JAMES J. GIBSON
CONSTANCY AND INVARIANCE IN PERCEPTION

Why do things look as they do? Why do they appear, on the whole, pretty much as they are? The makers of things-to-be-seen need to understand how we see. So also, for that matter, do the *consumers* of things-to-be-seen—those of us who enjoy looking at artifacts of any sort, particularly at pictures. Perceiving for its own sake—contemplating, registering, detecting, discriminating, and comprehending—is not only a pleasure, it is useful. It can even be thought of as a sort of discipline. Artists have long believed that in making "art" they learn to perceive "nature", and that they can thereby show the rest of us how to see it better. Artists are bound to be psychologists insofar as they are concerned with the psychological question of how a perceiver does what he does. By the same token the perception-psychologist ought to be concerned with art.

The psychology of perception, however, is a large field with a long history, a bewildering series of controversies, and a vast amount of evidence on both sides of the issues. It includes the study of the physiology of sensations at one extreme, and the philosophy of knowledge at the other. A bare introduction to the subject is given in a long and difficult book by E. G. Boring entitled *Sensation and Perception in the History of Experimental Psychology*.[1] Any artist who even attempts to read everything important that has been written about it is a bold man. Gombrich has recently done so in his study of the psychology of pictorial representation.[2] Although he found the literature illuminating, he also found it full of contradictions.

The central puzzle of perception, I believe, is the problem of what is called constancy. This term is not very familiar outside of psychology, and I shall try to explain what the term means and why it is considered so important.

Constancy is the tendency to perceive an object as the same despite changing sense-impressions. One sees the size of an object fairly well at quite different distances from the eye. One sees the shape of the face of an object correctly even when it is slanted or inclined to the line of sight, i.e., foreshortened. The impression of "extent" changes with distance and the impression of "form" changes with inclination, but the perception, on the whole, does not. Likewise both the perception of surface-color and of white-black do not seem to vary much with the color and intensity of the light entering the eye (which change with varying illumination, or shadows) but depend on a property of the surface—the differential absorbing and reflecting of incident light. These three facts are called size constancy, shape constancy, and color constancy, and a great number of experiments have been carried out in psychological laboratories to discover the conditions under which they occur, to measure the tendencies, and to test the theories which purport to explain them. It is fair to say that these facts are as much a puzzle now as they were half a century ago when they began to be studied.

The constancies of size, shape, and color are not the whole of the matter. It is becoming clear that there is a much larger set of constancies in perception, not so easily labelled. All of them involve a discrepancy between the sense impression and the experience of the ordinary naïve observer. Some of these are worth describing, in order to show the scope of the problem.

Constancy of perceived space. Not only is the size of a single rigid body seen to be the same at different distances and the shape the same from different viewpoints, but also the size and shape of two bodies in different places can be compared. Their dimensions and proportions are visible. Moreover, the apparent distances between bodies, as well as the objects themselves, remain constant. One can match the

separation of two things which are far off with two things nearby. Similarly, one can say whether the edges of a road running off into the distance (or two stretched strings controlled by an experimenter) are parallel or not. What remains constant in such situations is the scale of things and the intervals between things. It is the ground, not just the bodies resting on the ground, that keeps the same size and shape in experience.

One can put these facts in another way by saying that what men are conscious of are the surfaces of their environment, and the *layout* of these surfaces. So far as the evidence goes, this also holds true for children and, on behavioral evidence, for animals. Observers are not ordinarily conscious of the patchwork of colors in the field of view as this is determined by the laws of perspective. Instead they perceive the environing surfaces with their edges, corners, slants, convexities, concavities, and interspaces. These are, of course, the pathways and obstacles, the places and things, the goals and the dangers of the terrestrial world. They are identified by their surface properties, including texture and differential reflectance of light. They have to be seen as constant, where they are, in order to be identified for what they are.

Historically, the central problem of perception has been taken to be how we see depth and distance, the so-called third dimension of space. The psychologist and the painter have been led to ask what the clues or cues may be for tridimensional perception as distinguished from bidimensional sensation. It begins to be evident, however, that the heart of the problem is not so much how we see objects in depth as how we see the constant layout of the world around us. Space as such, empty space, is not visible but surfaces are.

The apparent rigidity of the phenomenal world. Whenever an observer moves from place to place, the pattern of his field of view—that is, the optic array that determines his retinal image[3]—undergoes a perspective transformation. This follows, of course, from the facts of perspective at a stationary point. There is a deformation of this array as a whole and in every part. The visual sense-impression, therefore, changes with every change of position. Unless an observer holds his head unnaturally still and fixes his eye, his visual field is alive with transformation. A sensation of "form", then, is an extreme rarity in life; what normally stimulates the eye is a continuous transformation in time. Nevertheless perception is of rigid things, of a rigid ground, and of constant separations between fixed things. The phenomenal world is not distorted during locomotion, although it ought to be if perception is based on sensation.

The same perspective transformation of the field that occurs when the head is moved from right to left appears simultaneously in one eye relative to the other when both eyes are open. This disparity of the two fields, the mismatch of pattern, yields a sensation called double imagery. But perception is not doubled and the mismatch can only be noticed if a man attends to his subjective sense-impressions instead of to the world.

Other discrepancies between sensation and perception. The set of color-patches which make up the visual field continually change in a jerky fashion as the eyes scan the array of ambient light. We move our eyes from one fixation point to another several times a second during waking life. We also blink frequently. The sense impressions, therefore, are highly unstable and interrupted. But an obvious

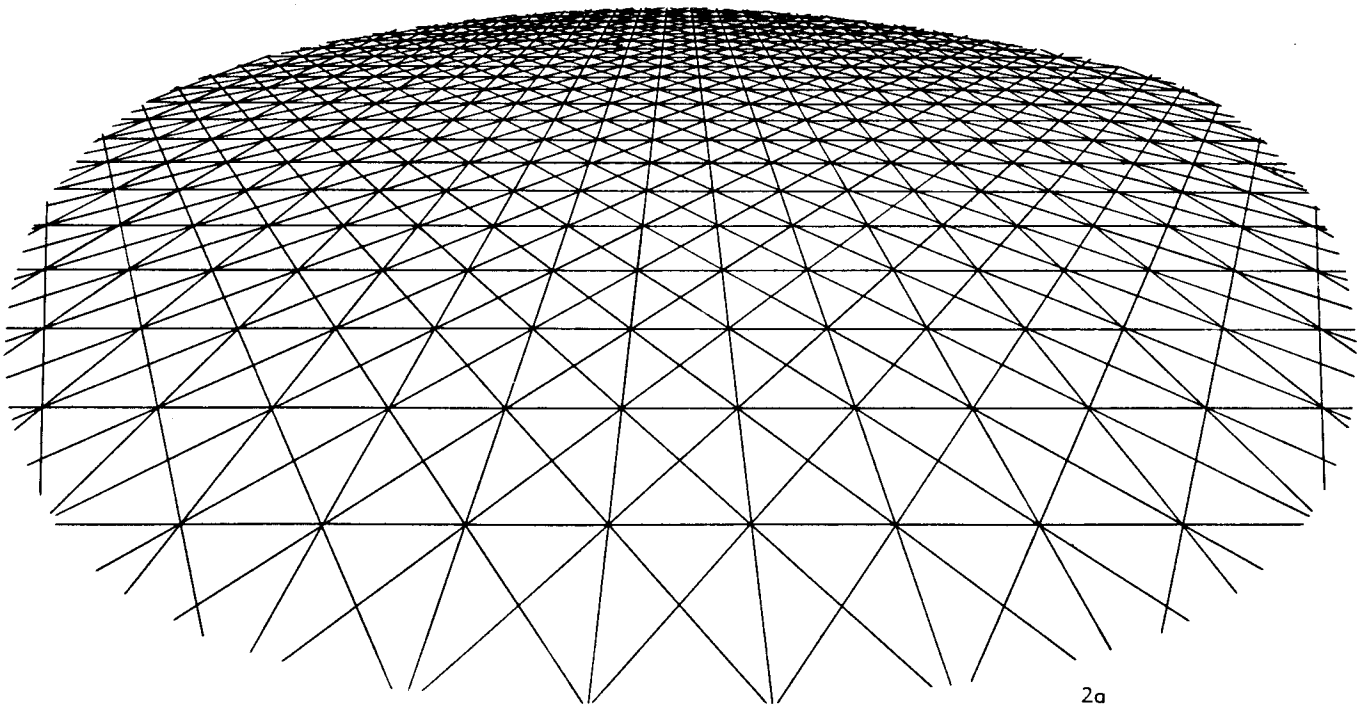

Fig. 1. *The Perspective of a Pavement.* Note that when this picture is observed from the proper station-point the ground tends to emerge in experience, not so the visual sensations. The blocks appear equal in size, the same in shape, with edges parallel. They recede into the distance instead of getting smaller or denser. The scale remains constant. The space is Euclidean. (From Gibson, *Perception of the Visual World,* Houghton Mifflin, 1950)

2a

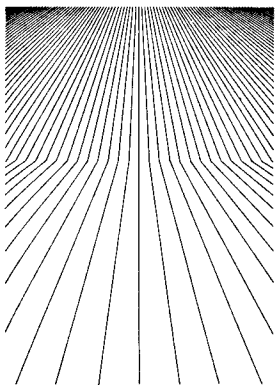

Fig. 2a. *The Visual Stimulus for a Corner.* Whenever the gradient of density in an optic array abruptly changes rate in this manner one sees a "corner", i.e., two adjoining surfaces at an angle. This one is concave. The mathematical information would be the same for any other texture. It is an invariant piece of information in the light to a station-point.

2b

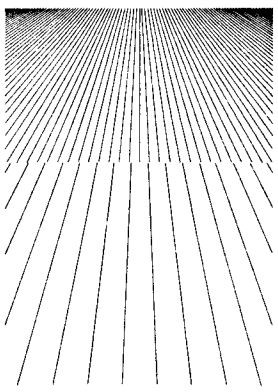

Fig. 2b. *The Visual Stimulus for an Edge.* Whenever the gradient of density in an optic array abruptly changes *density* in this manner, one sees an "edge" with depth behind it, i.e., a step downward, a falling-off place. This mathematical discontinuity is also an invariant. It carries external meaning —important meaning for anyone walking toward the edge. (From Gibson, *Perception of the Visual World*)

characteristic of perception is its stability and continuity. The world does not seem to move as the retinal image moves over the retina. Here is another sort of constancy.

Human visual sensation is clear only in the center of the field corresponding to the central fovea of the eye at each momentary fixation. It is a fleeting impression. But visual perception, being extended over time and depending on the whole array of ambient light, may be clear in all directions. The momentary sensation is bounded by the margins of the cone of light rays that can enter the eye. But visual perception is unbounded. We are aware of a world that surrounds us like a panorama, not a cone of rays. I have described these contrasts more fully in Chapter 3 of *The Perception of the Visual World.*[4]

Constancy in perception other than visual. The stability of perception with unstable sensations holds true for all the receptive systems, not only the visual system. Just as the visible world does not seem to rotate 90° when one lies down on his side, although the retinal image does, so the tangible ground and the direction of gravity do not seem to swing upward, although the tactual impression has shifted from one's feet to one's flank and the weights in the sacs of the inner ear now pull sideways to the head. Instead of the ground changing, one feels that *he* has changed and the earth has not.

Sensations of touch are often radically different from perceptions of touch. One gets a tactual impression whenever the skin is pushed in, usually by contact with a solid body. But when one touches an object with the fingers he feels the object, not the contacts. When you move your groping fingers over an unfamiliar object with eyes closed you will experience its shape, size, proportions, and rigidity but you will be almost wholly unaware of the sequence of cutaneous impressions. The same single object, a pencil for example, is felt whether you hold it with two, three, four, or five fingers, and this means with two, three, four, or five different sensations at different places on the skin. In short the perception is unitary despite diversity of the impressions. When you press on a surface lightly or heavily you do not feel the changing intensity of the impression; you feel only the unchanging solidity of the object. When you move your hand over the edges and corners of a rigid body, you do not feel the cutaneous motion over the skin; you feel a motionless object with a moving hand. The perception is constant and stable although the sensations are changing and mobile.

The channels for stimulus information that we have arbitrarily separated and called "senses" are normally active and exploratory, not passive and receptive. It begins to be clear, I think, that the passive arousal of sensations, as these have been studied by sensory physiologists, is not typical of the way perception works in life.

One more example may be offered, from hearing. The sensory qualities of auditory experience are said to be loudness, pitch, and tonal complexity. A great deal is known about the corresponding variables of physical sound. But the perceptions arising in auditory experience are of outer happenings, inanimate events, cries, and the speech of our fellow men. The perceiving of speech sounds is largely independent of loudness, pitch, and tonal complexity. The proof is that we hear the same speech whether it is whispered or shouted, voiced or sung, produced by male or female vocal organs. The critical speech sounds, the phonemes, depend on properties of sound that are invariant, that is, properties which do not change when the level of intensity or the level of pitch is altered.

Theories of the invariance of perception with varying sensations. How can the experience of a constant world arise from the ever changing flux of sensory impressions? This is the central puzzle. If the data of sense vary, how can the perception of unvarying places and things be explained? The constancies of size, shape, and color of objects together with all the other sorts of constancy are the principal reason for theories of perception.

Theories of perception go back for hundreds of years and have occupied the best efforts of some of the greatest thinkers in history. They cannot be here summarized, but they can be classified roughly. They seem to fall into three types: first, those that appeal to innate ideas or the rational faculties of the mind for making the sensory data intelligible; second, those that appeal to past experience, memory, or learning for supplementing and interpreting the sensory data; and third, a recent theory which asserts that the sensory chaos is organized by a spontaneous process of self-distribution in the brain.

It should be noted that all these theories assume without question that sense impressions are somehow the cause of perception but not a sufficient cause. They are taken to be the occasion for perception, the basis for it, or the raw material from which perception is constructed. These theories all take for granted the poverty of the senses and seek for a special process in the mind or the brain to supplement them. They assume that the organs of sense are passive, or merely receptive, accepting whatever physical stimulation enters as if they were merely windows. The activity of perception is supposed to be an internal or subjective process. Meaning is supposed to come from inside, not from outside. These assumptions, as will appear later, can be challenged. It might be that no special process is necessary to explain perception, and that in fact perception is not based on sensation. But this is getting ahead of the story.

The theory of innate ideas and faculties. It is possible to suppose that perceiving, although occasioned by having sensations, is chiefly a matter of intuition. The idea of a constant and fixed Euclidean space, for example, may be simply a part of the inborn capacity of the human mind. The continually changing perspectives of visual sensation are interpreted in terms of this abstract concept as rigid objects. But the mind is informed by its own preconceptions, not by the sense impressions.

The theory of the accumulation of past experience. By far the most popular theory of perception is that of empiricism, as contrasted with nativism. Concepts and general ideas are taken to be learned, not inborn. They are the residue of all the fleeting perspectives of the past. We have learned that objects are constant in shape and size and color and therefore we unwittingly interpret our sensations in accordance with what we have learned about objects. According to Helmholtz, the process is one of "unconscious inference", the sense data serving only as clues to the real nature of the objects.

This explanation, or one of its many variants, is so widely accepted that many people have never heard of any other. It has the virtue of emphasizing training or learning instead of the mysterious faculties of the mind. It allows for the possibility of the improvement of our perceptual abilities with the accumulation of memories. But as an explanation it is weak and it was destructively criticized a generation ago by the proponents of another theory, especially by Koffka.[5]

The theory of sensory organization. Gestalt theory. Koffka asked how a perceiver could be supposed to learn about objects if all he had to go on were the flux of meaningless sensations. How could a child learn to see an object without ever seeing one? A theory of association, or of the learning of clues, cannot explain it unless the objectivity of an object is given at some time. There was evidence to suggest, moreover, that inexperienced animals and children did not behave as if they confused a large far object with a near small object—in short that they did not see their retinal images in the first place.

The *Gestalt* theorists proposed that a process of sensory organization instead of association was the explanation of perception. The nature of the nervous system is such that organization takes place spontaneously. Experience is structured; it comes in a field and, at the very least, there is always a "figure" on a "ground". It is never wholly meaningless, even at the start.

Koffka also made the acute observation that, in perceiving an object, we do not separately see a retinal size and then a distance, or a retinal shape and then a slant; instead we see all at once a size-at-a-distance, or a shape-at-a-slant. The relationship between the members of these pairs is invariant in experience, he noted, the visual angle for a given object being reciprocal to distance and the perspective flattening being concomitant with slant. Form and space are linked together, as it were, not separable. It is only a step from this idea to the hypothesis that a shape-at-a-slant may actually be given as an invariant within two variables of optical stimulation, although this is a step that Koffka did not take.

The *Principles of Gestalt Psychology*[5] was certainly the most knowledgeable book on visual perception ever written, and it is still a good foundation for new knowledge. *Gestalt* theory has been called a modern form of nativism and there is a grain of truth in this observation, for the hypothetical "laws of visual organization" bear some analogy to innate forms of apprehension. But the trend or direction of the theory was novel and its emphasis on structure, order, articulation, pattern, and the "total field" of perception is still to be followed up. The proposed laws of organization have not been verified experimentally. But some kind of organization in perception is a fact. The question is where it comes from. Perhaps it comes from outside, not inside.

The relation between stimuli and their sources in the environment. It was only in the era of *Gestalt* theory that the paradox of perceptual constancy became clear, although it has been the root of the difficulty all along. In my terminology, it is the invariance of perception with varying sensations. The other side of the problem is, the invariance of physical objects with what seem to be varying physical stimuli.

One must not confuse the stimuli for the eyes, the ears, and the skin with the sources of these stimuli in the environment. The light, the sound, and the mechanical energy respectively must be distinguished from the objects that reflect light, emit sound, or come in contact with the skin. The impinging physical energy is called the *proximal* stimulus, the stimulus proper, and the external object or event is called the *distal* or *distant* stimulus. The sense organs are excited only by proximal stimuli but what the perceiver is aware of are distal stimuli. This is the other side of the paradox of perceptual constancy. The environment is constant, the stimuli are changing, the sensations are changing, and the perception is constant. The distance, depth, solidity, and permanence of the environment seem to be lost in the proximal stimuli and in the sensory impressions but they turn up again, almost miracu-

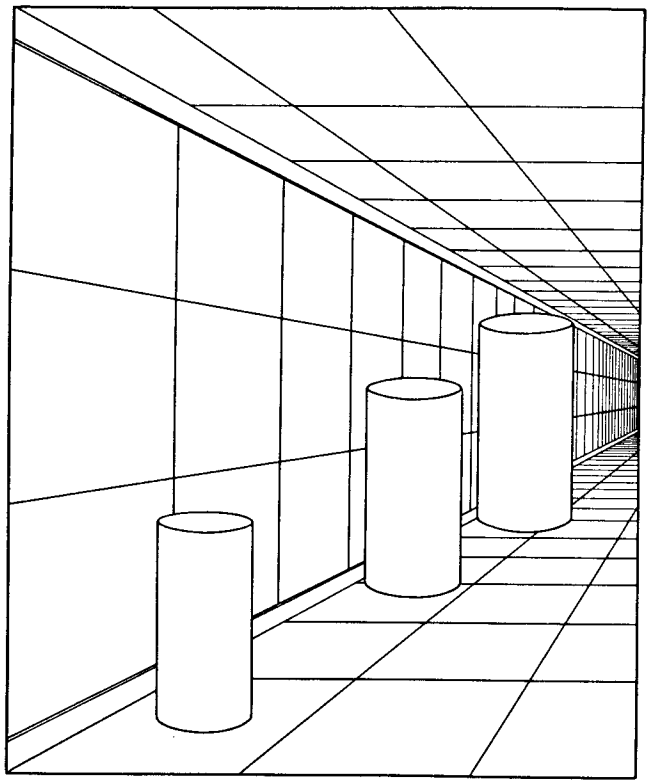

Fig. 3. *The Reciprocity of Size and Distance.* The pictorial sizes of the images of the cylinders in this drawing are equal, but the phenomenal sizes of the virtual cylinders in the space of the picture are different, the farthest being two and a half times larger than the nearest. One sees the size-at-a-distance, not the sensation of size. The distance of each cylinder is given by its place in the gradient of texture. The size of each cylinder is given by the number of texture elements it intercepts. The size of an object depends on the scale of the space where it is and this tends to remain constant even in pictorially represented space.
(From Gibson, *Perception of the Visual World*)

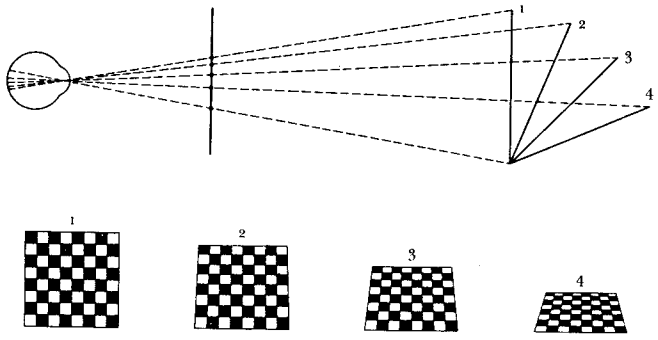

Fig. 4. *The Reciprocity of Shape and Slant.* The phenomenal shape of a checkerboard remains constant despite change of slant. One tends to see the square at a slant, not the trapezoidal sensation. One explanation is that the impression of trapezoidal shape is corrected or compensated by the clues for slant. But a simpler explanation would be that the information for both the shape and the slant of the surface is given by the texture. The surface-shape is what counts, not the ghost-shape of its projected edges.
(From Gibson, *Perception of the Visual World*)

lously, in perception. The third dimension of space is said to be lost in the two-dimensional visual image but restored by some activity of the mind (or the memory, or the brain) in visual experience.

The study of the senses has always involved the study of proximal stimuli, and experimenters have been applying energies to the sense organs of their subjects to see what happens for more than a century. A great deal more is now known about stimuli than when the theories were first formulated. It used to be thought that they were bits of energy, points and movements. But the evidence accumulates that the field of stimulation and the flow of stimulation are what is important in arousing sense organs. Both the field and the flow must have a pattern or structure. What excites the eye, for example, is not an even distribution of light but at least one contrast or margin; and not a fixed level of intensity but a change. The effective stimuli are gradients and transients, and this is as true for the ear and the skin as it is for the eye.

It used to be thought that stimuli could not possibly represent their sources in the world. Objects do not get into the eye; only light can enter. It is true that an object cannot be *replicated* in light rays, but the properties of its surface can be *specified* by them. Perspective carries some information about the object, and change of perspective carries still more.[6] An observer who has "looked at all sides of a thing" is one who has sensed it in all possible perspectives. As he moves around it his eyes are stimulated by a whole family of serial transformations. The perspective forms change from moment to moment, but note that they change in perfectly orderly and in completely reversible ways. It has been taken for granted that the ever changing form of the stimulus is a chaos which cannot possibly contain the solid form of the object. But perhaps it can. The true shape of the object may be implicit in the serial transformations. In that case, order does not have to be imposed on the momentary stimuli; it is already there in the sequential stimulus.

The relation between orderly stimuli and their sources, then, may not be as tenuous as we have thought in the past. The varying stimuli which go with unvarying objects in the world may have an unvarying component. And this leads us back to the theory of perception.

A new approach to the invariance of perception. I have a suspicion that the theories of perception have been on the wrong track. It has often been true in the history of thought that a puzzling problem cannot be solved but has to be reformulated, and perhaps this is true of the problem of constancy.

The invariance of perception with varying sensations ought to lead to the conclusion that sensations do not cause perceptions. But that seems to go against common sense, for we perceive only when the sense organs are stimulated—otherwise we only imagine or remember or guess the facts of the world. The way out may be this: that sensory experience is a special self-conscious kind of awareness while perceptual experience is unselfconscious and direct. The latter does not depend on the former. Perception is not mediated by sensations, nor based on sensations.

This requires that the stimuli causing sensations be different from those causing perceptions. This is a new idea, not at all evident but very promising when it is considered. The individual is bathed in a sea of energy at all times, and the stimulus energies that his receptors can pick up are a flowing array. Light, sound, and physical touchings are patterned, both simultaneously and successively. The hypothesis is that the flowing array has two components, one of change and one of non-change.

This hypothesis can be sharpened by borrowing from mathematics two notions. One is that of *transformations* and the other is that of *invariants under transformation*. These terms, although not taught us in beginning geometry, are fundamental. They are much more useful than the notion of "form" as the Greeks conceived it.[7] Stimuli are not static forms but serial transformations which are nevertheless lawful. A static form is simply a special case of continued non-transformation. Transformation, in mathematics, is not simply change but permanence in change. The specific hypothesis is that the invariant component in a transformation carries information about an object and that the variant component carries other information entirely, for example, about the relation of the perceiver to the object. When an observer attends to certain invariants he perceives objects; when he attends to certain variants he has sensations.

This hypothesis is incomplete (and probably the same thing can be said in another way) but it explains the constancy of objective perception without recourse to theories of a subjective process. It also explains how perception can be, in effect, focused on the sources of stimulation, although it is in fact wholly dependent on the stimulation itself. For the permanent properties of the outer world—its texture, edges, layout, solidity, stability, and the fact of gravity—are undoubtedly specified by invariant properties in the visual and tactual stimulus flux. We can thus understand why perceiving is so often correct without recourse to difficult philosophical theories.

Extra hypotheses are needed about what I have called the variant component of stimulus transformation. They are speculative, and need investigation. For one thing, we must suppose that there is a difference between perspective transformations and other transformations. In vision, a perspective transformation results from a movement of the object relative to the observer or of the observer relative to the object. But if the thing observed is fluid or viscous or ephemeral or changeable in itself, not simply moveable, a quite different sort of transformation occurs in the light to the eye. The difference is quite noticeable.[8] If the thing observed is broken or disrupted still another transformation occurs, a permutation of order in the stimulus, and the event is visible as such.

Another extra hypothesis is needed about self-produced transformations of the stimulus-array as contrasted with those not dependent on action of one's own. In moving one's eyes, or head, or body, or in manipulating external objects, one changes the optical stimulus array in whole or part. Each action has its own family of transformations, and each is accompanied by a family of other feelings. We can thus control our responses and explore the possibilities of new stimulation. Locomotion, for example, is guided in this way. We even hear our own footsteps. But some changes in the pattern of light or touch or sound are not self-produced in this circular fashion. They are object-produced. The absence of circularity may well be the feature of these stimuli that gives the corresponding perceptions their external reference. A clue to the whole muddle of explaining how experience can be both external and internal, both objective and subjective, may lie here.

Still another hypothesis is needed about the way in which perception develops in the child, and how discrimination improves in the adult. Presumably it is a matter of the growth and the education of attention. The theory that the infant has only meaningless sensations based on raw stimuli and later enriches these impressions with memories is not good enough. The child learns, but what he probably learns is to fix on the subtle variables of stimulation instead of only the crude ones. He does not have to construct a constant world out of ever changing perspectives but he does have to discover the finer properties of the world that lie hidden in these transformations.

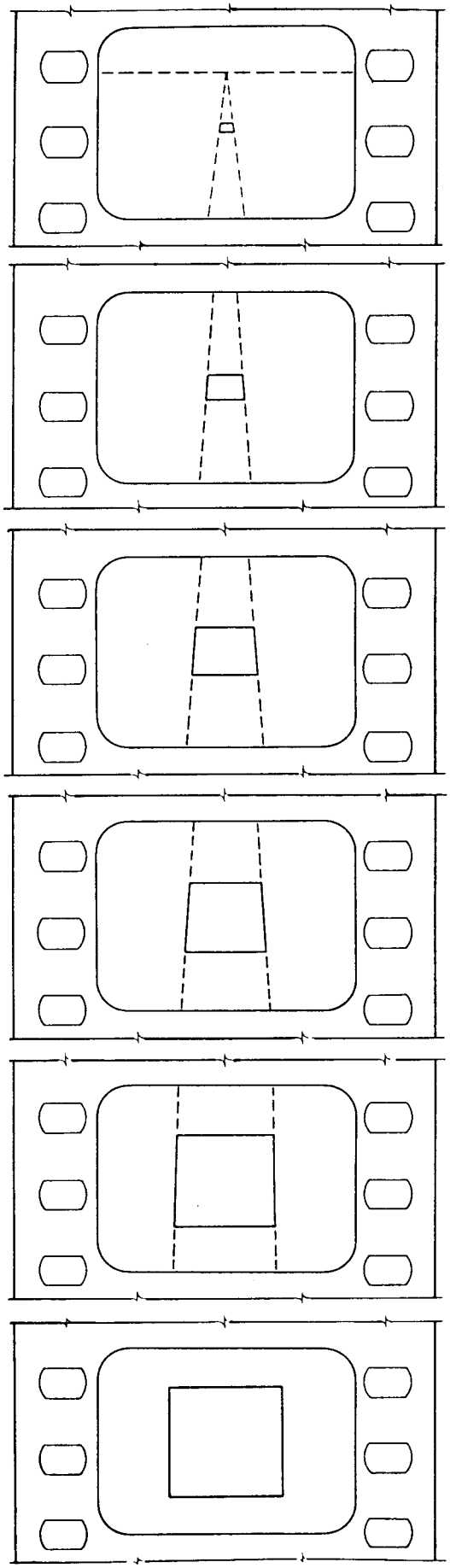

Fig. 5. *A Serial Transformation.* This sequence, from top to bottom, shows the perspective transformation of the projection of a unit of pavement as one walks closer to it. We are accustomed to think that the form changes over time, but we forget that certain important properties of the form do not change; the permanence is just as important for our perception as the change. The transformation specifies approach. The invariant under transformation specifies the constant object. (From *Perception of the Visual World*)

1. E. G. Boring, *Sensation and Perception in the History of Experimental Psychology*, New York, Appleton-Century-Crofts (1942).
2. E. H. Gombrich, *Art and Illusion: A Study in the Psychology of Pictorial Representation*, New York, Pantheon (1960).
3. J. J. Gibson, "Ecological optics," in *Vision Research* (1:1961), pp. 253–262.
4. J. J. Gibson, *The Perception of the Visual World*, Boston, Houghton Mifflin (1950).
5. K. Koffka, *Principles of Gestalt Psychology*, New York, Harcourt Brace (1935).
6. J. J. Gibson, *The Perception of the Visual World*, Ch. 7.
7. J. J. Gibson, "What is a form?," in *Psychological Review* (58:1951), pp. 403–412; idem., "Optical motions and transformations as stimuli for visual perception," in *Psychological Review* (64:1957), pp. 288–295.
8. K. von Fieandt and J. J. Gibson, "The sensitivity of the eye to two kinds of continuous transformation of a shadow-pattern," in *Journal of Experimental Psychology* (57:1959), pp. 344–347.