they cannot be shuffled. The sleeping places, eating places, meeting

places, hiding places, and falling-off places of the habitat are immobile. Place-learning is therefore different from other kinds.

### *Attached Objects*

I defined an *object* in Chapter 3 as a substance partially or wholly surrounded by the medium. An object attached to a place is only partly surrounded. It is a protuberance. It cannot be displaced without becoming detached. Nevertheless, it has a surface and enough of a natural boundary to constitute a unit. Attached objects can thus be counted. Animals and children learn what such objects are good for and how to distinguish them. But they cannot be separated from the places where they are found.

### *Detached Objects*

A fully detached object can be displaced or, in some cases, can displace itself. Learning to perceive it thus has a different character from learning to perceive places and attached objects. Its affordances are different. It can be put side by side with another object and compared. It can therefore be grouped or classed by the manipulation of sorting. Such objects when grouped can be rearranged, that is, permuted. And this means not only that they can be counted but that an abstract number can be assigned to the group.

It is probably harder for a child to perceive “same object in a different place” than it is to perceive “same object in the same place.” The former requires that the information for persistence-despite-displacement should have been noticed, whereas the latter does not.

Inanimate detached objects, rigid or nonrigid, natural or manufactured, can be said to have features that distinguish them. The features are probably not denumerable, unlike the objects themselves. But if they are compounded to specify affordances, as I argued they must be, only the relevant compounds need to be distinguished. So when it comes to the natural, nonrigid, animate objects of the world whose dimensions of difference are overwhelmingly rich and complex, we pay attention only to what the animal or person affords (Chapter 8).

### *Persisting Substances*

A *substance* is that of which places and objects are composed. It can be vaporous, liquid, plastic, viscous, or rigid, that is, increasingly “substantial.” A substance, together with what it affords, is fairly well specified by the color and texture of its surface. Smoke, milk, clay, bread, and wood are polymorphic in layout but invariant in color-texture. Substances, of course, can be smelled and tasted and palpated as well as seen.

The animal or child who begins to perceive substances, therefore, does so in a different way than one who begins to perceive places, attached objects, and detached objects. Substances are formless and cannot be counted. The number of substances, natural compositions, or mixtures is not fixed. (The number of chemical elements is fixed, but that is a different matter.) We discriminate among surface colors and textures, but we cannot group them as we do detached objects and we cannot order them as we do places.

We also, of course, perceive changes in otherwise persisting substances, the ripening of fruit, and the results of boiling and baking, or of mixing and hardening. But these are a kind of event.

### *Events*

As I used the term, an *event* is any change of a substance, place, or object, chemical, mechanical, or biophysical. The change may be slow or fast, reversible or nonreversible, repeating or nonrepeating. Events include what happens to objects in general, plus what the animate objects *make* happen. Events are nested within superordinate events. The motion of a detached object is not the prototype of an event that we have been led to think it was. Events of different sorts are perceived as such and are not, surely, reducible to elementary motions.

### *The Information for Perception*

*Information*, as the term is used in this book (but not in other books), refers to specification of the observer's environment, not to specification of the observer's receptors or sense organs. The qualities of objects are specified by information; the qualities of the receptors and nerves are specified by sensations. Information about the world cuts right across the qualities of sense.

The term *information* cannot have its familiar dictionary meaning of *knowledge communicated to a receiver*. This is unfortunate, and I would use another term if I could. The only recourse is to ask the reader to remember that picking up information is not to be thought of as a case of communicating. The world does not speak to the observer. Animals and humans communicate with cries, gestures, speech, pictures, writing, and television, but we cannot hope to understand perception in terms of these channels; it is quite the other way around. Words and pictures convey information, carry it, or transmit it, but the information in the sea of energy around each of us, luminous or mechanical or chemical energy, is not conveyed. It is simply there. The assumption that information can be transmitted and the assumption that it can be stored are appropriate for the theory of communication, not for the theory of perception.

The vast area of speculation about the so-called media of communication had a certain discipline imposed on it some years ago by a mathematical theory of communication (Shannon and Weaver, 1949). A useful measure of information

transmitted was formulated, in terms of “bits.” A sender and receiver, a channel, and a finite number of possible signals were assumed. The result was a genuine discipline of communications engineering. But, although psychologists promptly tried to apply it to the senses and neuropsychologists began thinking of nerve impulses in terms of bits and the brain in terms of a computer, the applications did not work. Shannon’s concept of information applies to telephone hookups and radio broadcasting in elegant ways but not, I think, to the firsthand perception of being in-the-world, to what the baby gets when first it opens its eyes. The information for perception, unhappily, cannot be defined and measured as Claude Shannon’s information can be.

The information in ambient light, along with sound, odor, touches, and natural chemicals, is inexhaustible. A perceiver can keep on noticing facts about the world she lives in to the end of her life without ever reaching a limit. There is no threshold for information comparable to a stimulus threshold. Information is not lost to the environment when gained by the individual; it is not conserved like energy.

Information is not specific to the banks of photoreceptors, mechanoreceptors, and chemoreceptors that lie within the sense organs. Sensations are specific to receptors and thus, normally, to the kinds of stimulus energy that touch them off. But information is not energy-specific. Stimuli are not always imposed on a passive subject. In life one *obtains* stimulation in order to extract the information (Gibson, 1966b, Ch. 2). The information can be the same, despite a radical change in the stimulation obtained.

Finally, a concept of information is required that admits of the possibility of illusion. Illusions are a theoretical perplexity in any approach to the study of perception. Is information always valid and illusion simply a failure to pick it up? Or is the information picked up sometimes impoverished, masked, ambiguous, equivocal, contradictory, even false? The puzzle is especially critical in vision.

In Chapter 14 of *The Senses Considered as Perceptual Systems* (Gibson, 1966b) and again in this book I have tried to come to terms with the problem of misperception. I am only sure of this: it is not one problem but a complex of different problems. Consider, first, the mirage of palm trees in the desert sky, or the straight stick that looks bent because it is partly immersed in water. These illusions, together with the illusion of Narcissus, arise from the regular reflection or refraction of light, that is, from exceptions to the ecological optics of the scatter-reflecting surface and the perfectly homogeneous medium. Then consider, second, the misperception in the case of the shark under the calm water or the electric shock hidden in the radio cabinet. Failure to perceive the danger is not then blamed on the perceiver. Consider, third, the sheet of glass mistaken for an open doorway or the horizontal sheet of glass (the optical cliff) mistaken for a void. A fourth case is the room composed of trapezoidal surfaces or the trapezoidal window, which look normally rectangular so long as the observer does not open both eyes and walk around. Optical misinformation

enters into each of these cases in a different way. But in the last analysis, *are* they explained by misinformation? Or is it a matter of failure to pick up *all* the available information, the inexhaustible reservoir that lies open to further scrutiny?

The misperceiving of affordances is a serious matter. As I noted in Chapter 8, a wildcat may look like a cat. (But *does* he look just like a cat?) A malevolent man may act like a benevolent one. (But *does* he exactly?) The line between the pickup of misinformation and the failure to pick up information is hard to draw.

Consider the human habit of picture-making, which I take to be the devising and displaying of optical information for perception by others. It is thus a means of communication, giving rise to mediated apprehension, but it is more like direct pickup than word-making is. Depiction and its consequences are deferred until later, but it can be pointed out here that picture-makers have been experimenting on us for centuries with artificial displays of information in a special form. They enrich or impoverish it, mask or clarify it, ambiguate or disambiguate it. They often try to produce a discrepancy of information, an equivocation or contradiction, in the same display. Painters invented the cues for depth in the first place, and psychologists looked at their paintings and began to talk about cues. The notions of counterbalanced cues, of figure-ground reversals, of equivocal perspectives, of different perspectives on the same object, of “impossible” objects—all these come from artists who were simply experimenting with frozen optical information.

An important fact to be noted about any pictorial display of optical information is that, in contrast with the inexhaustible reservoir of information in an illuminated medium, it cannot be looked at close up. Information to specify the display as such, the canvas, the surface, the screen, can always be picked up by an observer who walks around and looks closely.

### ***The Concept of a Perceptual System***

The theory of information pickup requires perceptual systems, not senses. Some years ago I tried to prove that a perceptual system was radically different from a sense (Gibson, 1966*b*), the one being active and the other passive. People said, “Well, what I mean by a sense is an *active* sense.” But it turned out that they still meant the passive inputs of a sensory nerve, the activity being what occurs in the brain when the inputs get there. That was not what I meant by a perceptual system. I meant the activities of looking, listening, touching, tasting, or sniffing. People then said, “Well, but those are responses to sights, sounds, touches, tastes, or smells, that is, motor acts resulting from sensory inputs. What you call a perceptual system is nothing but a case of feedback.” I was discouraged. People did not understand.

I shall here make another attempt to show that the senses considered as special senses cannot be reconciled with the senses considered as perceptual systems. The five perceptual systems correspond to five modes of overt attention. They

have overlapping functions, and they are all more or less subordinated to an overall orienting system. A system has organs, whereas a sense has receptors. A system can orient, explore, investigate, adjust, optimize, resonate, extract, and come to an equilibrium, whereas a sense cannot. The characteristic activities of the visual system have been described in Chapter 12 of this book. The characteristic activities of the auditory system, the haptic system, and the two related parts of what I called the “chemical value system” were described in Chapters 5–8 of my earlier book (Gibson, 1966*b*). Five fundamental differences between a sense and a perceptual system are given below.

1. A special sense is defined by a bank of receptors or receptive units that are connected with a so-called projection center in the brain. Local stimuli at the sensory surface will cause local firing of neurons in the center. The adjustments of the organ in which the receptors are incorporated are not included within the definition of a sense.

A perceptual system is defined by an organ and its adjustments at a given level of functioning, subordinate or superordinate. At any level, the incoming and outgoing nerve fibers are considered together so as to make a continuous loop.

The organs of the visual system, for example, from lower to higher are roughly as follows. First, the lens, pupil, chamber, and retina comprise an organ. Second, the eye with its muscles in the orbit comprise an organ that is both stabilized and mobile. Third, the two eyes in the head comprise a binocular organ. Fourth, the eyes in a mobile head that can turn comprise an organ for the pickup of ambient information. Fifth, the eyes in a head on a body constitute a superordinate organ for information pickup over paths of locomotion. The adjustments of accommodation, intensity modulation, and dark adaptation go with the first level. The movements of compensation, fixation, and scanning go with the second level. The movements of vergence and the pickup of disparity go with the third level. The movements of the head, and of the body as a whole, go with the fourth and fifth levels. All of them serve the pickup of information.

2. In the case of a special sense, the receptors can only receive stimuli, passively, whereas in the case of a perceptual system the input-output loop can be supposed to obtain information, actively. Even when the theory of the special senses is liberalized by the modern hypothesis of receptive units, the latter are supposed to be triggered by complex stimuli or modulated in some passive fashion.

3. The inputs of a special sense constitute a repertory of innate sensations, whereas the achievements of a perceptual system are susceptible to maturation and learning. Sensations of one modality can be combined with those of another in accordance with the laws of association; they can be organized or fused or supplemented or selected, but *no new sensations can be learned*. The information that is picked up, on the other hand, becomes more and more subtle, elaborate, and precise with practice. One can keep on learning to perceive as long as life goes on.

4. The inputs of the special senses have the qualities of the receptors being stimulated, whereas the achievements of the perceptual systems are specific to the qualities of things in the world, especially their affordances. The recognition of this limitation of the senses was forced upon us by Johannes Müller with his doctrine of specific “nerve energies.” He understood clearly, if reluctantly, the implication that, because we can never know the external causes of our sensations, we cannot know the outer world. Strenuous efforts have to be made if one is to avoid this shocking conclusion. Helmholtz argued that we must deduce the causes of our sensations because we cannot detect them. The hypothesis that sensations provide clues or cues for perception of the world is similar. The popular formula that we can interpret sensory signals is a variant of it. But it seems to me that all such arguments come down to this: we can perceive the world only if we already know what there is to be perceived. And that, of course, is circular. I shall come back to this point again.

The alternative is to assume that sensations triggered by light, sound, pressure, and chemicals are merely incidental, that information is available to a perceptual system, and that the qualities of the world in relation to the needs of the observer are experienced directly.

5. In the case of a special sense the process of attention occurs at centers within the nervous system, whereas in the case of a perceptual system attention pervades the whole input-output loop. In the first case attention is a consciousness that can be focused; in the second case it is a skill that can be educated. In the first case physiological metaphors are used, such as the filtering of nervous impulses or the switching of impulses from one path to another. In the second case the metaphors used can be terms such as *resonating*, *extracting*, *optimizing*, or *symmetricalizing* and such acts as orienting, exploring, investigating, or adjusting.

I suggested in Chapter 12 that a normal act of visual attention consists of scanning a whole feature of the ambient array, not of fixating a single detail of the array. We are tempted to think of attention as strictly a narrowing-down and holding-still, but actually this is rare. The invariants of structure in an optic array that constitute information are more likely to be gradients than small details, and they are scanned over wide angles.

### ***The Registering of Both Persistence and Change***

The theory of information pickup requires that the visual system be able to detect both persistence and change—the persistence of places, objects, and substances along with whatever changes they undergo. Everything in the world persists in some respects and changes in some respects. So also does the observer himself. And some things persist for long intervals, others for short.

The perceiving of persistence and change (instead of color, form, space, time, and motion) can be stated in various ways. We can say that the perceiver *separates* the change from the nonchange, *notices* what stays the same and what



does not, or *sees* the continuing identity of things along with the events in which they participate. The question, of course, is how he does so. What is the information for persistence and change? The answer must be of this sort: The perceiver extracts the invariants of structure from the flux of stimulation while still noticing the flux. For the visual system in particular, he tunes in on the invariant structure of the ambient optic array that underlies the changing perspective structure caused by his movements.

The hypothesis that invariance under optical transformation constitutes information for the perception of a rigid persisting object goes back to the moving-shadow experiment (Gibson and Gibson, 1957). The outcome of that experiment was paradoxical; it seemed at the time that a changing form elicited the perception of a constant form with a changing slant. The solution was to postulate invariants of optical structure for the persisting object, “formless” invariants, and a particular disturbance of optical structure for the motion of the object, a perspective transformation. Separate terms needed to be devised for physical motions and for the optical motions that specified them, for events in the world and for events in the array, for geometry did not provide the terms. Similarly, different terms need to be invented to describe invariants of the changing world and invariants of the changing array; the geometrical word *form* will not do. Perhaps the best policy is to use the terms *persistence* and *change* to refer to the environment but *preservation* and *disturbance* of structure to refer to the optic array.

The stimulus-sequence theory of perception, based on a succession of discrete eye fixations, can assume only that the way to apprehend persistence is by an act of comparison and judgment. The perception of what-it-is-now is compared with the memory of what-it-was-then, and they are judged *same*. The continuous pickup theory of perception can assume that the apprehension of persistence is a simple act of invariance detection. Similarly, the snapshot theory must assume that the way to apprehend change is to compare what-it-is-now with what-it-was-then and judge *different*, whereas the pickup theory can assume an awareness of transformation. The congruence of the array with itself or the disparity of the array with itself, as the case may be, is picked up.

The perception of the persisting identity of things is fundamental to other kinds of perception. Consider an example, the persisting identity of another person. How does a child come to apprehend the identity of the mother? You might say that when the mother-figure, or the face, is continually fixated by the child the persistence of the sensation is supported by the continuing stimulus. So it is when the child clings to the mother. But what if the mother-figure is scanned? What if the figure leaves and returns to the field of view? What if the figure goes away and comes back? What is perceived when it emerges from the distance or from darkness, when its back is turned, when its clothing is changed, when its emotional state is altered, when it comes back into sight after a long interval? In short, how is it that the phenomenal identity of a person agrees so well with the biological identity, despite all the vicissitudes of the figure in the optic array and all the events in which the person participates?

The same questions can be asked about inanimate objects, attached objects, places, and substances. The features of a person are invariant to a considerable degree (the eyes, nose, mouth, style of gesture, and voice). But so are the analogous features of other things, the child's blanket, the kitchen stove, the bedroom, and the bread on the table. All have to be identified as continuing, as persisting, as maintaining existence. And this is not explained by the constructing of a concept for each.

We are accustomed to assuming that successive stimuli from the same entity, sensory encounters with it, are united by an act of recognition. We have assumed that perception ceases and memory takes over when sensation stops. Hence, every fresh glimpse of anything requires the act of linking it up with the memories of that thing instead of some other thing. The judgment, "I have seen this before," is required for the apprehension of "same thing," even when the observer has only turned away, or has only glanced away for an instant. The classical theory of sense perception is reduced to an absurdity by this requirement. The alternative is to accept the theory of invariance detection.

### **THE EFFECT OF PERSISTING STIMULATION ON PERCEPTION**

We have assumed that perception stops when sensation stops and that sensation stops when stimulation stops, or very soon thereafter. Hence, a persisting stimulus is required for the perception of a persisting object. The fact is, however, that a truly persisting stimulus on the retina or the skin specifies only that the observer does not or cannot move his eye or his limb, and the sense perception soon fades out by sensory adaptation (Chapter 4). The persistence of an object is specified by invariants of structure, not by the persistence of stimulation.

The seeing of persistence considered as the picking up of invariants under change resolves an old puzzle: the phenomenal identity of the spots of a retinal pattern when the image is transposed over the retina stroboscopically. The experiments of Josef Ternus first made this puzzle evident. See Gibson (1950, pp. 56 ff.) for a discussion and references.

I used to think that the aftereffects of persisting stimulation of the retina obtained by the prolonged fixation of a display could be very revealing. Besides ordinary afterimages there are all sorts of perceptual aftereffects, some of which I discovered. But I no longer believe that experiments on so-called perceptual adaptation are revealing, and I have given up theorizing about them. The aftereffects of prolonged scrutiny are of many sorts. Until we know more about information pickup, this field of investigation will be incoherent.

The quality of familiarity that can go with the perception of a place, object, or person, as distinguished from the quality of unfamiliarity, is a fact of experience. But is familiarity a result of the percept making contact with the traces of past percepts of the same thing? Is unfamiliarity a result of not making such contact? I think not. There is a circularity in the reasoning, and it is a bad theory. The quality of familiarity simply accompanies the perception of persistence.

The perception of the persisting identity of places and objects is more fundamental than the perception of the differences among them. We are told that to perceive something is to categorize it, to distinguish it from the other types of things that it might have been. The essence of perceiving is discriminating. Things differ among themselves, along dimensions of difference. But this leaves out of account the simple fact that the substance, place, object, person, or whatever has to last long enough to be distinguished from other substances, places, objects, or persons. The detecting of the invariant features of a persisting thing should not be confused with the detecting of the invariant features that make different things similar. Invariants over time and invariants over entities are not grasped in the same way.

In the case of the persisting thing, I suggest, the perceptual system simply extracts the invariants from the flowing array; it *resonates* to the invariant structure or is *attuned* to it. In the case of substantially distinct things, I venture, the perceptual system must *abstract* the invariants. The former process seems to be simpler than the latter, more nearly automatic. The latter process has been interpreted to imply an intellectual act of lifting out something that is mental from a collection of objects that are physical, of forming an abstract concept from concrete percepts, but that is very dubious. Abstraction is invariance detection across objects. But the invariant is only a similarity, not a persistence.

### ***Summary of the Theory of Pickup***

According to the theory being proposed, *perceiving* is a registering of certain definite dimensions of invariance in the stimulus flux together with definite parameters of disturbance. The invariants are invariants of structure, and the disturbances are disturbances of structure. The structure, for vision, is that of the ambient optic array.

The invariants specify the persistence of the environment and of oneself. The disturbances specify the changes in the environment and of oneself. A perceiver is aware of her existence in a persisting environment and is also aware of her movements relative to the environment, along with the motions of objects and nonrigid surfaces relative to the environment. The term *awareness* is used to imply a direct pickup of the information, not necessarily to imply consciousness.

There are many dimensions of invariance in an ambient optic array over time, that is, for paths of observation. One invariant, for example, is caused by

the occluding edge of the nose, and it specifies the self. Another is the gradient of optical texture caused by the material texture of the substratum, and it specifies the basic environment. Equally, there are many parameters of disturbance of an ambient optic array. One, for example, is caused by the sweeping of the nose over the ambient optic array, and it specifies head turning. Another is the deletion and accretion of texture at the edges of a form in the array, and it specifies the motion of an object over the ground.

For different kinds of events in the world there are different parameters of optical disturbance, not only accretion–deletion but also polar outflow–inflow, compression, transformation, substitution, and others. Hence, the same object can be seen undergoing different events, and different objects can be seen undergoing the same event. For example, an apple may ripen, fall, collide, roll, or be eaten, and eating may happen to an apple, carrot, egg, biscuit, or lamb chop. If the parameter of optical disturbance is distinguished, the event will be perceived. Note how radically different this is from saying that if stimulus–event *A* is invariably followed by stimulus–event *B* we will come to expect *B* whenever we experience *A*. The latter is classical association theory (or conditioning theory, or expectancy theory). It rests on the stimulus–sequence doctrine. It implies that falling, colliding, rolling, or eating are not units but sequences. It implies, with David Hume, that even if *B* has followed *A* a thousand times there is no certainty that it will follow *A* in the future. An event is only known by a conjunction of atomic sensations, a contingency. If this recurrent sequence is experienced again and again, the observer will begin to anticipate, or have faith, or learn by induction, but that is the best he can do.

The process of pickup is postulated to depend on the input–output loop of a perceptual system. For this reason, the information that is picked up cannot be the familiar kind that is transmitted from one person to another and that can be stored. According to pickup theory, information does not have to be stored in memory because it is always available.

The process of pickup is postulated to be very susceptible to development and learning. The opportunities for educating attention, for exploring and adjusting, for extracting and abstracting are unlimited. The increasing capacity of a perceptual system to pick up information, however, does not in itself constitute information. The ability to perceive does not imply, necessarily, the having of an idea of what can be perceived. The having of ideas is a fact, but it is not a prerequisite of perceiving. Perhaps it is a kind of extended perceiving.

## The Traditional Theories of Perception: Input Processing

The theory of information pickup purports to be an alternative to the traditional theories of perception. It differs from all of them, I venture to suggest, in rejecting the assumption that perception is the processing of inputs. *Inputs* mean sensory or afferent nerve impulses to the brain.

Adherents to the traditional theories of perception have recently been making the claim that what they assume is the processing of information in a modern sense of the term, not sensations, and that therefore they are not bound by the traditional theories of perception. But it seems to me that all they are doing is climbing on the latest bandwagon, the computer bandwagon, without reappraising the traditional assumption that perceiving is the processing of inputs. I refuse to let them pre-empt the term *information*. As I use the term, it is not something that has to be processed. The inputs of the receptors have to be processed, of course, because they in themselves do not specify anything more than the anatomical units that are triggered.

All kinds of metaphors have been suggested to describe the ways in which sensory inputs are processed to yield perceptions. It is supposed that sensation occurs first, perception occurs next, and knowledge occurs last, a progression from the lower to the higher mental processes. One process is the filtering of sensory inputs. Another is the organizing of sensory inputs, the grouping of elements into a spatial pattern. The integrating of elements into a temporal pattern may or may not be included in the organizing process. After that, the processes become highly speculative. Some theorists propose mental operations. Others argue for semilogical processes or problem-solving. Many theorists are in favor of a process analogous to the decoding of signals. All theorists seem to agree that past experience is brought to bear on the sensory inputs, which means that memories are somehow applied to them. Apart from filtering and organizing, the processes suggested are cognitive. Consider some of them.

### ***Mental Operations on the Sensory Inputs***

The a priori categories of understanding possessed by the perceived, according to Kant

The perceiver's presuppositions about what is being perceived

Innate ideas about the world

### ***Semilogical Operations on the Sensory Inputs***

Unconscious inferences about the outer causes of the sensory inputs, according to Helmholtz (the outer world is deduced)

Estimates of the probable character of the "distant" objects based on the "proximal" stimuli, according to Egon Brunswik (1956), said to be a quasirational, not a fully rational, process

### ***Decoding Operations on the Sensory Inputs***

The interpreting of the inputs considered as signals (a very popular analogy with many variants)

The decoding of sensory messages

The utilizing of sensory cues

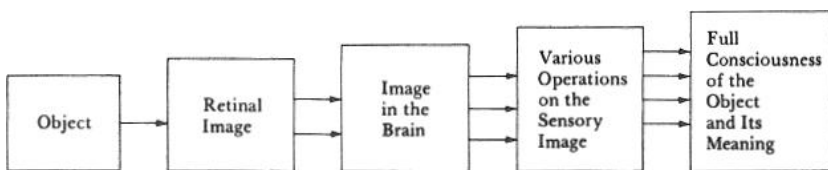
The understanding of signs, or indicators, or even *clues*, in the manner of a police detective

### ***The Application of Memories to the Sensory Inputs***

The “accrual” of a context of memory images and feelings to the core of sensations, according to E. B. Titchener’s theory of perception (1924).

This last hypothetical process is perhaps the most widely accepted of all, and the most elaborated. Perceptual learning is supposed to be a matter of enriching the input, not of differentiating the information (Gibson and Gibson, 1955). But the process of combining memories with inputs turns out to be not at all simple when analyzed. The appropriate memories have to be retrieved from storage, that is, aroused or summoned; an image does not simply accrue. The sensory input must fuse in some fashion with the stored images; or the sensory input is assimilated to a composite memory image, or, if this will not do, it is said to be assimilated to a class, a type, a schema, or a concept. Each new sensory input must be categorized—assigned to its class, matched to its type, fitted to its schema, and so on. Note that categories cannot become established until enough items have been classified but that items cannot be classified until categories have been established. It is this difficulty, for one, that compels some theorists to suppose that classification is a priori and that people and animals have innate or instinctive knowledge of the world.

The error lies, it seems to me, in assuming that either innate ideas or acquired ideas must be applied to bare sensory inputs for perceiving to occur. The fallacy is to assume that because inputs convey no knowledge they can somehow be made to yield knowledge by “processing” them. 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.



**FIGURE 14.1** The commonly supposed sequence of stages in the visual perceiving of an object.

All this should be treated as ancient history. Knowledge of the environment, surely, develops as perception develops, extends as the observers travel, gets finer as they learn to scrutinize, gets longer as they apprehend more events, gets fuller as they see more objects, and gets richer as they notice more affordances. Knowledge of this sort does not “come from” anywhere; it is got by looking, along with listening, feeling, smelling, and tasting. The child also, of course, begins to acquire knowledge that comes from parents, teachers, pictures, and books. But this is a different kind of knowledge.

### The False Dichotomy between Present and Past Experience

The division between present experience and past experience may seem to be self-evident. How could anyone deny it? Yet it is denied in supposing that we can experience both change and nonchange. The difference between present and past blurs, and the clarity of the distinction slips away. The stream of experience does not consist of an instantaneous present and a linear past receding into the distance; it is not a “traveling razor’s edge” dividing the past from the future. Perhaps the present has a certain duration. If so, it should be possible to find out when perceiving stops and remembering begins. But it has not been possible. There are attempts to talk about a “conscious” present, or a “specious” present, or a “span” of present perception, or a span of “immediate memory,” but they all founder on the simple fact that there is no dividing line between the present and the past, between perceiving and remembering. A special sense impression clearly ceases when the sensory excitation ends, but a perception does not. It does not become a memory after a certain length of time. A perception, in fact, does not *have* an end. Perceiving goes on.

Perhaps the force of the dichotomy between present and past experience comes from language, where we are not allowed to say anything intermediate between “I see you” and “I saw you” or “I am seeing you” and “I was seeing you.” Verbs can take the present tense or the past tense. We have no words to describe my continuing awareness of you, whether you are in sight or out of sight. Language is categorical. Because we are led to separate the present from the past, we find ourselves involved in what I have called the “muddle of memory” (Gibson, 1966a). We think that the past ceases to exist unless it is “preserved” in memory. We assume that memory is the bridge between the past and the present. We assume that memories accumulate and are stored somewhere; that they are images, or pictures, or representations of the past; or that memory is actually physiological, not mental, consisting of engrams or traces; or that it actually consists of neural connections, not engrams; that memory is the basis of all learning; that memory is the basis of habit; that memories live on in the unconscious; that heredity is a form of memory; that cultural heredity is another form of memory; that any effect of the past on the present is memory, including hysteresis. If we cannot do any better than this, we should stop using the word.

The traditional theories of perception take it for granted that what we see *now*, present experience, is the sensory basis of our perception of the environment and that what we have seen *up to now*, past experience, is added to it. We can only understand the present in terms of the past. But what we see *now* (when it is carefully analyzed) turns out to be at most a peculiar set of surfaces that happen to come within the field of view and face the point of observation (Chapter 11). It does not comprise what we see. It could not possibly be the basis of our perception of the environment. What we see *now* refers to the self, not the environment. The perspective appearance of the world at a given moment of time is simply what specifies to the observer where he is at that moment. The perceptual process does not begin with this peculiar projection, this momentary pattern. The perceiving of the world begins with the pickup of invariants.

Evidently the theory of information pickup does not need memory. It does not have to have as a basic postulate the effect of past experience on present experience by way of memory. It needs to explain learning, that is, the improvement of perceiving with practice and the education of attention, but not by an appeal to the catch-all of past experience or to the muddle of memory.

The state of a perceptual system is altered when it is attuned to information of a certain sort. The system has become sensitized. Differences are noticed that were previously not noticed. Features become distinctive that were formerly vague. But this altered state need not be thought of as depending on a memory, an image, an engram, or a trace. An image of the past, if experienced at all, would be only an incidental symptom of the altered state.

This is not to deny that reminiscence, expectation, imagination, fantasy, and dreaming actually occur. It is only to deny that they have an essential role to play in perceiving. They are kinds of visual awareness other than perceptual. Let us now consider them in their own right.

## A New Approach to Nonperceptual Awareness

The redefinition of *perception* implies a redefinition of the so-called higher mental processes. In the old mentalistic psychology, they stood above the lower mental processes, the sensory and reflex processes, which could be understood in terms of the physiology of receptors and nerves. These higher processes were vaguely supposed to be intellectual processes, inasmuch as the intellect was contrasted with the senses. They occurred in the brain. They were operations of the mind. No list of them was ever agreed upon, but *remembering*, *thinking*, *conceiving*, *inferring*, *judging*, *expecting*, and, above all, *knowing* were the words used. *Imagining*, *dreaming*, *rationalizing*, and *wishful thinking* were also recognized, but it was not clear that they were higher processes in the intellectual sense. I am convinced that none of them can ever be understood as an operation of the mind. They will never be understood as reactions of the body, either. But perhaps if they are



reconsidered in relation to ecological perceiving they will begin to sort themselves out in a new and reasonable way that fits with the evidence.

To perceive is to be aware of the surfaces of the environment and of oneself in it. The interchange between hidden and unhidden surfaces is essential to this awareness. These are existing surfaces; they are specified at some points of observation. Perceiving gets wider and finer and longer and richer and fuller as the observer explores the environment. The full awareness of surfaces includes their layout, their substances, their events, and their affordances. Note how this definition includes within perception a part of memory, expectation, knowledge, and meaning—some part but not all of those mental processes in each case.

One kind of remembering, then, would be an awareness of surfaces that have ceased to exist or events that will not recur, such as items in the story of one's own life. There is no point of observation at which such an item will come into sight.

To expect, anticipate, plan, or imagine creatively is to be aware of surfaces that do not exist or events that do not occur but that could arise or be fabricated within what we call the limits of possibility.

To daydream, dream, or imagine wishfully (or fearfully) is to be aware of surfaces or events that do not exist or occur and that are outside the limits of possibility.

These three kinds of nonperceptual awareness are not explained, I think, by the traditional hypothesis of mental imagery. They are better explained by some such hypothesis as this: a perceptual system that has become sensitized to certain invariants and can extract them from the stimulus flux can also operate without the constraints of the stimulus flux. Information becomes further detached from stimulation. The adjustment loops for looking around, looking at, scanning, and focusing are then inoperative. The visual system visualizes. But this is still an activity of the system, not an appearance in the theater of consciousness.

Besides these, other kinds of cognitive awareness occur that are not strictly perceptual. Before considering them, however, I must clarify what I mean by *imaginary* or *unreal*.

## The Relationship between Imagining and Perceiving

I assume that a normal observer is well aware of the difference between surfaces that exist and surfaces that do not. (Those that do not have ceased to exist, or have not begun to, or have not and will not.) How can this be so? What is the information for existence? What are the criteria? It is widely believed that young children are not aware of the differences, and neither are adults suffering from hallucinations. They do not distinguish between what is "real" and what is "imaginary" because perception and mental imagery cannot be separated. This doctrine rests on the assumption that, because a percept and an image both

occur in the brain, the one can pass over into the other by gradual steps. The only “tests for reality” are intellectual. A percept cannot validate itself.

We have been told ever since John Locke that an image is a “faint copy” of a percept. We are told by Titchener (1924) that an image is “easily confused with a sensation” (p. 198). His devoted student, C. W. Perky, managed to show that a faint optical picture secretly projected from behind on a translucent screen is sometimes not identified as such when an observer is imagining an object of the same sort on the screen (Perky, 1910). We are told by a famous neurosurgeon that electrical stimulation of the surface of the brain in a conscious patient “has the force” of an actual perception (Penfield, 1958). It is said that when a feeling of reality accompanies a content of consciousness it is marked as a percept and when it does not it is marked as an image. All these assertions are extremely dubious.

I suggest that perfectly reliable and automatic tests for reality are involved in the working of a perceptual system. They do not have to be intellectual. A surface is seen with more or less definition as the accommodation of the lens changes; an image is not. A surface becomes clearer when fixated; an image does not. A surface can be scanned; an image cannot. When the eyes converge on an object in the world, the sensation of crossed diplopia disappears, and when the eyes diverge, the “double image” reappears; this does not happen for an image in the space of the mind. An object can be scrutinized with the whole repertory of optimizing adjustments described in Chapter 11. No image can be scrutinized—not an afterimage, not a so-called eidetic image, not the image in a dream, and not even a hallucination. An imaginary object can undergo an *imaginary* scrutiny, no doubt, but you are not going to discover a new and surprising feature of the object this way. For it is the very features of the object that your perceptual system has already picked up that constitute your ability to visualize it. The most decisive test for reality is whether you can discover new features and details by the act of scrutiny. Can you obtain new stimulation and extract new information from it? Is the information inexhaustible? Is there more to be seen? The imaginary scrutiny of an imaginary entity cannot pass this test.

A related criterion for the existence of a thing is reversible occlusion. Whatever goes out of sight as you move your head and comes into sight as you move back is a *persisting* surface. Whatever comes into sight when you move your head is a *preexisting* surface. That is to say, it exists. The present, past, or future tense of the verb *see* is irrelevant; the fact is perceived without words. Hence, a criterion *for real* versus *imaginary* is what happens when you turn and move. When the infant turns her head and creeps about and brings her hands in and out of her field of view, she perceives what is real. The assumption that children cannot tell the difference between what is real and what is imaginary until the intellect develops is mentalistic nonsense. As the child grows up, she apprehends more reality as she visits more places of her habitat.

Nevertheless, it is argued that dreams sometimes have the “feeling” of reality, that some drugs can induce hallucinations, and that a true hallucination in

psychosis is proof that a mental image can be the same as a percept, for the patient acts as if he were perceiving and thinks he is perceiving. I remain dubious (Gibson, 1970). The dreamer is asleep and cannot make the ordinary tests for reality. The drug-taker is hoping for a vision and does not want to make tests for reality. There are many possible reasons why the hallucinating patient does not scrutinize what he says he sees, does not walk around it or take another look at it or test it.

There is a popular fallacy to the effect that if you can touch what you see it is real. The sense of touch is supposed to be more trustworthy than the sense of sight, and Bishop Berkeley's theory of vision was based on this idea. But it is surely wrong. Tactual hallucinations can occur as well as visual. And if the senses are actually perceptual systems, the haptic system as I described it (Gibson, 1966*b*) has its own exploratory adjustments and its own automatic tests for reality. One perceptual system does not *validate* another. Seeing and touching are two ways of getting much the same information about the world.

## A New Approach to Knowing

The theory of information pickup makes a clear-cut separation between perception and fantasy, but it closes the supposed gap between perception and knowledge. The extracting and abstracting of invariants are what happens in both perceiving and knowing. To perceive the environment and to conceive it are different in degree but not in kind. One is continuous with the other. Our reasons for supposing that seeing something is quite unlike knowing something come from the old doctrine that seeing is having temporary sensations one after another at the passing moment of present time, whereas knowing is having permanent concepts stored in memory. It should now be clear that perceptual seeing is an awareness of persisting structure.

Knowing is an *extension* of perceiving. The child becomes aware of the world by looking around and looking at, by listening, feeling, smelling, and tasting, but then she begins to be *made* aware of the world as well. She is shown things, and told things, and given models and pictures of things, and then instruments and tools and books, and finally rules and short cuts for finding out more things. Toys, pictures, and words are aids to perceiving, provided by parents and teachers. They transmit to the next generation the tricks of the human trade. The labors of the first perceivers are spared their descendants. The extracting and abstracting of the invariants that specify the environment are made vastly easier with these aids to comprehension. But they are not in themselves knowledge, as we are tempted to think. All they can do is facilitate knowing by the young.

These extended or aided modes of apprehension are all cases of information pickup from a stimulus flux. The learner has to hear the speech in order to pick up the message; to see the model, the picture, or the writing; to manipulate the instrument in order to extract the information. But the information itself is largely independent of the stimulus flux.

What are the kinds of culturally transmitted knowledge? I am uncertain, for they have not been considered at this level of description. Present-day discussions of the “media of communication” seem to me glib and superficial. I suspect that there are many kinds merging into one another, of great complexity. But I can think of three obvious ways to facilitate knowing, to aid perceiving, or to extend the limits of comprehension: the use of instruments, the use of verbal descriptions, and the use of pictures. Words and pictures work in a different way than do instruments, for the information is obtained at second hand. Consider them separately.

### ***Knowing Mediated by Instruments***

Surfaces and events that are too small or too far away cannot be perceived. You can of course increase the visual solid angle if you approach the item and put your eye close to it, but that procedure has its limits. You cannot approach the moon by walking, and you cannot get your eye close enough to a drop of pond water to see the little animals swimming in it. What can be done is to *enlarge* the visual solid angle from the moon or the water drop. You can convert a tiny sample of the ambient optic array at a point of observation into a magnified sample by means of a telescope or a microscope. The structure of the sample is only a little distorted. The surfaces perceived when the eye is placed at the eyepiece are “virtual” instead of “real,” but only in the special sense that they are very much closer to the observer. The invariants of structure are nearly the same when a visual angle with its nested components is magnified. This description of magnification comes from ecological optics. For designing the lens system of the instrument, a different optics is needed.

The discovery of these instruments in the seventeenth century enabled men to know much more about very large bodies and very small bodies than they had before. But this new knowledge was almost like seeing. The mountains of the moon and the motions of a living cell could be observed with adjustments of the instrument not unlike those of the head and eyes. The guarantees of reality were similar. You did not have to take another person’s word for what he had seen. You might have to learn to use the instrument, but you did not have to learn to interpret the information. Nor did you have to judge whether or not the other person was telling the truth. With a telescope or a microscope you could look for yourself.

#### **THE UNAIDED PERCEIVING OF OBJECTS IN THE SKY**

Objects in the sky are very different from objects on the ground. The heavenly bodies do not come to rest on the ground as ordinary objects do. The

rainbow and the clouds are transient, forming and dissipating like mists on earth. But the sun, the moon, the planets, and the stars seem permanent, appearing to revolve around the stationary earth in perfect cycles and continuing to exist while out of sight. They are immortal and mysterious. They cannot be scrutinized.

Optical information for direct perception of these bodies with the unaided eye is lacking. Their size and distance are indeterminate except that they rise and set from behind the distant horizon and are thus very far away. Their motions are very different from those of ordinary objects. The character of their surfaces is indefinite, and of what substances they are composed is not clear. The sun is fiery by day, and the others are fiery at night, unlike the textured reflecting surfaces of most terrestrial objects. What they afford is not visible to the eye. Lights in the sky used to look like gods. Nowadays they look like flying saucers.

All sorts of instruments have been devised for mediating apprehension. Some optical instruments merely enhance the information that vision is ready to pick up; others—a spectroscope, for example—require some inference; still others, like the Wilson cloud chamber, demand a complex chain of inferences.

Some measuring instruments are closer to perception than others. The measuring stick for counting units of distance, the gravity balance for counting units of mass, and the hourglass for time are easy to understand. But the complex magnitudes of physical science are another matter. The voltmeters, accelerometers, and photometers are hard to understand. The child can see the pointer and the scale well enough but has to learn to “read” the instrument, as we say. The direct perception of a distance is in terms of whether one can jump it. The direct perception of a mass is in terms of whether one can lift it. Indirect knowledge of the metric dimensions of the world is a far extreme from direct perception of the affordance dimensions of the environment. Nevertheless, they are both cut from the same cloth.

### ***Knowing Mediated by Descriptions: Explicit Knowledge***

The principal way in which we save our children the trouble of finding out everything for themselves is by describing things for them. We transmit information and convey knowledge. Wisdom is handed down. Parents and teachers and books give the children knowledge of the world at second hand. Instead of having to be extracted by the child from the stimulus flux, this knowledge is communicated to the child.

It is surely true that speech and language convey information of a certain sort from person to person and from parent to child. Written language can even be stored so that it accumulates in libraries. But we should never forget that this

is information that has been put into words. It is not the limitless information available in a flowing stimulus array.

Knowledge that has been put into words can be said to be *explicit* instead of *tacit*. The human observer can verbalize his awareness, and the result is to make it communicable. But my hypothesis is that there has to be an awareness of the world before it can be put into words. You have to see it before you can say it. Perceiving precedes predicating.

In the course of development the young child first hears talk about what she is perceiving. Then she begins herself to talk about what she perceives. Then she begins to talk to herself about what she knows—when she is alone in her crib, for example. And, finally, her verbal system probably begins to verbalize silently, in much the same way that the visual system begins to visualize, without the constraints of stimulation or muscular action but within the limits of the invariants to which the system is attuned. But no matter how much the child puts knowledge into words all of it cannot be put into words. However skilled an explicator one may become one will always, I believe, see more than one can say.

Consider an adult, a philosopher, for example, who sees the cat on the mat. He knows *that* the cat is on the mat and believes the proposition and can say it, but all the time he plainly sees all sorts of wordless facts—the mat extending without interruption behind the cat, the far side of the cat, the cat hiding part of the mat, the edges of the cat, the cat being supported by the mat, or resting on it, the horizontal rigidity of the floor under the mat, and so on. The so-called concepts of extension, of far and near, gravity, rigidity, horizontal, and so on, are nothing but partial abstractions from a rich but unitary perception of *cat-on-mat*. The parts of it he can name are called concepts, but they are not all of what he can see.

### ***Fact and Fiction in Words and Pictures***

Information about the environment that has been put into words has this disadvantage: The reality testing that accompanies the pickup of natural information is missing. Descriptions, spoken or written, do not permit the flowing stimulus array to be scrutinized. The invariants have already been extracted. You have to trust the original perceiver; you must “take his word for it,” as we say. What he presents may be fact, or it may be fiction. The same is true of a depiction as of a description.

The child, as I argued above, has no difficulty in contrasting real and imaginary, and the two do not merge. But the factual and the fictional may do so. In storytelling, adults do not always distinguish between true stories and fairy stories. The child herself does not always separate the giving of an account from the telling of a story. Tigers and dragons are both fascinating beasts, and the child will not learn the difference until she perceives that the zoo contains the former but not the latter.

Fictions are not necessarily fantasies. They do not automatically lead one astray, as hallucinations do. They can promote creative plans. They can permit vicarious learning when the child identifies with a fictional character who solves problems and makes errors. The “comic” characters of childhood, the funny and the foolish, the strong and the weak, the clever and the stupid, occupy a great part of children’s cognitive awareness, but this does not interfere in the least with their realism when it comes to perceiving.

The difference between the real and the imaginary is specified by two different modes of operation of a perceptual system. But the difference between the factual and the fictional depends on the social system of communication and brings in complicated questions. Verbal descriptions can be true or false as predications. Visual depictions can be correct or incorrect in a wholly different way. A picture cannot be true in the sense that a proposition is true, but it may or may not be true to life.

### ***Knowing and Imagining Mediated by Pictures***

Perceiving, knowing, recalling, expecting, and imagining can all be induced by pictures, perhaps even more readily than by words. Picture-making and picture-perceiving have been going on for twenty or thirty thousand years of human life, and this achievement, like language, is ours alone. The image makers can arouse in us an awareness of what they have seen, of what they have noticed, of what they recall, expect, or imagine, and they do so *without converting the information into a different mode*. The description puts the optical invariants into words. The depiction, however, captures and displays them in an optic array, where they are more or less the same as they would be in the case of direct perception. So I will argue, at least. The justification of this theory is obviously not a simple matter, and it is deferred to the last chapters of this book, Part IV.

The reality-testing that accompanies unmediated perceiving and that is partly retained in perceiving with instruments is obviously lost in the kind of perceiving that is mediated by pictures. Nevertheless, pictures give us a kind of grasp on the rich complexities of the natural environment that words could never do. Pictures do not stereotype our experience in the same way and to the same degree. We can learn from pictures with less effort than it takes to learn from words. It is not like perceiving at first hand, but it is *more* like perceiving than any verbal description can be.

The child who has learned to talk about things and events can, metaphorically, talk to himself silently about things and events, so it is supposed. He is said to have “internalized” his speech, whatever that might mean. By analogy with this theory, a child who has learned to draw might be supposed to picture to himself things and events without movement of his hands, to have “internalized” his picturemaking. A theory of internal language and internal images might be based on this theory. But it seems to me very dubious. Whether or not

it is plausible is best decided after we have considered picturemaking in its own right.

## Summary

When vision is thought of as a perceptual system instead of as a channel for inputs to the brain, a new theory of perception considered as information pickup becomes possible. Information is conceived as available in the ambient energy flux, not as signals in a bundle of nerve fibers. It is information about both the persisting and the changing features of the environment together. Moreover, information about the observer and his movements is available, so that self-awareness accompanies perceptual awareness.

The qualities of visual experience that are specific to the receptors stimulated are not relevant to information pickup but incidental to it. Excitation and transmission are facts of physiology at the cellular level.

The process of pickup involves not only overt movements that can be measured, such as orienting, exploring, and adjusting, but also more general activities, such as optimizing, resonating, and extracting invariants, that cannot so easily be measured.

The ecological theory of direct perception cannot stand by itself. It implies a new theory of cognition in general. In turn, that implies a new theory of noncognitive kinds of awareness—fictions, fantasies, dreams, and hallucinations.

Perceiving is the simplest and best kind of knowing. But there are other kinds, of which three were suggested. Knowing by means of instruments extends perceiving into the realm of the very distant and the very small; it also allows of metric knowledge. Knowing by means of language makes knowing explicit instead of tacit. Language permits descriptions and pools the accumulated observations of our ancestors. Knowing by means of pictures also extends perceiving and consolidates the gains of perceiving.

The awareness of imaginary entities and events might be ascribed to the operation of the perceptual system with a suspension of reality-testing. Imagination, as well as knowledge and perception, can be aroused by another person who uses language or makes pictures.

These tentative proposals are offered as a substitute for the outworn theory of past experience, memory, and mental images.