TWELVE

LOOKING WITH THE HEAD AND EYES

We modern, civilized, indoors adults are so accustomed to looking at a page or a picture, or through a window, that we often lose the feeling of being *surrounded* by the environment, our sense of the *ambient* array of light. Even when outdoors under the sky, one is apt to be driving an automobile and looking only through the windshield, or traveling in a vehicle where the window to the outside world is constricted to a small angle. We do not look *around*.

We live boxed-up lives. Our ancestors were always looking around. They surveyed the environment, for they needed to know where they were and what there was in all directions. Children pay attention to their surroundings when allowed to do so. Animals must do so. But we adults spend most of our time *looking at* instead of *looking around*. In order to look around, of course, one must turn one's head.

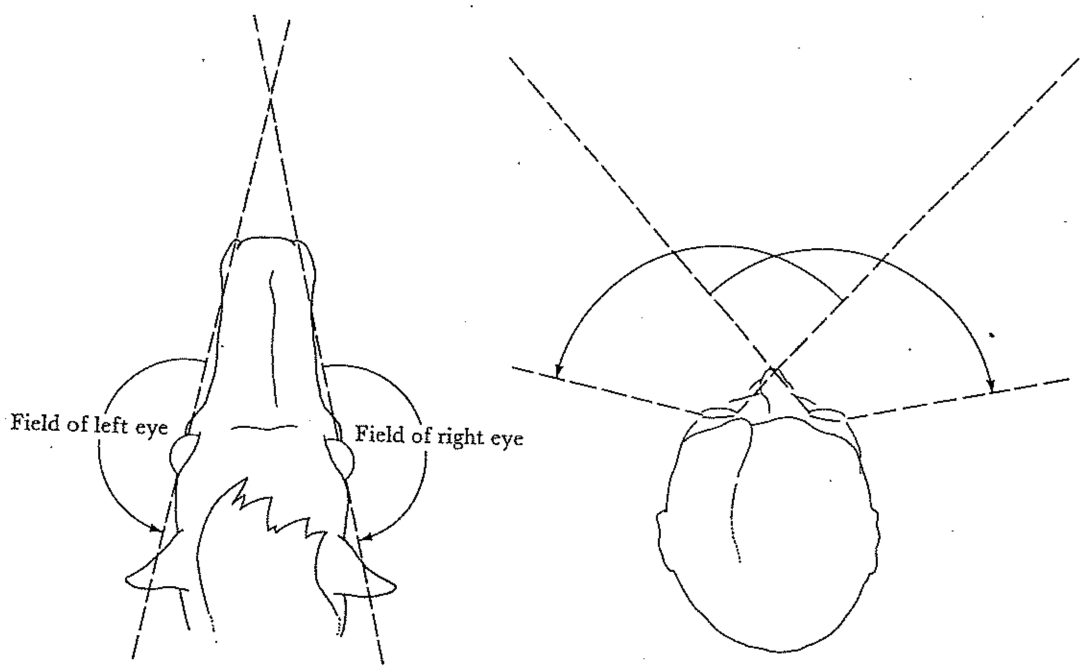
LOOKING AROUND AND LOOKING AT

The reason why humans must turn their heads in order to look around is that their eyes are set in the front of their heads instead of on either side, as they are in horses or rabbits. The orbits in the human skull are frontal, not lateral. The horse can see most of its surroundings (but not all) without having to turn its head; it can *see* around fairly well without having to *look* around. Thus, an enemy can sneak up on a person from behind, sometimes, but the hunter cannot sneak up on a rabbit. It has been suggested that animals who are preyed upon need a more panoramic field of view, whereas predatory animals such as cats can afford to have eyes in the front of the head (Walls, 1942). It has also been argued that the frontal eyes of primates living in the trees afford better "depth perception," but this argument presupposes the entrenched fallacy of depth perception that this book has been at such pains to destroy. Even if depth were perceived, it would be another error to assume that the only kind of depth perception is "binocular," that is, the kind that rests on binocular disparity.

With lateral eyes, the blind region behind the animal is minimal, but the overlap of the fields of view ahead of the animal is sacrificed. With frontal eyes, the overlap of the fields of view ahead is maximal, but the scope of the field of view is sacrificed and the blind region is large. Complete simultaneous ambience of perception is impossible. There has to be some gap in the combined field of view simply because the body of the animal itself is there, that is, its body is bound to hide some of the surfaces of the surrounding environment. Simultaneous ambience of perception is unnecessary in any case if the animal can always turn its head. There is no need to perceive everything at once if everything can be perceived in succession.

The gap in the combined field of view of the eyes is that portion of the ambient optic array filled by the head and body of the observer himself. It is a visual solid angle with a closed envelope, *a closed boundary in the array that specifies the body.* It has a meaning, and it carries information. I have already made this point at some length in Chapter 7 on the self. The portion of the environmental layout that is hidden by the body is continually interchanging with the unhidden remainder as the head turns and the body moves.

Figure 12.1
The lateral eyes of a horse and the frontal eyes of a man and the respective fields of view approximately.



The difference between the way a horse perceives its environment and the way a human does is therefore not so profound as you might assume. The blind region caused by the head and body of the horse is a small part of the ambient array, whereas the blind region caused by the head of a human is a large part, in fact, a visual solid angle of about 180°, approximating a hemisphere of the array. But it is not actually a blind region, of course; it is the head. A small turn of the head enables the horse to see what is behind, and a large turn of the head enables a person to see what is behind, but in both cases the observer sees himself in the middle of an environment. Are you doubtful that a horse can see itself? Why shouldn't the horse see itself just as much as the human does, if vision yields proprioception as well as exteroception? The horse's blind area is differently shaped than the human's—the boundaries of its field of view are different—but the blind area means the same thing to the horse as it does to the human. Egoreception and exteroception are inseparable kinds of experience. The seeing of oneself is not a complex intellectual experience but a simple primitive one. The orthodox dogma that no animal but the human animal has *self-consciousness* is surely false.

WITH WHAT DOES ONE SEE THE WORLD?

We human observers take it for granted that one sees the environment with one's eyes. The eyes are the organs of vision just as the ears are the organs of hearing, the nose is the organ of smelling, the mouth is the organ of tasting, and the skin is the organ of touching. The eye is considered to be an instrument of the mind, or an organ of the brain. But the truth is that each eye is positioned in a head that is in turn positioned on a trunk that is positioned on legs that maintain the posture of the trunk, head, and eyes relative to the surface of support. Vision is a whole perceptual system, not a channel of sense (Gibson, 1966*b*). One sees the environment not with the eyes but with the eyes-in-the-head-on-the-body-resting-on-the-ground. Vision does not have a *seat* in the body in the way that the mind has been thought to be seated in the brain. The perceptual capacities of the organism do not lie in discrete anatomical parts of the body but lie in systems with nested functions.

Even so, it might be argued, one surely *looks* with the eyes even if one does not *see* with the eyes. But looking with the eyes alone is mere looking *at*, not looking *around*. It is the scanning of an object, a page of print, or a picture. One also looks with the head, not just with the eyes, more exactly with the head-eye system, as I said at the outset.

The exclusive concern with eye turning to the neglect of head turning is one of the deep errors of the snapshot theory of vision and goes back at least a century.

Helmholtz asserted in *Physiological Optics* that "the intent of vision is to see as distinctly as possible various objects or parts of an object in succession. This is accomplished by so pointing the eyes that an image of the given object is projected on the fovea of each retina. The governing of the ocular movements is wholly subordinated to this end; both eyes are adjusted and accommodated together so as to permit this light absorptive pointing. Any . . . eye movement not having for its end the attaining of distinct imaging of an object cannot be performed" (Helmholtz, trans. 1925, p. 56). He assumed that objects and parts of objects are what we perceive and that these are limited to objects in the fixed field of view. He would be astonished at the assertion that a man perceives his surroundings, including the environment behind his head, for that is not "the intent of vision."

THE AWARENESS OF THE ENVIRONMENT AND THE EGO

Despite what Helmholtz said, some psychologists have insisted that a man *is* aware of the environment behind his head. Koffka was one who did so. Phenomenal space, he said, extends to the sides; yonder is the wall of the room and there are walls to the right and left, but phenomenal space also extends *behind*. You would be vividly aware of the space behind if the edge of a cliff were there. "Behavioral space does not confront me but encloses me." What is it that lies between the "in front" and the "behind"? It is, he says, "just that phenomenal object we call the Ego." It is a segregated object, like others in phenomenal space (Koffka, 1935, p. 322). It is only a step from this description to the theory that the head and body of the observer *hide* the surfaces of the world that are outside the occluding edges of the field of view. Koffka made no reference to head turning and failed to recognize the interchange of the hidden and the unhidden, and that is an important step, but he did recognize a fact of perception.

When I distinguished, years ago, between the visual field as one kind of experience and the visual world as a radically different kind (Gibson, 1950*b*, Ch. 3), I was elaborating on Koffka. The visual field, I suggested, consists of a patchwork of colors something like a picture, whereas the visual world consists of familiar surfaces and objects one behind another. The *visual field* has boundaries, roughly oval in shape, and it extends about 180° from side to side and about 140° up and down. The boundaries are not sharp, but they are easily observed when attended to. The *visual world*, however, has no such boundaries; it is unbounded, like the surface of a sphere extending all the way around me. The visual field is clear in the center and vague in the periphery—that is, less definite toward the boundaries—but the visual world has no such center of definition and is everywhere clear. The oval boundaries of the visual field sweep across the array whenever I turn my head and wheel over the array whenever I tilt my head, but the visual world is perfectly stationary and always upright. The patchwork of the visual field deforms as I move and, in particular, flows outward from a center when I move in the direction of that center, but the phenomenal surfaces of the world are always perfectly rigid.

The visual field is a special kind of experience that can arise from a sample of the ambient array taken with the head and eyes fixed. In its purest form, the visual field arises with a single fixed eye. The visual world is the kind of experience that arises naturally from the whole ambient array when one is looking around and looking with two eyes at two slightly different points of observation. The field of view of the two eyes is a sort of mixed cross-section of the overlapping solid angles registered by the eyes. The field of one eye would correspond to a plane picture cutting the solid angle for that eye. It would correspond in the sense that a faithful picture could be substituted for the angular sample so as to yield almost the same phenomenal experience. But the visual world is a kind of experience that *does not correspond to anything,* not any possible picture, not any motion picture, and not even any "panoramic" motion picture. The visual world is not a *projection* of the ecological world. How could it be? The visual world is the outcome of the picking up of invariant information in an ambient optic array by an exploring visual system, and the awareness of the observer's own body in the world is a part of the experience.

The awareness of "out there" and of "here" are complementary. The occluding boundary of the field of view constitutes "here." The content and details of the field of view are "out there," and the smaller the detail the farther away it is.

THE VISUAL EGO

Ecological optics distinguishes between an unoccupied point of observation in the medium and an occupied point (Chapter 7). The former is a position where an observer *might* be situated and the latter is a position where an observer *is* situated. The ambient optic array is then altered, for it includes a solid angle filled by the observer, having a boundary that is unique to the observer's particular anatomy. It is called the *blind*

Terminological Note

I should never have entitled my 1950 book *The Perception of the Visual World,* for it has promoted confusion. A better title would have been *The Visual Perception of the World.* The term *visual world* should be reserved for the awareness of the environment obtained by vision.

region in physiological optics. But it is blind only for exteroception, not for proprioception. It *looks like oneself.* Its shape depends on the shape of one's nose, the shape of one's head, and the shape of one's limbs. It is altered when a person puts on eyeglasses or when a horse is made to wear blinders. Thus, whenever a point of observation is occupied, the occupier is uniquely specified, whether adult or child, monkey or dog.

An observer perceives the position of *here* relative to the environment and also his body as *being* here. His limbs protrude into the field of view, and even his nose is a sort of protuberance into the field. Undoubtedly, the length of a man's nose determines how he sees himself. (Consider the ego of a baboon in this respect, and think what the ego of an elephant would be!) For us, the nose is the leftward edge of the right eye's field of view and the rightward edge of the left eye's field of view. Hence, it yields a kind of subjective sensation called a *double image* in the theory of binocular vision, in fact, the maximum limit of *crossed diplopia.* It would therefore be a theoretical zero for the dimension of distance from here.

Since the occupied point of observation is normally a moving position, not a stationary one, the animal sees its body moving relative to the ground. It sees that part of the enviironment toward which it is moving; it sees the movements of its feet, relative to its body and also over the ground. When it looks around during locomotion, it sees the turning of its head. These are all cases of visual kinesthesis.

THE PERSISTING ENVIRONMENT: PERSISTENCE, COEXISTENCE, AND CONCURRENCE

To say that one is aware of the environment behind one's head is to say that one is aware of the *persistence* of the environment. Things go out of sight and come into sight as the head turns in looking around, but they persist while out of sight. Whatever

THE INFORMATION FOR PERSISTENCE

The perceiving of the persistence of the environment is not, of course, an achievement of the visual system alone. It is a nonmodal form of perception, cutting across the perceptual systems and transcending the "senses." Touching and listening accompany looking. The young child who goes for a walk and looks around at the strange wide world can cling to the mother's hand, confirming her persistence while she is temporarily out of sight. Similarly, the persistence of the mother when she goes around the corner, or goes out of sight in the dark, is confirmed by hearing her voice. The information to specify the continued existence of something may be carried by touch or sound as well as by light. Incessant stimulation is not necessary for the perceiving of persistence.

leaves the field as one turns to the right re-enters the field as one turns to the left. The structure that is deleted is later accreted; this is a reversible transition, and therefore the structure can be said to be *invariant* under the transition. To pick up the invariant is to perceive the persistence of a surface, so my argument runs. If this is true, there is no need to appeal to a concept of "object permanence" or to any theory of how the concept might develop.

To perceive the persistence of surfaces that are out of sight is also to perceive their coexistence with those that are in sight. In short, the hidden is continuous with the unhidden; they are *connected*.

Separated places and objects are perceived to coexist. This means that separated *events* at these places are perceived to be concurrent. What happens at one end of a corridor is seen to co-occur with what happens at the other end, even though one must look back and forth between the two. Different concurrent events, thus, can be sampled in succession without destroying their concurrence, just as different coexisting places can be sampled in succession without destroying their coexistence.

HOW DOES THE EYE-HEAD SYSTEM WORK? OUTLINE OF A NEW THEORY

Looking-around and looking-at are acts that naturally go together, but they can be studied separately. In fact, looking-at has been studied almost exclusively by visual physiologists. What they have recorded and measured are so-called eye movements, that is, movements of the eyes *relative to the head.* The head is usually fixed in an apparatus. The eyes are then allowed to scan a display of some sort within the field of view of the stationary head, a pattern of luminous points in the dark, or a line of print on a page, or a picture. The eyes rotate in rapid jumps from one fixation to another, and these are called *saccadic movements.* In terms of the retinal image theory, the fovea of each retina is moved so that an image of the particular "object of interest" falls on the retinal point of highest acuity where the photoreceptors, the cones, are most densely packed together. The anatomical fovea corresponds to the psychological "center of clearest vision." The fine details of the optical image are said to be best "resolved" at the fovea. All this is implied in the quotation from Helmholtz, above.

THE RECOGNIZED TYPES OF EYE MOVEMENT

There are other kinds of ocular movements besides scanning, however, and the accepted classification goes back to R. Dodge (1903), who was the first investigator to record and measure them by photography. They have since been studied with ever-increasing ingenuity and precision, but Dodge's list has never been challenged by physiologists. He never doubted the eye-camera analogy; he only forced us to consider that the eyes were *movable* cameras at the ends of flexible cables leading to the brain. The list is approximately as given below.

1. *Fixation*: Not strictly a "movement," fixation is nevertheless an important kind of ocular behavior. It should be called a *posture* of the eye, a *pointing at.*

2. *Saccadic movement*: A saccadic movement is a rapid rotation of the eyeball from one fixation to another. It has long been taken for granted that the movement is a response of the eye muscles to a stimulus at the periphery of the retina such as to bring that stimulus to the center of the retina, the fovea. But I shall challenge this assumption.

3. *Pursuit movement*: This kind of movement is said to be fixation of the eye on a moving object in the world, often nowadays called *tracking*. It is much slower than a saccadic movement.

4. *Convergence and divergence*: Convergence is the inward rotation of each eye so as to permit both eyes to fixate on the same near object. Divergence is the opposite, a return of the ocular axes to parallel so as to permit both eyes to fixate on the same distant object. In retinal image optics, it is assumed that these movements occur so that the two similar but more or less disparate images of the object can be "fused" in the brain to yield a single phenomenal object with depth. They are said to be governed by what is called a *fusion reflex*, but this is not consistent with the notion of a reflex as a response to a stimulus. Note that in both saccadic and pursuit movements the two eyes fixate together and rotate together as if they were linked. They are said to be *conjugated*. But they rotate in opposite directions during the vergence movements. All three types of movement, however, can be said to work in the interest of fixation.

5. *Compensatory movement*: This movement is quite different from the others. Like them, it is a rotation of each eye in the head but in precisely the opposite direction from that of the head, and to exactly the same degree. It *compensates* for the turning of the head. Thus, it is a *nonmotion* of the eyes relative to the *environment*, a posture, like fixation. Anyone can note how exact this compensation normally is by looking at one eye in his mirror image and then moving his head around, left and right, up and down; the eye never swerves from its fixed orientation in space. It is as if anchored to the environment. When the head starts, the eye starts; when the head stops, the eye stops.

If the head turns through an angle too great for compensation, the eye jumps rapidly to a new orientation and holds it. Thus, a man on a mountaintop who turns around completely, taking several seconds for the act, keeps his eyes anchored to the dual ambient array for the whole period except for a small part of the time, totaling only a fraction of a second, during which the jumps have occurred. This is what happens in "looking around," and the result is a vivid perception of the whole environment. This is the natural exploratory activity of the visual system.

What experimenters do, however, is to put the subject in a rotating chair and turn him passively. In this unnatural situation a reflex response of the eyes is aroused to the stimulus of acceleration in the semicircular canals of the inner ear. It is called *nystagmus*. The compensatory turning of the eyes then has a certain latency; it does not begin with the turning of the head as does the compensatory turning of the eyes with an *active* head movement. The latter is not a reflex to a stimulus but a coordination. The head turning and the eye turning are concurrent movements of a single act. The active turning of the head *involves* the opposite turning of the eyes in much the same way that the contracting of the extensor muscles of a limb involves the relaxing of the flexor muscles. The neck muscles and the eye muscles are innervated at the same time, reciprocally. But the passive movement of the eyes in response to a passive movement of the head has received by far the most attention from experimenters. Physiologists are preoccupied with reflexes to stimuli, probably because they assume that reflexes are basic for behavior.

Experiments on ocular nystagmus with passive rotation often bring about a kind of disorientation of the eyes to the environment called *vertigo*. After stopping such rotation, the eyes will compensate for a nonexistent turning of the head. The experimenter has overstrained the capacity of the system. The observer reports that the world seems to be going around and usually that he feels as if his body were also being rotated. These two experiences are inconsistent. He usually just says that he is dizzy. He is at any rate disoriented to the environment: he cannot point to things, he will stagger, and somtimes fall down. I have described the limitations of the vestibular apparatus in my chapter on the basic orienting system (Gibson, 1966*b*, Ch. 4.). The study of dizziness, however interesting and important for neurology, tells us nothing about the normal working of the eye-head system. I would explain it by saying that the normal complementarity of exteroception and proprioception has broken down.

A RECONSIDERATION OF EYE MOVEMENTS

Ecological optics as distinguished from eyeball optics calls for a re-examination of the traditional eye movements. We must consider how the visual system works, not just how the eyes move. Eyeball optics is appropriate for visual physiology and the prescribing of eyeglasses but not for the psychology of visual perception.

Fixation The prolonged fixing of the eyes on an "object or part of an object," the bringing of its image to the fovea and keeping it there, does not occur in life. It is a laboratory artifice, brought about when an experimenter tells an observer to stare at a "fixation point" that is usually of no interest to her. No one stares at a fixed point in the world for long unless she is so preoccupied that she is actually not seeing what she looks at. Seeming exceptions arise in the aiming of a rifle or the threading of a needle, but these are actually cases where different objects are *aligned*, not where a single object is fixated. The eyes normally search, explore, or scan, and there are seldom fewer than several saccadic jumps per second. They *look at* but do not *fixate*.

Even when fixation is artificially prolonged in the laboratory, it turns out not to be pure fixation, a steady posture. The eye is never literally fixed. It undergoes a series of miniature movements or microsaccades. The recording of such eye movements has become very precise in recent years, and the evidence now suggests that looking at a tiny thing consists of making tiny movements. If so, looking is always exploring, even so-called fixating. On the smallest scale, the eyes could never be perfectly steady, for the eye muscles that control their posture undergo tremor and the eyes tremble in the same way that the hand does when you hold it out. There does not seem to be any clear separation between large saccades, small saccades, microsaccades, and tremor. Perhaps the general conclusion should be that *an eye-posture is nothing but movements that are very small.*

This conclusion is consistent with my conception of the ambient optic array. It consists of adjacent visual solid angles that are *nested*, each solid angle having its base in a feature or face or facet of the environmental layout, the features being *themselves* nested in superordinate and subordinate units. The eyes can explore the large details of large solid angles. And the eye-head system can explore the hemispheric solid angles of the sky and the earth or of the mountains to the east and the valley to the west. We perceive a large mural painting with sweeps of the eyes. We perceive a page of print with small saccades. And one puts a thread into the eye of a needle with the tiniest saccades of all.

Saccadic Movement The jump of the eyeball from one fixation to another, it seems, can vary from an angle of many degrees to one of a few minutes of arc or less. So, just as there is no pure fixation, there is also no pure movement. There are postures of the eyes that are relatively stable and movements of the eyes from one such posture to another, but they grade into each other. Moving and fixating are complementary. They combine in the act of scanning.

It is certainly a fallacy to assume that a saccadic movement is a response to a "stimulus" on the periphery of the retina that brings it to the fovea. There are no stimuli in an optic array. That assumption comes from experiments in which a point of light is flashed on in utter darkness; the eyes then turn so as to foveate it, but this experimental situation does not apply to everyday vision. Visual physiologists, however, presuppose an array of stimuli and assume that a localizing movement, a "fixation reflex," tends to occur for each retinal point when it is stimulated.

It is also a fallacy, if a little more plausible, to assume that a series of fixations is a series of acts of selective attention to the different *objects* in the world. Each fixation would then be a centering of foveal attention on one object to the exclusion of other objects. Each saccade must then be a *movement* of attention from one object to another. But the truth is that attention is not only selective, it is also integrative. Attention can be distributed as well as being concentrated. The awareness of details is not inconsistent with the awareness of wholes. Each in fact implies the other. One can perfectly well pay attention to some aspect of the environment that extends over a large angle of the ambient array, such as the gradient of the ground that goes all the way from one's feet out to the horizon. Hence, a whole series of fixations can be a single act of attention.

Pursuit Movement Not just a fixation of the eyes on a moving object in the world, pursuit movement is also, and usually, an adjustment of the ocular system to the flowing ambient array during locomotion of the observer. The centrifugal outflow of the optic array from the direction in which one is traveling must be attended to in order to see where one is going, and in order to control one's locomotion. The eyes are pointed at one element of the flowing array so that all the other elements of the flow pattern that fall into lawful relationships to it can be picked up. This is what happens when you drive your car down the road: your eyes fix on a piece of the layout and track it downward, then jump ahead to a new piece. These drifts and jumps are somewhat similar to the compensatory nystagmus with head turning, but the drifts during locomotion are not the same as the drifts during compensation.

Convergence and Divergence Retinal image optics assumes that if one object in space has an image in each of the two eyes the two images have to be fused into one picture in the brain. It further assumes that convergence or divergence of the eyes somehow works in the interests of this fusion process. If the physiological images were not combined or unified, we should see two objects instead of one. Ecological optics makes no such assumptions, rejecting the very idea of a physiological image transmitted to the brain. It supposes that two eyes have no more difficulty in perceiving one object than two hands do in feeling one object, or than two ears do in perceiving one event. The dual ocular system registers both the *matching* of structure between the optic arrays at the different points of observation of the two eyes and the perspective *mismatch* of their structure, both the congruence and the disparity, at the same time. The two eyes are not two channels of sensation but a single system. The converging and diverging of the eyes presumably work in the interests of picking up the congruity/disparity information.

Note that two arrays could not possibly be fused in the sense of being united in one location. Neither could two optical images be mixed or combined. They do not need to be. The fallacy of the traditional theory comes from supposing that two *physiological* images have to be fused in the brain, as if one picture were picked up and superposed on the other and then compared, in the manner of a photographer who puts one transparent film on top of another and looks to see if they match. The error is to assume that a unitary mental image can only arise from a unitary brain image, a process in the brain that occurs in one locus.

The human binocular system extracts the similarities of structure between two arrays, I suggest, just as each eye extracts the invariants of structure in its own array. Varying convergence of the binocular system is a kind of exploration, like the varying fixation of each monocular system. The dual array is available for exploration just as much as the single array is. The difference in perspective structure between two arrays is the same as the *change* in perspective structure of one array when one eye moves sideways through the interocular distance. This disparity is neither identity at one extreme nor discrepancy at the other. If the structures were completely identical, nothing would be specified but a hypothetical and ecologically impossible surface called a *horopter*. If the structures were completely discrepant, another kind of impossibility would be specified, and this discrepancy can actually be imposed on the binocular system with a device called a *haploscope*. For example, the array to one eye may consist of vertical stripes and the array to the other of horizontal stripes. In this case the binocular system fails, and the use of one eye is suppressed. The system becomes monocular. Typically, the suppression shifts from one eye to the other, and the result is called *binocular rivalry*. You see horizontal stripes or vertical stripes, or horizontal in one part of the field and vertical in the other, but never horizontal stripes and vertical stripes in the same place at the same time.

This kind of contradiction is very interesting. It is not logical contradiction of the

---

On Binocular Disparity

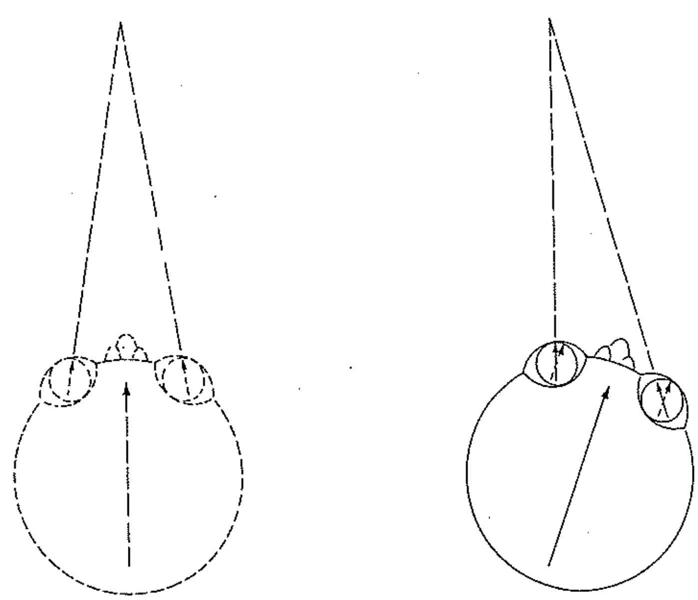
The idea of disparity between two arrays is quite new. It is not the same as the old idea of disparity between two retinal images defined by noncorresponding points on two retinas considered as receptive mosaics. Array disparity rests on ecological optics instead of on physiological optics.

The application of the new optics to binocular disparity has been worked out by Barrand (1978). Although anticipated by the gestalt theorists in the assumption that binocular disparity was "relational," it is a departure from the classical theory of stereopsis. It can handle the neglected fact of occluding edges in stereopsis, for example, which the classical theory cannot.

sort that philosophers have studied since Aristotle. It might be called *ecological contradiction.* It is a discrepancy of information. There will be more about this in Chapter 14 on depiction.

Compensatory Movement When we consider the ambient optic array at the point of observation occupied by an eye instead of the retinal image of an object formed in the eye by light rays, we begin to understand the purpose of the compensatory eye-head coordination. Its purpose is to keep the eye oriented to the unchanging features of the environment for as much of the time as possible while the observer looks around and gets about. It prevents both eyes from wandering or drifting aimlessly. They are linked to the layout of the surfaces. Only if they are stable relative to the world can they look at the world. The eyes do tend to drift or wander when the ambient light is homogeneous, as it is in the presence of an unstructured total field like the blue sky or a dense fog or in ambient darkness. The compensation is automatic, but it is not a reflex response to a stimulus. If the eyes were not anchored, the phenomenal world would "swing" instead of being the fixed frame of reference it is. Indeed, the experience of

Figure 12.2
The turning of the eyes in the head to compensate for the turning of the head in the world.
As the head turns to the right, the eyes turn to the left through the same angle. In this diagram the eyes are converged. (From *The Perception of the Visual World* by James Jerome Gibson and used with the agreement of the reprint publisher, Greenwood Press, Inc.)



vertigo does arise whenever the coordinate compensation breaks down after the stopping of prolonged passive rotation in a swivel chair.

The so-called *swinging of the scene* can be artificially induced in another way when the structure of the field of view of the eye is distorted or reversed by a prism or lens attached to the head by a spectacle frame. The field of view no longer sweeps over the ambient array in the normal manner when the head turns. Thus, the compensatory eye movements no longer serve their purpose, for the sampling of the array by head turning has been disturbed. The eyes are no longer anchored to the environment. The results of experiments on perception when distorting spectacles are worn can only be understood in the light of this fact (Kohler, 1964).

OTHER ADJUSTMENTS OF THE VISUAL SYSTEM

The movements and postures of the eye-head system have now been described. But vision is a highly tuned and elaborate mode of perception, and several other kinds of adjustment occur in the activity of looking. The eyes blink, the tear glands secrete, the pupils enlarge or contract, the lens accommodates, and the retina adjusts for either daytime or nighttime illumination. All these adjustments subserve the pickup of information.

Eye-blinking The eyelids close and reopen at intervals during waking hours to keep the transparent surface of the cornea washed clean and prevent it from drying out. The retina is of course deprived of stimulation during these brief moments of eye-closing, but the interesting fact is that no sensation of darkening is noticed, although even the briefest dimming of the illumination when the electric light falters is noticed. The explanation, perhaps, is simply that this particular kind of flicker is propriospecific. A dimming obtained with an eye-blink is experienced as an eye-blink; a dimming imposed by the illumination is experienced as coming from the world.

The ordinary eye-blink is not a triggered reflex. It may sometimes be stimulated by a puff of air or a cinder on the cornea, but it usually operates to prevent stimulation, not to respond to it. Like the closing of the eyes during sleep, it is an adjustment.

The eyelids work in cooperation with the tear glands. The reason for keeping the surface of the cornea clear is (by analogy with a windshield wiper) that dirt or foreign particles reduce its transparency. The structure of the optic array when the air is clear can be extremely fine, and these very small solid angles specify both the small-scale structure of the near environment and the large-scale structure of the far environment. A dirty cornea still admits light to the eye but degrades the information in the array of light.

The Accommodation of the Lens The combined cornea and lens of the eye constitute a lens system that is said to focus an image of an object on the retina, in accordance with the classical theory formulated by Johannes Kepler. To each radiating point on the near surface of the object there corresponds, ideally, one focus point in the retinal image. The function of the lens is to make it a true point instead of a "blur circle," whatever the distance of the object. The lens accommodates for distance and minimizes the blur (Chapter 4).

The theory of ambient light and its structure is not consistent with this, or at least I do not now understand how it could be made consistent. The notion of nested solid angles based on a nested layout of surfaces, the solid angles being ever changing and never frozen, is of a different order from the notion of radiation from the atoms of a surface and the bringing to a focus of a pencil of these rays from each point of the surface. I do not understand how the former notion could be reduced to the latter, for they are in different realms of discourse.

The focusing of the lens of a photographic camera for a given distance or range of distances is not as similar to the accommodating of the lens of an eye as we have been taught. There is just enough similarity to make optometry and the prescribing of eyeglasses a useful technology. But the eye's lens works as part of the exploratory mechanism of the visual system, along with fixation and convergence, and nothing in the photographic camera is comparable to this. Photographic film does not scan, or look at, or pick up disparity. We are so accustomed to think of deficiencies in accommodation in terms of an acuity chart that we tend to forget this fact. Distinct vision with a fixated eye is not the only test for good visual perception, but that is all the optometrist tries to measure.

The function of the retina is to register *invariants of structure,* not the *points of an image.* The point-to-point correspondence of the theory of image formation does not apply. Ecological optics will have to explain the action of the ocular lens in a different way than does classical geometrical optics. The explanation is not simple.

The Adjustment of the Pupil In Chapter 4, I distinguished carefully between stimulus energy and stimulus information, between ambient light and the ambient array. Light as energy is necessary if the photochemicals in the photoreceptors of the retina are to react, but light as a structural array is necessary if the visual system is to pick up information about the world. Although a clear distinction should be made, ~~it~~ must not be forgotten that stimulus information is *carried* by stimulus energy. There is no information in utter darkness. And, at the other extreme, perception fails in blazing illumination. The photoreceptors are then swamped by the intense light, and the information cannot be extracted by the perceptual system. We describe the accompanying sensation as *dazzle,* and it is propriospecific. The contracting of the pupil of the eye is an adjustment that reduces the tendency of the photoreceptors to be overwhelmed by excessive stimulation.

Physiological optics, concerned with receptors and stimulus energy, is adequate to explain the pupillary adjustment. Ecological optics, concerned with perceptual organs and the information in light, is not required. The different levels of optics correspond to different levels of activity in the visual system. It should nevertheless be noted that the contraction of the pupil in strong illumination and its enlargement in weak illumination work in the interests of information pickup. And the continuous adjustment of pupil size to light intensity is not a series of responses to stimuli but an optimizing process.

The Dark Adaptation of the Retina One kind of adjustment of the retina to the level of illumination involves no movement at all. This is a shift between the functioning of one set of photoreceptors and another set contaidning different photochemicals and with a different level of sensitivity. We have what is called a *duplex retina,* and the duplicity theory of the retina is one of the triumphs of the study of vision at the cellular level. The cones provide for daylight vision and the rods for night vision. The shift of function from cones to rods and the reverse is supplementary to the adjustment of the pupil, which by itself is insufficient for the million-to-one intensity difference between daylight and nightlight.

I have described the advantages of a night retina, a day retina, and a duplex retina in the chapter on the evolution of the visual system in *The Senses Considered as Perceptual Systems* (Gibson, 1966*b,* Ch. 9). Animals of our sort are able to perceive well enough in either a brightly or a dimly lighted environment and thus do not have to make a choice between a diurnal and a nocturnal way of life.

CONCLUSION: THE FUNCTIONS OF THE VISUAL SYSTEM

The anatomical parts of the visual system are, approximately, the body, the head, the eyes, the appurtenances of an eye (eyelid, pupil, and lens), and finally the retina of an eye, which is composed of photocells and nerve cells. The body includes all the other parts, and the cell includes none of the others. All these components are connected with the nervous system, and all are active. All are necessary for visual perception. Both the parts and their activities form a hierarchy of organs. At the top is the body, then the head, and then the eyes. Being equipped with muscles, the parts can move, each in its own way—the eyes relative to the head, the head relative to the body, and the body relative to the environment. Hence, all move relative to the environment, and I suggested that their purpose is perceptual exploration. At the level of the single eye, the eylid wipes, the lens accommodates, and the pupil adjusts. Muscles are also required for these activities, but they are not bodily movements in the sense used above. At the bottom level, the retina and its cells adapt to external conditions but the activity of the retina does not depend on muscles. At all levels the activities are *adjustments* of the system instead of reflex reactions to stimuli, or "motor" responses, or responses of any kind, for that matter.

The body explores the surrounding environment by locomotion; the head explores the ambient array by turning; and the eyes explore the two samples of the array, the fields of view, by eye movements. These might be called *exploratory adjustments*. At the lower levels, eyelid, lens, pupil, and retinal cells make what might be called *optimizing adjustments*. Both the global structure and the fine structure of an array constitute information. The observer needs to look around, to look at, to focus sharply, and to neglect the amount of light. Perception needs to be both comprehensive and clear. The visual system *hunts* for comprehension and clarity. It does not rest until the invariants are extracted. Exploring and optimizing seem to be the functions of the system.

THE FALLACY OF THE STIMULUS SEQUENCE THEORY

The traditional assumption has been that we perceive the world by means of a sequence of stimuli. When we look at the scene in front of us, we see it in a succession of *glimpses* analogous to snapshots, each glimpse corresponding to a pure fixation. Similarly, when we look around at the whole environment, we perceive it as a sequence of *visual fields* analogous to pictures, each field corresponding to a posture of the head.

THE LORGNETTE TACHISTOSCOPE

I once devised a sort of test of what perception would be like if it really consisted of a sequence of snapshots. I mounted the shutter of a camera on a handle so that it could be held close to one eye and triggered with a finger, giving a wide-angle glimpse of the environment for a fifth of a second or less. The other eye was covered. The subject was led up to a table on which was a collection of familiar objects and told to keep looking until he knew what was there. Because he couldn't scan the table with his eye, he had to scan with his head and trigger the shutter for each new fixation.

Perception was seriously disturbed, and the task was extremely difficult. What took only a few seconds with normal looking required many fixations with the lorgnette tachistoscope, and there were many errors. I now begin to understand why.

Both the glimpses and the pictures of the world have been vaguely identified with retinal images. But this assumption that we perceive in a sequence of pictures, either glimpses or fields, is quite false.

A visual fixation is not at all comparable to a snapshot, that is, a momentary exposure. The eye has no shutter. The eye scans over the field. The fovea is transposed over the sample of the array, and the structure of the array remains invariant. Not even a visual field at a head-posture is comparable to a picture in a sequence of pictures (although I used to think it was). The field sweeps over the ambient array with progressive gain and loss at its leading and trailing edges, and the ambient structure remains invariant. No succession of discrete images occurs, either in scanning or in looking around.

The transposition of the fovea over the sample of the array and the sweeping of the edges of the field of view over the ambient array are propriospecific; they specify eye turning and head turning respectively, which is precisely what they should do. The former is visual kinesthesis for an eye movement, and the latter is visual kinesthesis for a head movement.

The formula of visual kinesthesis for the exploratory movements of the eyes, and of the head and eyes together, resolves a number of long-standing puzzles concerning visual sensations. It cuts a Gordian knot. The century-old problem of why the world does not seem to move when the eyes move and the analogous problem of why the room does not appear to go around when one looks around are unnecessary. They only arise from the assumption that visual stimuli and visual sensations are the elements of visual perception. If the visual system is assumed instead to detect its own movements along with extracting the information about the world from the ambient light, the puzzles disappear. I shall have more to say about this later.

The false problems stem from the false analogy between photography and visual perception that everyone has taken for granted. A photograph is an arrested moment of a changing array. The film has to be exposed, and the so-called latent image must be developed, fixed, and printed before it becomes a picture. But there is nothing even faintly comparable to a latent image in the retina. It is misleading enough to compare the eye with a camera, but it is even worse to compare the retina with a photographic film.

The stimulus-sequence theory of perception underlies much of modern thought, not only the thinking of philosophers, psychologists, and physiologists but that of the man in the street. It is reinforced by comic strips and cartoons and news photographs, and the movies above all. As children we do learn much about the world at second hand from picture sequences, so much that we are strongly tempted to interpret firsthand experience in the same way. Everybody knows what pictures are, and textbooks tell us that retinal images are pictures. I said so myself in my book on the visual world, and the only problem that bothered me was how a *sequence* of images could be converted into a *scene* (Gibson, 1950*b*, Ch. 8, pp. 158 ff.). I did realize that something was wrong with this assertion, but it took me years to detect the fallacy.

THE THEORY OF THE CONVERSION OF A SEQUENCE INTO A SCENE

It sounds plausible to assert that a sequence of images is converted into a scene. "At the circus, for example, you may watch the tightrope walker, then look at the performing seals, pause to observe a clown, and return to the tightrope walker. Although you have had a succession of impressions the events are perceived as coexisting" (Gibson, 1950*b*, p. 158). You are aware of three concurrent events in three different places, all going on at the same time, but you are not aware of the successive order in which they have been fixated. An adjacent order of places, a whole scene, must have been obtained from a successive order of sensory inputs, a sequence, by some sort of *conversion*. The sequence of smaller fixations with which you observe the tightrope walker in his smaller situation is noticed even less. You look back and forth between his feet and his hands, say, but what you see is the whole act.

The hypothesis of *conversion* is consistent with the traditional theory that successive inputs of a sensory nerve are *processed*, that a series of signals is *interpreted*, or that the incoming data of sense are *operated on* by the mind. Sensations are converted into perceptions, and the question is, how does this come about? In the case of successive retinal images, the process is supposed to be that of memory. It may be called short-term or primary memory, or immediate memory as distinguished from long-term memory, but the basic assumption is that each image has to be held over, or stored in some sense, in order for the sequence to be integrated, that is, combined into a unit. The present percept is nothing without past percepts, but past percepts cannot combine with the present except as memories. Every item of experience has to be carried forward into the present in order to make possible perception in the present. Memories have to *accumulate*. This is the traditional theory of memory made explicit. It is full of difficulties, but it has seemed to provide the only explanation of how images could be integrated.

The error was to suppose in the first place that perception of the environment is based on a sequence of discrete images. If it is based instead on invariance in a flow of stimulation, the problem of integration does not arise. There is no need to unify or combine different pictures if the scene is *in* the sequence, is specified by the invariant structure that underlies the samples of the ambient array.

The problem of explaining the experience of what I once called the unbounded visual world (Gibson, 1950*b*, Ch. 8) or what I would now call the surrounding environment is a false problem. The retinal image is bounded, to be sure, and the foveal image has even smaller bounds, but the ambient array is unbounded. If the stimulation of the retina, or that of the fovea, is accepted as basic, another problem arises as well, how to explain the experience of a *stable* visual world. The stimulation of the retina is continually shifting, but this is also a false problem, for the structure of the ambient array is quite stable.

SUMMARY

One sees the environment not just with the eyes but with the eyes in the head on the shoulders of a body that gets about. We look at details with the eyes, but we also look around with the mobile head, and we go-and-look with the mobile body.

A theory of how the eye-head system works has been formulated in this chapter. A theory of how the system works during locomotion was formulated in the last chapter. The exploratory adjustments of the eye-head system (fixation, saccadic movements, pursuit movements, convergence-divergence, and compensatory movements) are easier to understand. Even the optimizing adjustments of the lens, the pupil, and the photoreceptors are more intelligible when we consider optical information instead of stimuli.

The flow of optical stimulation is not a sequence of stimuli or a series of discrete snapshots. If it were, the sequence would have to be converted into a scene. The flow is sampled by the visual system. And the persistence of the environment together with the coexistence of its parts and the concurrence of its events are all perceived together.

THIRTEEN

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 mahines 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 kinesthesis 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 KINESTHESIS 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 kinesthesis 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 kinesthesis specifies locomotion relative to the environment, whereas the other kinds of kinesthesis 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 kinesthesis but the same visual kinesthesis. 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

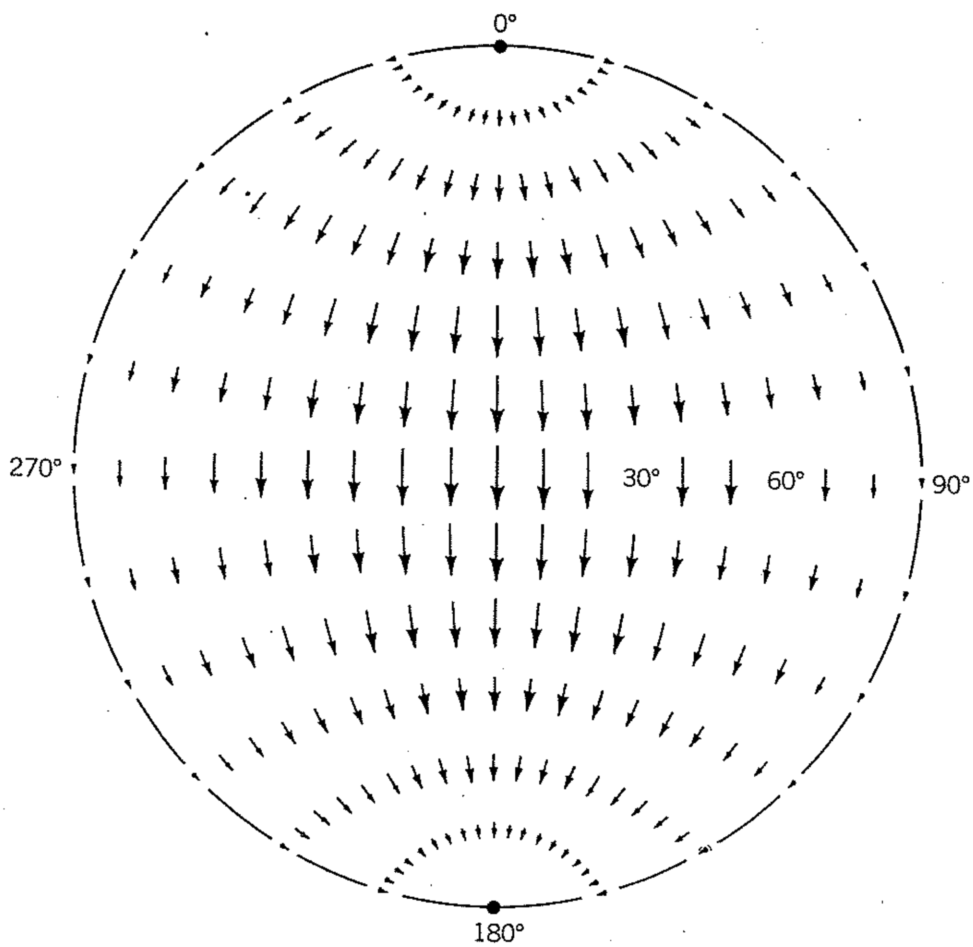
sample only one or the other, by looking "ahead" or by looking "behind." Note that a reversal of the flow pattern specifies a reversal of locomotion.

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

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.)



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.

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,

---

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.

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.

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°. 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 kinesthesis*, 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.

FOURTEEN

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, 1966*b*, 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, 1966*b*) 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

---

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, p. 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.

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 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 therefor 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

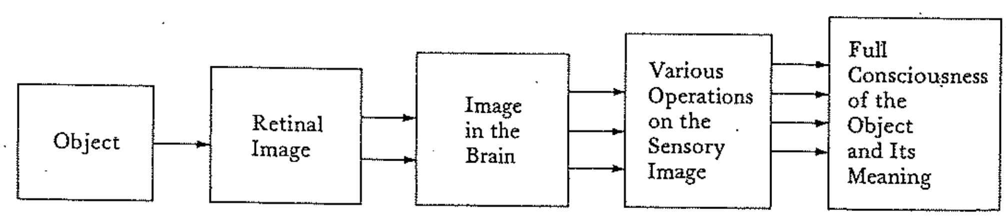
The understanding of signs, or indicators, or even *clues*, in the manner of a police detective

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

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



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.

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, 1966*a*). 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 *pre-existing* 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.