THE PERCEPTION OF THE VISUAL WORLD

Gibson

PRINT
612.84
G
Gibson, James J.
THE PERCEPTION OF THE VISUAL WORLD.

THE RIVERSIDE PRESS

Cambridge, Massachusetts

Printed in the U.S.A.

Editor's Introduction

Artists and philosophers as well as physicists, physiologists, and psychologists have long been interested in isolating the factors which make possible man's visual world. Physicians who are interested in the eye and its diseases, illuminating engineers, photographers, designers of photographic and motion picture equipment, and many others in pure and applied science are also concerned with certain aspects of this complex subject.

The present book represents the culmination of nearly a quarter of a century of study of visual phenomena by its able author. He has approached the problem in an eclectic manner. In its pages the point of view of the student who is being introduced to the subject is never forgotten.

The author emphasizes the fact that a fundamental condition for seeing is an array of physical surfaces which reflect light that is then projected on the retina. He further gives new emphasis to the importance of considering the retinal images of each eye as involving steps and changes in gradients of light.

The student will find in this volume an interesting discussion of the old and difficult problem of the nature of visual depth. The author also deals with the constancy of the characteristics of perceived objects in relation to geometric space and many other related topics.

Throughout the book theories of perception are carefully evaluated. Certainly the present volume can be recommended to all artists and

This book has a great deal to say about the physical stimuli which are the correlates of perception, but relatively little to say about the activities of the sense organs and the brain which are also the correlates of perception. The writer has elected to study psychophysics rather than psychophysiology because he believes that it offers the more promising approach in the present state of our knowledge. This is not to minimize the importance of physiology. Such books as Bartley's *Vision: A Study of its Basis* (6) are essential to an understanding of the complete process. What we lack, however, is an application of the psychophysical methods to perception.

A psychophysics of perception may sound to some readers like a contradiction in terms. This book undertakes, however, to justify and make possible such a science. For many years, experimental evidence has accumulated about the effect of the observer's attitude on perception, the influence of culture on perception, and the roles of past experience and of sensory organization in perception. All these experiments, however revealing, leave out of account the simple question of the relation of the stimulus to perception. Until this question is settled the other evidence will be hard to evaluate.

Several recent currents of psychological thought have influenced the writing of this book: the ideas of Gestalt psychology, of American functionalism, and of what might be called dimensionalism. The twentieth century scientists to whom I am most in debt are Kurt Koffka, Leonard T. Troland, and Edwin G. Boring. The hypotheses I have adopted were precipitated by research in the field of military aviation, carried out during the war.

Every book is a collaboration of its writer with others. The hardest collateral labor that went into these pages was performed by Eleanor J. Gibson, my wife, whose scientific conscience is stricter than my own and to whom the reader ought to be very grateful. This book is for her, with thanks and affection. The text has also been combed by Leonard Carmichael, editor of psychological books for Houghton Mifflin and one of my earliest teachers, with so much insight and erudition that I can never repay him. At an early stage of the manuscript it was carefully read by S. Rains Wallace who made the kind of detailed and penetrating comments that only a genuine friend is capable of. I am

grateful likewise to a number of other friends and colleagues who have gone over large or small parts of the manuscript: Robert B. MacLeod, John Volkmann, T. A. Ryan, Edwin G. Boring, Wolfgang Kohler, Hans Wallach, Annalies A. Rose, Fritz Heider, R. T. Sollenberger, H. E. Israel, Mervin Jules, and Oliver W. Larkin. Thanks are especially due to Frederick N. Dibble, who has worked with me for several years in testing experimentally some of the hypotheses to be described and who helped formulate them. Finally, my debt must be acknowledged to Robert M. Gagne and the co-workers of my wartime research unit who performed the feat of behaving like scientists in a military community.

This is a book intended to interest anyone who has ever acquired a sense of the awe-inspiring intricacy of vision. No realm of inquiry offers more strange and wonderful discoveries.

JAMES J. GIBSON

HOW TO LOOK AT THE ILLUSTRATIONS IN THIS BOOK

Many of the pictures are intended to give an impression of depth or distance. This effect will generally be clearer and more vivid if you will close one eye, look at the center of the picture, and hold it somewhat closer than you are accustomed to. You may have to wait a few seconds for the full effect to occur. This rule applies to the photographs and drawings but not to the cross-sectional diagrams.

Contents

Editor's Introduction
v

Preface
vii

1
Why Do Things Look as They Do?
1

2
Theories of Perception
12

3
The Visual Field and the Visual World
26

4
The Formation of Retinal Images
44

5
A Psychophysical Theory of Perception
59

6
The Stimulus Variables for Visual Depth and Distance – Momentary Stimulation
77

7

The Stimulus Variables for Visual Depth and Distance – The Active Observer
117

8

The Problem of the Stable and Boundless Visual World
145

9

The Constant Sizes and Shapes of Things
163

10

Geometrical Space and Form
188

11

Meaning
197

12

Learning
214

13

Spatial Perception and Spatial Behavior
223

References
231

Index
239

The Perception of the Visual World

For this end the visive sense seems to have been bestowed on animals, to wit, that by the perception of visible ideas .. they may be able to forsee the damage or benefit which is like to ensue upon the application of their own bodies to this or that body which is at a distance; which foresight how necessary it is to the preservation of an animal, everyone's experience can inform him."

George Berkeley, Bishop of Cloyne, *An Essay Towards a New Theory of Vision*, 1709

1

Why Do Things Look as They Do?

The Theoretical Approach . . . . The Initial Hypotheses of a "Ground Theory" of Space Perception . . . . Sensation and Perception

This is a book about how we see. There are, as everybody knows, a number of conditions which have to be fulfilled before anyone can see: there must be light to see by; the eyes must be open; the eyes must focus and point properly; the sensitive film at the rear of each eyeball must react to light; the optic nerves must transmit impulses to the brain. Just so long as one of these conditions is not fulfilled, the seeing person is blind. People who have not thought about the problem find it difficult to realize that sight depends on such a complicated chain of circumstances, for seeing does not "feel like" that. It "feels as if" things were simply there. Nevertheless, such is the case.

Normal sight is an astonishingly good guide for getting about and doing things. A seeing man can walk without colliding with obstacles. He can use tools as fine as a jeweler's needle and as large as a steamshovel. He can read print, or look at pictures, or identify faces. He can discriminate objects which resemble one another even at a considerable distance. All these a blind person cannot do. A seeing man can climb a cliff, drive an automobile, fly an airplane, or even leap through the air at the top of a circus tent. He can match colors and draw representations of things. He can design and build machines, and he can change the appearance of the environment almost to suit himself. Or, as another possibility, he can simply sit and look at the scenery.

This last, in a way, is the most astonishing performance of all, for the view of a room or a countryside which one gets when he simply looks at it in a receptive mood has great scope and, at the same time, the most minute detail. The number of items that can be described in such a view is enormous. What is most astonishing is that it is in every detail a nervous process. The panorama is utterly and entirely a performance of the living organism. If the brain is injured in a particular way, a partial blindness results, and the kind of blindness is related to the particular injury. If the optic nerve or the retina of the eye is damaged in some part, sight suffers damage in a precisely corresponding way. The simplest experiment is to close one's eyes and reflect on the fact that the visual panorama vanishes.

The problem of visual perception has a long history. For hundreds of years men have felt the need for some explanation of why things are seen. Among the many puzzles to which the problem leads, perhaps the oldest and most general of all is this: how can one account for the richness of sight considering the poverty of the image within the eye? Vision depends on this retinal picture. But what an inadequate thing the image seems to be when compared with the result! The visible scene has depth, distance, and solidity; the image is flat. How can vision depend on the pictures in the eyes and yet produce a scene which extends to the horizon? The physical environment has three dimensions; it is projected by light on a sensitive surface of two dimensions; it is perceived nevertheless in three dimensions. How can the lost third dimension be restored in perception?

This is the problem of how we perceive space. The question is put in terms of the geometrical dimensions of height, width, and depth. In a sense, this book is about space perception. The plan of these chapters, however, is to end with the problem of abstract space rather than to begin with it. The space to be considered first is not a void with three lines intersecting at right angles but the space of rooms, streets, and regions, and the space of men who walk, drive, or fly an airplane. The puzzle of the third dimension can be much better understood if we first examine the scenes we actually see and the ones which are of practical importance for human behavior.

The problem of how we perceive space implies a good many other problems, and this book will also be concerned with them. For example, how do we see the form or shape of a thing? This question is not at all easy to answer. The search for an answer takes one so far afield that it provided, some thirty years ago, the basis for a new approach to psychology — the theory of Gestalt psychology. For another example, how do we see the motion of a thing? Still more fundamental, how do we see a thing — the mere object as distinguished from its general background? Probably this last reduces to two questions: first, how can we see an outline as separated from its background, and second, how do we see a solid surface? There are many other such questions, not easy to formulate scientifically because they are all more or less interrelated. Why do things have location, that is, how can we see where they lie? How do we see fine detail, and what are the limits of this acuity? Why do things look right side up? Why does the world always appear level even when we lie down?

There are also a whole set of practical problems which depend on the solution of the theoretical problems. How can men see to fly airplanes and drive automobiles? What does the artist see when he paints a picture? Why is a photograph so astonishingly like the scene at which the camera was pointed? How far must the movies inevitably fall short of natural seeing? Can vision be improved by training? What is visual education and how may it be used to advantage in school and college? These practical problems will be touched upon, but it is fruitless to look for their solution without first laying the groundwork of a scientific theory.

FIGURE 1. A Scene for Analysis

There is another problem with which we shall be indirectly concerned since upon our knowledge of it everything else depends – how do we see light and how can we perceive color? Light and color are, in a way, the raw material of vision. The perception of an object in space would be impossible if we were not sensitive to the light reflected from the object and to the brightness and hue of this light. There is a vast accumulation of evidence about brightness and hue. Nevertheless, this evidence is not enough to provide answers to the other questions, inasmuch as the seeing of an object is an ability quite different from the seeing of abstract color. Seldom or never does one see a color as such. This is primarily a book about objects. In the accepted terminology, our problem is that of perception, not of sensation.

All these problems can really be summed up in a single general question: How do we get the experience of a concrete visual world? The visual world can be described in many ways, but its most fundamental properties seem to be these: it is extended in distance and modelled in depth; it is upright, stable, and without boundaries; it is colored, shadowed, illuminated, and textured; it is composed of surfaces, edges, shapes, and interspaces; finally, and most important of all, it is filled with things which have meaning. If we could account for the perception of these properties of the visual world, we should at least be

well on the way to explaining the whole panorama of visual experience.

Examine the scene reproduced in Figure 1. It represents an uninteresting dry river bed of the sort common in the Southwest, bordered by a tall growth of bushes, with two men standing in the foreground. It will serve as an example of what is meant by a concrete visual world. Let us consider it abstractly without regard to the familiar meanings which can be applied to it. The bottom part of the picture (the "ground") looks solid whereas the upper part (the "sky") does not. The solid part looks generally level and this surface appears to extend to a great distance. Actually it is a compound of visual surfaces (ground, bushes, men) separated by contours. One of the most prominent contours is the horizon. The various surfaces have the quality of texture, sometimes fine and sometimes coarse, although the sky does not have this quality. Some of them have closed contours or shapes and they are located with reference to the ground. Parts of the ground appear to be illuminated and other parts shadowed. Most abstractly of all, the whole scene is composed of a pattern of light and dark, that is, an enormously complicated mosaic of grays, blacks, and whites, with variations (not represented in the photograph) of yellow and brown, dusty green, and vivid blue.

Granting that the picture, although it fails in some ways to look like the actual scene, is quite similar to it, what makes it such a good substitute? Analysing it, the properties which give it the appearance of concrete visual reality seem to be just those listed; surface-quality, solidity, horizontal character, texture, distance, contour, shape, adjacent location, illumination, and shading. The list is tentative and incomplete but it illustrates the kind of problems with which the contemporary study of space perception is concerned. There are, of course, other properties of an actual scene which do not show up in a photograph but are nevertheless important for space perception. Chief of these are the stereoscopic impressions dependent on vision with two eyes, and the vivid qualities of depth which occur when the head is moved. The contributions of these impressions to the perception of space have been known for a long time but they are not, as is sometimes believed, the exclusive basis of our perception of a three-dimensional world.

In contrast with the substantial world represented in Figure 1, let us imagine the perception obtained by an observer in the nearest possible approach to empty visual space. Assume that his environment consisted wholly of atmosphere without any opaque objects. He could live for some time at the center of such a sphere of air although, without the gravity of the earth, he could not maintain a posture or change his location. Suppose that this environment is illuminated by external sources but that his world of air is so large as to diffuse the light evenly, as our familiar sky tends to diffuse the light of the sun. If he opens his eyes he can see, and the question is *what* will he see?

There are experiments which yield a reasonably sure answer to the question. The light which stimulates the retinas of his eyes will be homogeneous (67), that is, the same at all points. He can turn his eyes in any direction but they will

not focus or converge and he cannot fixate or look at anything for there is nothing to fixate on. He will see luminosity or color but it is the kind of color which Katz has named film-color as distinguished from surface-color (61). It is unlocalized in the third dimension; its distance is indeterminate. The sea of light around him might vary from bright to dark and from one hue to another but the quality of color would be neither that of a surface on the one hand nor would it be extended in depth on the other. It is neither near nor far. The space he sees is certainly not two-dimensional in the sense of being flat but it is also not three-dimensional in the sense of being deep. Assuming the atmosphere to be cloudless, without dust or fog-particles, it has no texture, no arrangement, no contours, no shapes, no solidity, and no horizontal or vertical axes. The observer might as well be in absolute darkness so far as he can *see* anything. The results of this hypothetical experiment suggest, then, that what an observer would perceive in a space of air would not be space but the nearest thing to no perception at all. The suggestion is that *visual* space, unlike abstract geometrical space, is perceived only by virtue of what fills it.[1]

The hypothetical man at the center of a sphere of pure air is even further instructive. Although he would presumably have no impressions of far or near, and no sense of his surroundings as being either flat like a picture or modeled like a sculpture, that is not all he would lack. Almost certainly he would have no impression of up and down. Since the pull of gravity on his body and the resistance of his legs against the substratum are wholly lacking, he would have no equilibrium and could not maintain a posture. He would feel as if he were floating. Although he could look toward or away from his feet and could see his right hand and his left hand, these acts would probably have lost much of their normal meaning of up or down, right or left, and he would experience a profound and complete disorientation. He could thrash about but could not change his position in phenomenal space, and in fact he would have no position in a visible environment. His sense of the vertical and horizontal directions (ordinarily given by the stimuli for his postural reflexes and by the main lines of his retinal images of the horizon and of trees, tables, and rooms) would be wholly lacking. Since he would have no axes of reference for his space it is questionable

---

[1] Besides the experiments of Katz on film-color (61) there are also the results of Metzger on homogeneous light stimulation over the total visual field, the *Ganzfeld* (81). Taken together, they suggest the above results for a hypothetical observer floating in air. Katz studied "aperture-colors" — the appearance of a hole in a surface behind which is another surface too distant to yield the perception of a surface. Metzger studied the appearance of a uniform surface which filled the whole field of view. He made it homogeneous by reducing the illumination until nothing was seen but film or fog. The conclusion of these experiments was that a visual surface depends on the perception of "microstructure," that is, the minute inhomogeneities of reflected light which give it a texture or grain. These results are to be contrasted with the theory of Bühler that space might be given by a hypothetical "air-light" — a direct seeing of the atmosphere dependent on molecular particles (16). This supposition has never received confirmation, and Bühler himself abandoned it.

whether he could be said to perceive even an abstract geometrical space.[2]

The Theoretical Approach

At the beginning of World War II there was a sudden need to understand the perception of depth and distance as they applied to aviation. The critical task of estimating distance from the ground when a flier is landing an airplane was particularly important. Research was begun and studies of space-perception multiplied, based on what psychologists already knew about it from the great experiments of the 19th century. A list of the *clues* or *cues* for the perception of the distance of an object had resulted from these experiments and this list had gained acceptance. The cues were classified as binocular or monocular according to whether they depended on the use of two eyes or one eye.

The typical means of experimenting was to employ a stereoscope, or a depth-perception apparatus, or a dark room in which points of light or similar isolated stimuli appeared. The points, lines, or objects whose distance was to be judged usually appeared against a homogeneous background. The fact was, however, that these experiments failed to clarify the practical problems of how a man lands an airplane. Many tests were devised but none of them predicted a prospective flier's success or failure at this task. Many suggestions for training were made but none of them made the performance substantially easier. Toward the end of the war it began to be evident to psychologists working on problems of aviation that the usual approach to the problem of depth-perception was incorrect. Experiments needed to be performed *outdoors*. The stimuli to be judged ought to be those of a *natural environment*. A hypothesis with a vast set of new implications (new at least to the writer) began to assert itself – *the possibility that there is literally no such thing as a perception of space without the perception of a continuous background surface*. This hypothesis might be called a "ground theory" to distinguish it from the "air theory" which seemed to underlie the earlier research. A few experiments were performed by the writer and his collaborators before the war ended using outdoor situations, photographs, and motion pictures, in which a level ground was always visible (39).

The basic idea is that visual space should be conceived not as an object or an array of objects in air but as a continuous surface or an array of adjoining surfaces. The spatial character of the visual world is given not by the objects in it but by the background of the objects. It is exemplified by the fact that the airplane pilot's space, paradoxical as it may seem, is determined by the ground and the horizon, not by the air through which he flies. This conception leads to a radical reformulation of the

[2] There have been almost no experiments which study the effect of eliminating completely the force of gravity on the perceptions of a human observer. They are needed, since the rocket-passenger outside the earth's gravitational field will meet just this condition, and it is no longer fantastic to be concerned with the problem. A man falling freely toward the earth satisfies the condition, but volunteers for such an experiment are rare. In any event, there has been no instance in which a man without postural stimulation has also been presented with absolutely homogeneous visual stimulation. A free-falling parachutist can always see the horizon. The description above is therefore speculative, although consistent with such evidence as exists (42).

FIGURE 2. The Look of the World from the Air

stimuli or cues for depth and distance. Instead of investigating the *differences* in stimulation between two objects, the experimenter is led to investigate the *variations* in stimulation corresponding to a continuous background. This shift of emphasis has a great many implications, and these will be explored in the ensuing chapters.

This "ground theory" of visual space is the organizing scheme of the present book. The classical problems and facts of perception will be considered, but not in the order in which they were discovered and not under the usual headings. If our scientific conception of space perception was inapplicable to aviation, what we need is a new theory rather than new evidence. The "air theory" of visual space is actually inconsistent with a good many experimental results. But, as Conant has remarked of the history of science, "a theory is only overthrown by a better theory, never merely by contradictory facts" (24).

The Initial Hypotheses of a "Ground Theory" of Space Perception

What are the main principles of such a theory? Since they determine the plan of the book, it might be well to summarize them at the outset. Their explanation and factual status will be given in later chapters.

1. *The elementary impressions of a visual world are those of surface and edge.* These are the fundamental sensations of space, the stimuli for which need to be discovered. They are elementary, however, not in the sense that atoms or units are supposed to be elementary, but only in the sense in which a variable or quality of something is essential to understanding it. These candidates for the status of sensations are very different from the elementary impressions of *location* assumed by the traditional approach to space perception. Such elements were arranged from right to left, from up to down, and from near to far according to the abstract coordinates or dimensions of geometry.

The impression of a continuous surface may account for visual space conceived as a background. The impression of an edge may account for an outline or figure against the background—the "figure-ground phenomenon"—and together with the surface enclosed may account for the perception of an object.

2. *There is always some variable in stimulation (however difficult it may be to discover and isolate) which corresponds to a property of the spatial world.* This hypothesis says that even complex perceptual qualities must have stimuli. It is an extension of the principle of psychophysical correspondence to visual perception—the principle which has served so well in the study of sensation.

This rule suggests that a "stimulus" can be found for the impression of a surface. Probably it is a textured retinal image. A stimulus ought to be discoverable also for the quality of distance or depth over a continuous surface. Perhaps this is a gradual change along an axis of the retinal image, an increase or decrease, for instance, in the density of the texture of the image. Likewise, a stimulus ought to be discoverable for an edge or contour and for the impression of depth at a contour. Perhaps this is a jump or discontinuity in a gradient of the retinal image.

The policy of searching for a stimulus variable with which some quality of experience may prove to be in correspondence is the policy which underlies psychophysical methods in psychology (40). It is the first step in the *explanation* of experience. Some would argue that there is no real explanation of perception until the physiological mechanisms have been discovered, but this is a matter of preference. There are laws relating perception to physical stimulation as well as laws relating it to physiological processes. Explanation is a matter of lawfulness, although there are different levels of explanation. The level to be aimed for in the present book is a psychophysical theory, not a physiological theory.

3. *The stimulus-variable within the retinal image to which a property of visual space corresponds need be only a correlate of that property, not a copy of it.* The qualities of solidity and depth, for instance, do not have any replica in the two-dimensional retinal image but they may very well prove to have correlates there. An assumption will be borrowed from geometry which states that when a three-dimensional physical world is projected optically, the slant and shape of its surfaces undergo a mathematical transformation in the projection but that they do not on this account vanish or disappear.

There is a naive theory of perception to the effect that the outer world somehow gets into the eye. Almost the first principle the beginning student learns is that nothing gets into the eye but light. This third assumption can be sharpened by saying that, in a special sense, the outer world *does* get into the eye. It implies that at least the surfaces, slopes, and edges of the world have correlates in the retinal image specifically related to their objective counterparts by a lawful transformation. If this is correct, the problem of the restoration of the lost third dimension in perception is a false problem.

There is another naive theory of the visual process to the effect that a retinal *picture* is transmitted to the brain by the optic nerve. In a more sophisticated form it is tempting even to the visual scientists, although it leads to difficulties. According to the first part of the hypothesis, however, there is no need for a picture-theory of psychophysical correspondence since perception may be a correlate, not a copy, of the image. If the image is neither a replica of the world nor a picture for the perception but a complex of variations, it may prove easier to trace its specific correspondence to both.

4. *The inhomogeneities of the retinal image can be analysed by the methods of number theory and modern geometry into a set of variables analogous to the variables of physical energy.* This says, in effect, that the order or pattern of the retinal image can be considered a stimulus. It is the most debatable and least developed of the hypotheses being summarized. The problem of the abstract nature of a differentiated visual image is variously named. How do we perceive form, pattern, configuration, order? Why is vision organized, structured, detailed, precise? The greatest achievement in the theoretical struggle with this problem has been reached by Koffka in his *Principles of Gestalt Psychology* (67). An attempt will he made in Chapter 5, however, to follow a different theoretical path and to suggest that a so-called pattern of stimuli is itself a stimulus. The term pattern is vague and unanalysed. The mathematical conception of order, as exemplified by the number-series, is more exact. An effort will be made to show that a few simple variables of pattern – texture, contour, and density of texture – are definable as variations of *adjacent order* in the retinal image.

The experimental study of what was called *inhomogeneity* or *differentiation* of the retinal image has mostly been carried out under a different name and with quite a different intention. It has been called the study of visual *acuity*. A great many experiments have been carried out on acuity, but in them only a few kinds of inhomogeneity have been studied: the separateness of two adjacent spots or two parallel bars, the gap in a broken ring, the impression of a single line, a grating of dark and light bands, and the familiar letters of the acuity test. It can be argued that these are artificial rather than natural types of stimulation. An attempt to connect acuity, or "resolving power," with the more general idea of a differentiated, patterned, or textured image will be made in Chapter 6.

5. *The problem of how we perceive the visual world can be divided into two problems to be considered separately, first, the perception of the substantial or spatial world and, second, the perception of the world of useful and significant things to which we ordinarily attend.* The first is the world of colors, textures, surfaces, edges, slopes, shapes, and interspaces. The second is the more familiar world with which we are usually concerned, a world of objects, places, people, signals, and written symbols. The latter shifts from time to time depending on what we are doing at the moment, whereas the former remains a more or less constant background for our experience, and a sort of support for maintaining posture and for moving about. The world of significant things is too complex to be attended to all at once, and our perception of it is *selective.* Certain features stand out prominently, others are neglected. It is sometimes said that our perception is distorted and falsified by this fact. This kind of perception can be called *schematic,* whereas the first kind can be called *literal.*

Before one can fully understand schematic perception one must understand literal perception since it provides the fundamental repertory of impressions for all experience. This is primarily a book about literal perception, therefore, and only secondarily a treatment of schematic perception. The discussion of the meaningful visual world is deferred until the end and does not pretend to be complete. Although it is true that everyday perception tends to be selective, creative, fleeting, inexact, generalized, stereotyped, and to have all the other defects so commonly ascribed to it, the best hope of understanding these defects is first to examine the respects in which perception is adequate and exact.

The method of investigating adequate impressions of a substantial or spatial world is the psychophysical experiment. This is, essentially, a procedure of isolating and then systematically varying a feature of the physical stimulus for an observer who makes judgments of "more" or "less," or otherwise shows that he discriminates the variation. Although this method has been very little used in the study of perception (as distinguished traditionally from sensation) there is every reason to think that it can be applied (40). The attempt to do so can be called a psychophysical approach to the study of perception. It involves searching for some feature of the physical stimulus with which to set up an experiment.

The method usually employed in the past for the study of perception is fundamentally different from that of the psychophysical experiment. It was a policy of searching for *discrepancies* rather than *correlations* between the stimulus and the perception. Assuming that sensation is dependent on stimulation but that perception is not, the policy of the experimenter has been to isolate and study these discrepancies. A favorite device for enhancing them has been the tachistoscope which presents an image to the observer for only a fraction of a second. The method is one of "impoverishing" the stimulus, or reducing the optimal conditions for literal perception which characterize the psychophysical experiment. Brief exposure, low illumination, many stimuli in succession,

and the use of indefinite, ambiguous, or equivocal stimuli have all been employed in this kind of research. The resulting body of facts is very large, but only a small part of it can be considered in the present book.

Sensation and Perception

The approach outlined above is not consistent with the usual meanings of the terms *sensation* and *perception.* Obviously these terms will have to be either discarded or redefined. The second hypothesis implies that perception, at least of the type called "literal," is primarily dependent on stimulation rather than on meaning or mental elaboration. This hypothesis contradicts the traditional conception that, whereas sensation depends only on immediate stimulation, perception depends also on past stimulation, or memory.

The rejection of this distinction is neither so novel nor so radical as it may sound. Although a generation ago it was still possible to suppose that sensations and perceptions were essentially different, the discoveries of Gestalt psychology have overthrown the logical basis for the distinction. The seemingly vast difference between a sensation and a phenomenal object has been slowly vanishing in recent years. Instead of the doctrine that perceptions were built up out of elementary sensations, a more defensible idea has been gaining ground: that of variables or dimensions of all experience, perceptual as well as sensory. Such variables as the texture and slant of a surface are, no doubt, a far cry from the variables of hue, brightness, and saturation of color. But, if it is no longer to be assumed that the mind constructs the surface out of bits of color, the qualities of a surface need to be analysed as the qualities of color were analysed many years ago, and the first problem is to search for variables of the retinal image with which these qualities might prove to be in correspondence.

A substitute for the distinction between sensation and perception will be offered in Chapter 3, a substitute intended to retain what is verifiable in the classical distinction and eliminate what has been theoretically misleading. We can attend either to color-impressions or to object-impressions, generally speaking. Introspection of the first sort yields an experience of the visual *field.* Introspection of the second sort, called "phenomenological," yields an experience of the visual *world.* Both these kinds of experience must be accounted for if we are to understand vision, but the latter is the subject of this book. How can we see the world? Why do things look as they do?

2

Theories of Perception

The Distinction between Sensation and Perception . . . . Nativism and Empiricism . . . . Extensity and Location . . . . Form or Shape in Two Dimensions . . . . Depth and Distance . . . . The Theory of Cues . . . . Gestalt Theory . . . . The Fact of Perceptual Constancy . . . . Summary

The traditional explanation of vision is that perceiving things depends on first having sensations. Sensations are supposed to be the raw material of human experience and perceptions the manufactured product. Sensations are only colors, sounds, touches, odors, and tastes; objects and space depend upon perception. A certain hue, a feeling of warmth, and a smell of smoke are not things in themselves. Only when they are combined in a perception do they make us experience a fire. The eyes furnish us with an array of colors, the ears with a flow of sounds. That is all they can do, and the rest of experience is a matter of combining, ordering, and uniting the sensations into things and events. The play of light within the eye can give us color but not things. Things are a product of a mental capacity called perception.

This explanation has so much age and respectability that there is a temptation to forget that it is only a theory. As a matter of fact it is not consistent with a great deal of accumulated evidence in psychology. It would be worth while to consider this evidence, and we may start by inquiring how the distinction between sensation and perception arose.

The Distinction between Sensation and Perception

Around the latter part of the seventeenth century, the imagination of men began to be stirred by the theory that all human knowledge comes through the senses and from no other source. In short, we *learn* our ideas instead of discovering them implanted in our minds by God. It follows, for example, that every man can acquire his ideas for himself, and that he himself is the best judge of their truth. The doctrine was given a special impetus by John Locke in 1690 in *An Essay Concerning Human Understanding.* The mind at birth is a blank page – a *tabula rasa* – on which experience writes its record. If knowledge could exist in mind only by way of sense, it was obvious that the sensory capacities of man needed to be carefully investigated. Since vision was the principal sense, scholars began to concern themselves with the optics of seeing, and to note what they

themselves could see under controlled conditions. But here they encountered a difficulty. The visual sense was simply not adequate to account for all visual knowledge, especially of three dimensional space. Either, then, some knowledge of the world does *not* come through the senses, or the visual sense must be supplemented in some way by the mind. There must exist a special mental process over and above the visual sensations: a process which in some way constructs the world out of the "raw data" presented to the mind. Such a process might be one of association and inference; the alternative would be a kind of intuitive understanding of the data of sense which would imply a retreat toward the dogma of innate ideas. This argument was a rationale for the theory of perception which still underlies our thinking. The nature of this special mental process has puzzled some of the best thinkers and scientists in western civilization for two hundred years.

The obvious puzzle in giving any exact account of perception was the visual third dimension. A very knotty question arose: How can we apprehend the "real" world as distinct from the world of sense, or, in other words, the world which appeared to be "external" as distinct from the play of light within the eye? Various criteria of the visual reality of objects were described, such as maintaining their position despite eye movements and conforming with impressions of touch, but the distance and depth of objects were their most obvious features, and these it seemed impossible to explain. The 18th century scholars understood that the eye can obtain an image of an object but cannot sense the external object at a distance – the object "itself." The paradox was that the latter is nevertheless apprehended. There arose among philosophers a dispute, now centuries old, over whether and how we can *believe in* an external world. If objects with solidity and distance were creations or constructions of the mind, then it could be inferred, for example, that they were mental objects. Physical objects either did not exist or, if they did, were unknowable. If they were nevertheless known, the explanation must be supernatural. A vast amount of intellectual effort and ingenuity has been devoted to this type of controversy or to some means of escaping from the dilemma on which it was founded. And the dilemma itself appears to rest, in part at least, on the conviction that such properties as distance and solidity cannot be sensed and that the apprehension of them poses a unique and special problem. If a sensory basis for such properties could be discovered in the retinal image, however, the dilemma might collapse and the whole intellectual superstructure would fall with it.

The accepted view of perception is still that the percept is never completely determined by the physical stimulus. Instead, the percept is something essentially subjective in that it depends on some contribution made by the observer himself. Perception goes beyond the stimuli and is superposed on sensations. The sensations are basic and, being parts of our organic equipment, tend to be the same for all. Perceptions, however, are secondary and, depending on the peculiarities and past experience of the individual, may vary from one observer to another.

When this doctrine of perception is applied to such abilities as the apprehension of meaning—to the understanding of language for instance—it works very well and accounts for most of the experimental facts which psychologists have accumulated. Meanings do depend upon the past history of the individual. But when it is applied to the apprehension of material objects and of the spatial environment, it is less satisfactory. *For the visual worlds of different observers are more alike than they ought to be if the doctrine were the complete truth.* The evidence accumulates that men, and moreover even animals, appear to react to the spatial environment with an accuracy and precision too great for any known theory of space perception to be able to explain. The fundamental modern difficulty is this. If the solid visual world is a contribution of the mind, if the mind constructs the world for itself, where do the data for this construction come from, and why does it agree so well with the environment in which we actually move and get about? If space perception is a subjective process then why are we so seldom actually misled by illusory perceptions? Why are the optical illusions of the textbooks actually the exception rather than the rule?

Nativism and Empiricism

The history of past attempts to account for the process of space perception is protracted, involved, and difficult. Even at the risk of oversimplifying, however, its main issues need to be sketched if we are to clear the way for any novel approach to the problem. It is the history of a controversy. On the one side, a group of British philosophers in the eighteenth century and experimental psychologists in the nineteenth strove to explain perception with as little appeal as possible to intuition or innate ideas. Such theories they considered mystical and not consistent with a scientific psychology. Visual space, they were convinced, must be somehow learned. On the other side, many philosophers and some experimental psychologists could find no satisfactory way of understanding how this could occur. At least some features of visual space, they argued, are so immediate, simple, and clear in our consciousness that they must be either intuitions which are fundamental to "mind itself" or else must be innate features of the sensations themselves. The speculations and debates of these two groups make up what Boring calls the "long and barren controversy" over nativism and empiricism (12). [1]

In order to understand what was meant by *space* in this controversy, it is necessary to remember the scientific conception of the world which began to be current at the beginning of the eighteenth century. The discovery of gravity by Sir Isaac Newton led him to conceive a physical universe so logical and simple that it became the wonder of the age (85). The facts of astronomy and physics were united in it; these and many other facts became predictable from a few simple laws. This physical universe consisted of three things only, space, time, and matter, and

[1] The writer follows the usage of Boring in using the term empiricism as the alternative of both nativism and rationalism. "Empirism" is the term employed by Gestalt psychologists.

from these realities everything else could be deduced. Events were reducible to matter varying in space with time, i.e. in motion, and could therefore be analysed in terms of grams, centimeters, and seconds. The space of this universe, it may be noted, was empty Euclidean space, defined by the three dimensions of the Cartesian coordinates. Inevitably, then, the problem of how we can observe the world was formulated as the problem of how we can apprehend the Newtonian universe, and by space perception the eighteenth century philosophers and the nineteenth century psychologists meant *geometrical* space perception.

This presupposition influenced the psychologists' analysis of the problem and dictated the terms in which theories could be propounded. For both nativists and empiricists, perceived space seemed to divide up naturally into certain geometric categories. First there was *extensity* in two dimensions: the bare characteristic of space as being spread out. This corresponded to the plane of the vertical and horizontal axis in geometry. Then there was the aspect of *location* in two dimensions, or the localization of points in the visual field. This corresponded to the x and y coordinates of geometry. Next there was the aspect of *shape* or *form* in the visual field. This corresponded to the abstract forms of Greek geometry. Finally there was the aspect of *depth* or *distance*, the third dimension of space, and this corresponded to the third dimension of geometry. Extensity, location, shape, and distance: these were the primary constituents of visual space. They do not, it may be noted, constitute anything very similar to what has been called, in this book, the visual world.

Both nativists and empiricists agreed that the visual sensations were innate. Sensations were the data of, or what was "given" to, the mind. They disagreed over whether perception was a matter of learning or of intuition. But they also disagreed from the very beginning over what was sensed and what was perceived. The simplest and most logical doctrine was to suppose that only color could be sensed and that *all* the constituents of space were perceived, including extensity. This implied that a color sensation could only be a spot or point of color, and that an *area* of color was the sum of these elementary sensations. As thus conceived, the sensations corresponded with the focused points of light in terms of which optics had analysed the retinal image. This was the theory of Wundt, the most consistent sensationalist. Another doctrine was to suppose that extensity was sensed (or was an "attribute" of sensation) but that the *location* of points in the extended field was not sensed and therefore had to be learned by experience. As a third possibility, not only unshaped areas but also shaped areas, or forms, might be considered to be data of sense. William James for example, although he did not actually assert that a form was a sensation, did believe that a visual line was a simple datum rather than a row of point sensations. As a last possibility, it might have been assumed that *all* constituents of space were sensed. But actually no one ever supposed that depth and distance were simple sensations, and the visual third dimension was and remained

a phenomenon which only perception could explain. Keeping in mind these variations in conceiving the sensory elements with which perception had to work, let us examine the efforts of the empiricists to explain the constituents of space without resort to innate ideas.

Extensity and Location.

Although it was logically possible to assume that a visual field filled with pure color, such as the blue sky, was a mosaic of spaceless points which looked continuous because they had been associated together in past experience, this seemed highly improbable. A more plausible theory was that a color sensation was extended by its very nature: that color simply came that way. The blue sky, then, was a simple sensation and no problem to the empiricist. The commoner kind of visual field filled with patches of different color, however, was a different matter. This was something like a space with objects in it and to this the special process of perception might apply. Such a field possessed order, arrangement, or pattern as we would say today. But to the early psychologists it seemed that the way to start analysing it was not in terms of *order* but in terms of location. How did the spots of different color get their position or place in the extended field of view? If the position of all points in the field could be perceived, they reasoned, everything in the field could be perceived.

The space of the physicist was a space of points whose position could be defined by the Cartesian coordinates. To the psychologists, therefore, it was clearly necessary to develop a theory of "local signs" in order to account for a visual field. A local sign was the unique accompaniment of every point in the field, determining its position in the up-down and right-left dimensions. Since every point could be separately localized, or pointed to by the observer, each must have its own locality-characteristic distinguishing it from every other point. The question which divided the empiricist and the nativist was whether this differentiating characteristic became associated with its appropriate retinal point through experience or had been intrinsically connected with that point from birth onward. The variations of opinion on this question need not be described. A possible explanation for the learning of these locality-signs, in general terms, was that each point on the retina got associated with the movement of the eye just necessary to bring its stimulus to the fovea. It was practice in fixating points (or locating them with the eyes) which made their location possible when the eyes were motionless.

Form or Shape in Two Dimensions.

To the empiricist psychologists the perception of solid objects required two stages of explanation: first a theory of plane geometrical shapes, and second a theory of their three-dimensional character. Since the retinal image was two-dimensional, this seemed the most reasonable approach, and it was reinforced by the psychologists' tendency to see things pictorially when they analysed their own perceptions introspectively. The term *shape* thus came to mean primarily *projected* shape or, more specifically, the *projected shape as the object is commonly*

viewed. Form in this sense of the term can be experimented with, since it can be conveniently represented on paper. A box for example can be drawn as a square. Black lines on a white background are not, in actual fact, much like the edges and contours of objects in a visual field, but to civilized vision they are good equivalents. The study of drawn shapes, accordingly, has been pursued for centuries and it constituted an obvious problem for the early psychologists.

If sensations were points of color, then a shape must be a mosaic of such point-sensations associated with one another during the course of past experiences with that particular shape. A line, for instance, was a row of contiguous black spots. An extended shape was an array of colored points. But, as we have already noted, only the most radical of empiricists were explicit in believing that the sensations of vision were points. The more common opinion was that color possessed extensity as an innate attribute. The formed or shaped character of a "piece" of extensity might, however, be learned even if the extensity itself were not, and this is what empirically-minded psychologists have tended to believe up to the present day. But no one has yet demonstrated precisely how such learning could occur, or has even explained just why, if extensity is an unlearned feature of experience, form should be a learned one.

The experimental evidence on whether or not we have to learn to perceive forms has proved, over the years, to be not very conclusive. One can study the behavior of infants systematically and make inferences about their first visual perceptions. But the evidence obtained cannot be interpreted as proof that at the outset they either do or do not see shapes. The implication of the reactions which babies first make to faces and other visual objects is that they see them as forms, but of a sort incomprehensible to any adult: forms which can only be called *indeterminable* or *undifferentiated* from one another. These terms do not mean that vision in the infant is what adults would call blurred, or that the contours and details of things appear as they do in an out-of-focus photograph. They can only suggest, not describe, what the perceptions of the infant are probably like. There is evidence, for instance, that the typical baby at 3 to 5 months can see human faces as clearly distinguished from other things but not as distinguished from one another (99). The development of perception seems to proceed from the seeing of gross differences to the seeing of fine differences. Whether this development is principally a matter of learning or principally the result of the natural growth of the optic nervous system is not now known. In any event the learning process, if it is that, is not like the learning of geometry which proceeds logically from points to lines to planes and thence to solids in a wholly different kind of sequence.

Figure 3 shows what a nine-months-old baby is supposed to see when his mother plays peekaboo with him. An ingenious attempt has been made to suggest how the visual field becomes progressively less determinate from the center to the periphery, when the viewer fixates an object of interest, a fact as true for adults as for babies. The photograph is increasingly blurred away from the center. The baby's

(By Life *photographer Herbert Gehr. Copyright* Time, *Inc.)*

FIGURE 3. An Attempt to Represent the Vision of an Infant

perception, however, may be indefinite without being optically out of focus; this the photograph fails to convey. The evidence will be discussed in Chapter 11, (p. 207).[2]

Toward the end of the nineteenth century a few psychologists began to emphasize the fact that a form may be transposed on the retina, as the observer scans the object he is observing, without its making any difference in the perception. Although the sensory elements differed, the form did not. Moreover, the form was the same whether the color it was made of

[2]For a description of the year-old infant's visual behavior, see A. Gesell, F. L. Ilg, and G. Bullis, *Vision: Its Development in Infant and Child* (Harper and Brothers, 1949).

consisted of white on black or black on white. The form, they reasoned, must therefore be independent of the anatomical retinal points of the image and also independent of the color stimulation of the points. It must in fact be a "form-quality" *(Gestaltqualität)* somewhat analogous to the color qualities of hue and brightness and therefore incapable of being analysed into sensations. Form, then, was something that fell into neither the category of perception nor sensation; it was irreducible and elementary like a simple sensation but, unlike a sensation, it had no comprehensible stimulus-equivalent in the retinal image. What *could* be the stimulus for a visual form? Either it must be a set of point-stimuli, and this was not easy to understand, or it was the *form* of these points, and this was a mere tautology. The dilemma was one which, as we shall see, the Gestalt theorists attempted to resolve.

Depth and Distance, The Theory of "Cues"

On the basis of the sensations of color, conceived either as points or in some vague way as formless and sizeless extents, the empiricists supposed that human beings somehow construct a three-dimensional world in perception, or, in the terms of the philosophers, that we *have knowledge* of a three-dimensional world. How could this occur? Specifically, what information could the eye transmit on which such perception or knowledge could be based?

Considering the problem as one of Cartesian geometry, it seemed obvious that a single eye could not yield *any* information about the third dimension since the latter consisted of the line of sight itself, i.e. a line represented on the retina as a single point. Any external point on the line of sight would be optically the same as any other point. There was nothing to indicate whether it was near or far, or even for that matter outside the eye. The data for perceiving the distance of a point must therefore be provided by the use of *two* eyes.

Since both eyes are always aimed at an objective fixation-point so that there is a clear image of it on the exact center of each retina, the distance might be known by a sort of triangulation. The eyes might operate as a surveyor does when he, in effect, aims two telescopes at a distant object from the two ends of a fixed base line, or as a gunner does when he operates a range-finder. The visual process in the brain would have to include a kind of automatic reasoning not unlike the computing mechanism of a range finder, which can solve problems in trigonometry automatically. Helmholtz called the process "unconscious inference."

The sensory data for this estimation of distance could only be the eye-muscle sensations which accompany the converging or diverging of the eyes according as near or far points are fixated; the muscle sensation, then, was a "criterion" or "cue"[3] for the estimate. This idea can be credited to Bishop Berkeley who based his "new theory of vision" on it in 1709.

---

[3]In American psychology, as Boring has pointed out (48), the words "cue" and "clue" have both been used to mean a kind of sense-fact on which to base perception or behavior. "Clue" implies reasoning whereas "cue" implies the touching off of some response, but their meaning has never been clearly distinguished.

He added to it the idea that sensations of accommodation (the adapting of the lens which brings the point of fixation to a focus) might be supplementary cues for distance. The theory has had a long scientific history. It is not however, as later computations proved, adequate to account for estimates of distance as far away as we can actually judge distance. The far-accommodation of the lens approaches a maximum limit at around fifteen feet and the convergence of the eyes approaches a zero limit at about fifty feet. For lack of a better theory, however, the cues of convergence and accommodation continued to be, and still are, given as a partial explanation of depth perception in the textbooks.

In 1833 a new correlate of visual depth was discovered. In contrast with previous theorizing based only on self-observation, this was a truly experimental discovery. With a theory in mind, Wheatstone invented an optical device to test it, which he called a stereoscope. His idea was that the discrepancy between the two retinal images of an object on which the two eyes converged was not simply the paradox it had previously been considered (How can we see two different views as the same thing?) but was a basis for perceiving the object in depth.

The stereoscope produced a synthetic disparity of the two images, and this could be modified at will. The experimenter could draw pairs of geometric figures, one for each eye, differing in various ways and the instrument would project each upon its appropriate retina. If a lawful relationship could be established between the disparity and the perceived depth of the optically combined figure, then disparity was a cause of depth-perception. Everyone who has looked at stereograms knows how strikingly this theory was verified. The fact of binocular image-disparity at once became accepted and still remains the chief explanation of how we see the third dimension. Whether the disparity should be thought of as a clue for an interpretive perception of depth or as a kind of binocular sensation yielding depth immediately was not easy to decide. It was in any event a demonstrable

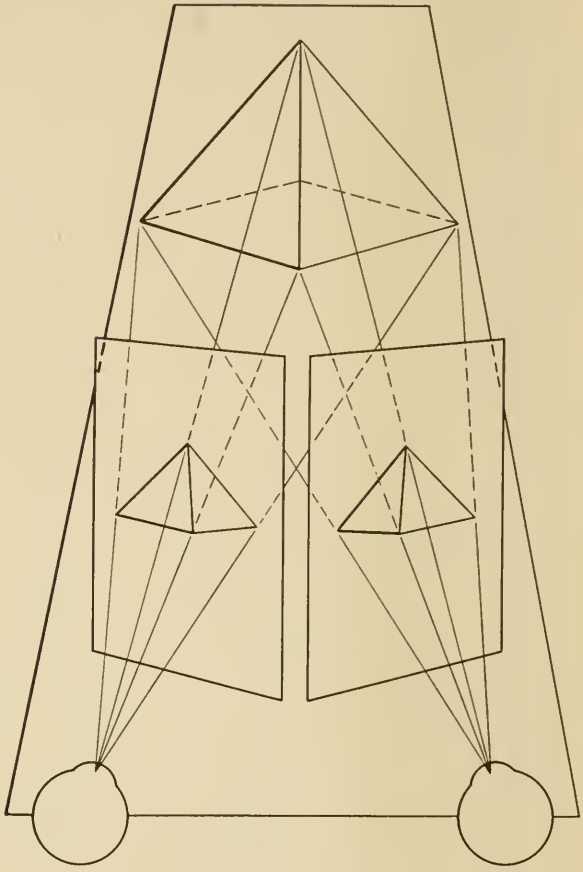


FIGURE 4. The Disparate Views of an Object by the Two Eyes

correlate of the experience of depth, unlike the hypothetical sensations of convergence and accommodation, and its discovery made the dogma of an innate intuition of space – of space as an inner condition of all experience – less likely than ever before.

In the outcome, the three classes of data derived from convergence, accommodation, and above all from retinal disparity were taken as the primary criteria for distance and depth and as the only discoverable basis for the perception of abstract three-dimensional space. There did exist, to be sure, in any concrete visual world such as a view of the countryside or a scene which a painter might choose to represent, a number of other clues for detecting the distance of things. If one object seems to "cover" another, it must be nearer. If edges known to be parallel seem to converge, they must really recede; and if objects known to be of similar size seem progressively smaller, they must really be progressively farther away. If one thing appears above another it is probably not suspended in the air but merely lying on the ground at a greater distance. If an object seems bluish and blurred it must be distant like the hills on the horizon. If an object is partly in light and partly in shadow its surface cannot be flat but must really be curved or bent. If a thing seems to move, or be displaced across other things when the observer moves his head from side to side, it must really be nearer than the other things in proportion to its relative motion. All these clues had been known long before the perception of distance ever became a philosophical issue. With the exception of the last named, they had been employed for centuries by painters in their effort to reproduce a segment of the world on a flat surface. When the philosophers and psychologists began to examine their visual sensations they inevitably began to view the world pictorially, as artists had learned to do, and these rules of picturing were recognized as being indicators or signs of distance for a retinal picture as well as for a painted picture. But these clues could not be given the same explanatory value that could be ascribed to convergence and binocular disparity. They were themselves perceptions, it appeared, not data of sensation; even the most convinced nativist could not argue that they were pure intuitions of space; they must obviously therefore be learned by experience. They were called *secondary* cues for depth and distance to distinguish them from the *primary* cues of convergence, accommodation and retinal disparity. Since they did not depend on the existence of two eyes, they became known also as the *monocular* cues, while convergence and disparity were *binocular* in origin. Although they have been described many times, re-observed by successive generations of curious men, and have passed into common knowledge as facts having to do with the perception of space, they have never been systematically controlled, varied, and subjected to experiment. In succeeding chapters we will have much more to say about them, and their significance may then appear in a new light.

The theory of *cues* as the explanation of our perception of the world has proved, in the eighty years since Helmholtz per-

fected it, more convincing than the alternative theories. Complicated as it is, it has seemed, to Americans at least, to be the only *scientific* explanation, for it did keep open the possibility of investigation whereas any appeal to intuition rendered experimentation impossible. It assumed that sensing and knowing were two different things and that all knowledge came through sense. Many phrases of common-sense psychology are reflections of this assumption; the "messages" of the sense organs, or the "information" or "facts" that they supply to the mind, imply a set of clues and a process of interpretation. The mind, it was assumed, is intelligent and acts on the sensations somewhat as a geometer or a logician would act, combining, computing, and comprehending the data it gets in much the same way as did the philosophers themselves when they invented the theory.

Gestalt Theory

The theory that sensations were data or cues for perception lasted a long time, but it had troublesome implications. For one thing, unless perception were purely intuitive, it had to be a kind of compounding or putting together of elementary sensations by means of associative learning. But these sensory elements could never be specified. They could hardly be points of color corresponding to the single spots of excitation on the retina since, after all, points are nothing but geometric fictions; at the same time no one could discover how they could plausibly be anything else. Furthermore, the theory of cues could never really explain how we *see* the world, or why it looks the way it does, but only how we can make *judgments* about the world. Both of these objections were raised some twenty-five years ago by the Gestalt psychologists.

The Gestalt theory started with the problem of how we can see visual form. Instead of simply adding a "form-quality" to the list of sensations, however, it took a new line of thought and asserted that a form was not compounded of sensations at all. Experience is not reducible to elements or additive units, the argument went, and when it is analysed introspectively into sensory components it is falsified. But if not constituted of sensations, how is a unitary perception of this sort to be accounted for? That there had to be a special perceptual process of some sort, the Gestalt psychologists never doubted. Observing that under experimental conditions visual patterns or dimly seen forms tended to be perceived as symmetrical, connected, completed, and meaningful, even though the drawings presented to the observer were not, they concluded that these tendencies were laws of the perceptual process in general and were indicative of its nature. Forms seemed to occur spontaneously in perception even when the picture constructed by the experimenter was objectively incoherent and meaningless. The theory of perception which occurred to them was that the process was one of relatively spontaneous *sensory organization*. The process of organization was assumed to occur in the brain, presumably at the level of the cerebral cortex. It was conceived as a process in a field, analogous to the visual field itself, and the parts of the field (the contour of the form and its background) were

united or separated by forces of attraction and repulsion similar to electro-magnetic forces. A perceived form in this theory, is a brain-form. The retinal image yields isolated single excitations. Only when these are projected on the cortex do the field-forces begin to operate among them and only then do they unite in a Gestalt. The causes of sensory organization are to be sought in what is sometimes called field-theory.

The Gestalt theory is not as explicit about the perception of space as it is about the perception of form. But it was based on a description of what the world looks like, not what it ought to be like geometrically, and it therefore asserted that all visual perception is tridimensional from the outset. The theory of perception as organization led to the following reasoning. The brain is a three-dimensional organ and the neural process of dynamical organization must therefore occur in a three-dimensional field. The perception itself, then, would naturally be three-dimensional if the underlying physiological events were. The reader may or may not find this argument convincing. In any event, this was about as far as the Gestalt theory could go with space, except for Koffka's analysis of the hypothetical field forces which might underlie binocular retinal disparity (67).

Perhaps the greatest contribution of the Gestalt theorists was that, having taken an unprejudiced look at the visual world they were trying to explain, they formulated problems for space-perception which were genuinely relevant. How is a figure separated in perception from its background? What is a surface? What is a contour? Why does the world look upright? How is the phenomenal ego located in it? Why do things appear to have very nearly their true size and color despite the variations in their retinal images? These were questions about phenomena of a wholly different kind from the geometrical points and lines of the nineteenth century psychologists, and these were the questions which the Gestalt psychologists asked. They were questions about the characteristics of the visual *world*. The only difficulty is whether the hypothetical process of sensory organization yields the answers to them.

The Fact of Perceptual Constancy.

The trend of thought which the Gestalt psychologists represented was responsible for more than a new theory of perception; it resulted in a massive amount of experimental evidence. An important part of this concerned the problem of what was called perceptual "constancy." By this term was meant the fact that perceptions, or phenomenal objects, kept their identity and their objective size, shape, and color despite variations in the retinal images with which they corresponded. Although the retinal image was a poor indicator of objective shape (so it seemed, inasmuch as it changed from one aspect to another as the observer moved) the perception was nevertheless in good agreement with the objective shape. In short, it tended to remain *constant*.

This kind of fact could be tested by experiment and measured; moreover, it was meaningful in terms of human behavior and it escaped from the atmosphere of

respectable unreality which clung to the nineteenth century problems of space perception. The interests of twentieth century psychologists began to shift toward the question of why perception was objective and away from the purely theoretical aspects of this paradox.

If the objectivity of perception could be studied in the laboratory it was no longer a speculative question of epistemology but a matter for experimental investigation. In a size-constancy experiment, for example, an observer is required to equate the size of a wooden stick near at hand with that of a stick placed at a distance. In general this task can be performed with some accuracy. The size of the retinal image of the far stick, however, may be only a quarter or an eighth that of the near stick, the rule of optics being that doubling the distance of the object halves the projected size of the image. Since the impression of size obviously does not depend on the image, on what can it depend? A reasonable answer would be that it depends on the whole stimulating situation or, more specifically, on the stick-image in relation to its background-image of three-dimensional space.

It is only a step from this kind of reasoning to the proposal made in the next chapter: that there exist, as extremes, two kinds of seeing, (1) the experience of a visual world in which objects stay the same size wherever they are and in which parallel edges do not converge, and (2) the experience of a visual field in which the principles of perspective hold true. Constancy of size would then be a corollary of the visible depth and distance of the visual world.

Summary

If everything we are aware of comes through stimulation of our sense organs, and if some things nevertheless have no counterparts in stimulation, it is necessary to assume that the latter are in some way synthesized. How this synthesis occurs is the problem of perception. Our awareness of the world of objects and space is particularly difficult to account for but also particularly important, since it permeates nearly all kinds of experience. Theories of the perception of objects and space, therefore, have a long history. Nativism assumed that the synthesis was intuitive or innate. Empiricism explained the synthesis as learned or inferred from past experience. More recently, Gestalt theory has suggested that it is produced by a characteristic achievement of the central nervous system which may be termed sensory organization.

The difficulty in postulating a consistent learning theory is that many kinds of perception seem to occur in children and animals who have had no opportunity to learn. Sensory organization, as a descriptive term, appears to fit these facts somewhat better. If it is necessary to assume some kind of synthesis of visual stimuli, "organization" is a better word to use than "reasoning" or "inference." As a theory of what might go on in the nervous system, however, "organization" is less valuable. It is true that physical and biological processes are often characterized by organization (the tendency of electric circuits to reach an equilibrium and the subordination of parts of an organism to the whole during the growth of the embryo) but when this concept is applied to the

physiology of visual perception it has a fundamental weakness: it does not in itself explain why a perception is like its object. The characteristic of perception is that the result is not so much spontaneous as it is faithful to the thing perceived. The question is not how a percept gets organized but why it is always organized like the particular entity toward which the eye happens to be pointing.

The Gestalt psychologists made much of the spontaneous character of the process of perception, but they were aware of the problem *of some kind* of correspondence between retinal stimulation and our awareness of things. Koffka, in his *Principles of Gestalt Psychology*, spoke of a "more comprehensive correspondence between the total perceptual field and the total stimulation" (67, p. 96) and implied that this correspondence would be clarified when the laws of sensory organization were known. What this book attempts is a direct explanation of this comprehensive correspondence. If the total stimulation contains all that is needed to account for visual perception, the hypothesis of sensory organization is unnecessary.

3

The Visual Field and the Visual World

The Bounded Visual Field . . . . The Gradient of Clarity . . . . The Effect of Eye and Head Movements . . . . The Location of After-Images . . . . The Apparent Size and Distance of After-Images . . . . The Effect of the Posture of the Head and Body . . . . The Apparent Size and Shape of Objects . . . . The Apparent Convergence of Parallel Lines . . . . The "Eclipsing" of Forms . . . . The Visual Field during Movement of the Observer . . . . The Awareness of Distance . . . . Summary . . . . The Problem of the Visual World

If we are to understand the problem of why the visual world looks as it does the first thing to do is to look at it. What actually *does* it look like? This question is not as easy as it sounds. It requires that we carefully examine our experience and then find the essential terms in which to describe it. The description needs to be carried out without preconceptions and without reference to theories as to how vision might occur. The known facts of vision may be kept in mind, but the known theories and their implicit terms should be disregarded. The problem is to state without any theoretical prejudgment what we see when we say that we perceive the environment.

Try making this observation for yourself. First look around the room and note that you see a perfectly stable scene of floor and walls, with an array of familiar objects at definite locations and distances. Every part of it is fixed relative to every other part. If you look out the window, there beyond is an extended environment of ground and buildings or, if you are lucky, "scenery". This is what we shall call the *visual world.* It is the familiar, ordinary scene of daily life, in which solid objects look solid, square objects look square, horizontal surfaces look horizontal, and the book across the room looks as big as the book lying in front of you. This is the kind of experience we are trying to account for.

Next look at the room not as a room but, insofar as you can, as if it consisted of areas or patches of colored surface, divided up by contours. To do so, you must fixate your eyes on some prominent point

and then pay attention not to that point, as is natural, but to the whole range of what you can see, keeping your eyes still fixed. The attitude you should take is that of the perspective draftsman. It may help if you close one eye. If you persist, the scene comes to approximate the appearance of a picture. You may observe that it has characteristics somewhat different from the former scene. This is what will here be called the *visual field.* It is less familiar than the visual world and it cannot be observed except with some kind of special effort. The fact that it differs from the familiar visual world is the source of a great deal of confusion and misunderstanding about vision. It is the experience on which the doctrine of visual *sensations* is based. It is strictly an introspective or analytic phenomenon. One gets it only by trying to see the visual world in perspective and to see its colors as a painter does.

Both the visual world and the visual field are products of the familiar but still mysterious process known as seeing. Both depend upon light stimulation and upon a properly functioning eye. But the differences between them are so great as to suggest two kinds of seeing. Let us try to list and describe these differences. Most of them can readily be observed without special apparatus, and the reader should therefore check them for himself as we go along.

The Bounded Visual Field

In the first place, the visual field has boundaries, whereas the visual world has none. If you keep your eyes fixed but put your attention on the periphery of the field (a trick that may require practice) you can observe that things are visible only to a limited angle out to the right and left and to an even more limited angle upwards and downwards. These boundaries, it is true, are not sharp like the margins of a picture and they are hard to notice, since all vision is unclear in such eccentric regions, but they are nevertheless present. The field is roughly oval in shape. When measured, it extends about 180 degrees laterally and 150 degrees up and down. If you close one eye you will notice that about a third of the field on that side disappears and also that the boundary is now the outline of your nose. Many an otherwise observant individual does not realize that his nose is represented in his visual field. Even if shadowy, however, it has always been there and its discovery only illustrates the unfamiliarity of this kind of seeing as compared with the familiar reality of ordinary perception.

What Ernst Mach, analyzing his sensations, called the phenomenal ego is illustrated in Figure 5. It is a literal representation of his visual field, with his right eye closed, as he reclined in a nineteenth century chaise longue. His nose delimits the field on the right and his moustache appears below. His body and the room are drawn in detail, although he could not see them in detail without moving his eye. The margins of the field are shown as definite and clear whereas of course their actual appearance was very vague. The point of fixation cannot be shown in the drawing; actually it is the center of the field and this should be the only part shown as wholly clear.

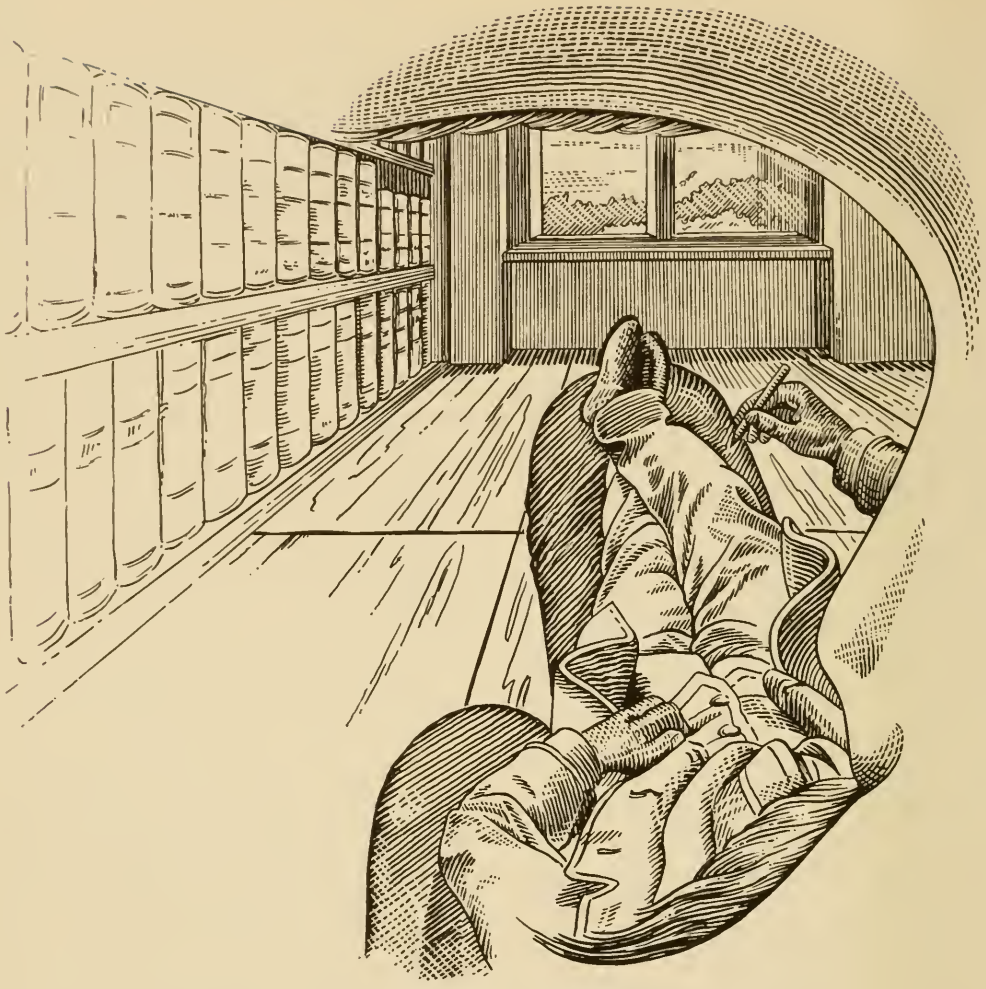


FIGURE 5. The Monocular Visual Field of Ernst Mach

The visual world, on the other hand, is certainly not delimited by an oval-shaped boundary. Floors, walls, and terrain are visibly continuous. As Koffka has pointed out (67), one is ordinarily aware of a world which extends backward behind the head as well as forward in front of the eyes. The world, in other words, surrounds us for the full 360°, in contrast to the visual field which is confined to about 180°. Whether the world which includes this space behind us is a strictly visual world or not is a question of definition rather than a matter of ordinary observation. It cannot be answered by inspection for the reason that in the effort to examine the experienced world one finds oneself inspecting the visual field instead. The visual world, as we shall discover, will not bear up under much introspection and analysis without changing its character. It is at least clear that the visual world

does not have boundaries. It has a panoramic character which the field does not possess.[1]

The Gradient of Clarity

A second characteristic of the visual field is that it is sharp, clear, and fully detailed at the center, but progressively vaguer and less detailed toward its boundaries. For instance, the contours and patterns of the array of surfaces in your field can be observed to become gradually less determinate as you attend to those out toward the periphery. So difficult are the latter to see that the impulse to turn the eyes and fixate them may seem almost irresistible.[2] If you move your eyes down this page of print, for example, and fixate at random one letter of a single word, you will probably find that you can perceive that word and the words adjacent to it on the right and left and above and below, but no more. The visual field, therefore, possesses a central-to-peripheral gradient of clarity. The visual world does not. It does not even have a center, which agrees with the fact that it does not have boundaries. The world is ordinarily perceived by *scanning*, that is, by moving the eyes rapidly from point to point, and the objects and surfaces which compose it are always clear and fully detailed. If the objection be advanced that they are in fact only clear and detailed *when fixated*, the answer is that the objector gets this fact from an inspection of his visual field, not his visual world.

The Effect of Eye and Head Movements

The visual field shifts whenever the eyes are moved from one fixation point to another, since the eyes normally play over the visual environment in much the same way that a searchlight moves over a night sky except that light is being absorbed by them instead of emitted. Scanning movements of this sort are termed saccadic eye movements, and are rapid jerks of very brief duration. If the shifts of fixation are wide the head also moves in the same direction as the eyes and, as a result, the boundaries of the visual field formed by the eyelids and nose sweep across the array of colored patches. If

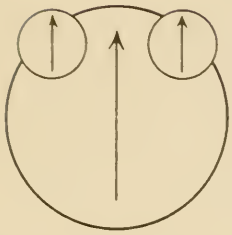
---

[1]To the reader familiar with Koffka's distinction between the "behavioral world" and the "geographical world" (67), it should now be clear that his is a quite different distinction from the one now being made. The "behavioral world" for Koffka was the whole field of visual experience. The point of this chapter is that visual experience needs to be subdivided into a bounded or field-like kind of experience and an unbounded or world-like kind of experience. The "geographical world" was Koffka's name for the physical environment. That there *is* a physical environment, neither Koffka nor the writer nor, presumably, the reader doubts.

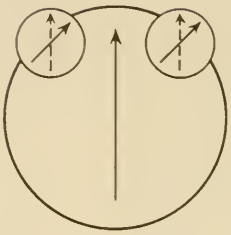
It should also be clear that the visual field as here defined is not the same thing as the "phenomenal field" as this term is employed by many writers, or the kind of field conceived by what is called "field-theory." These latter usages fail to distinguish between the delimited and the panoramic kind of experience. The distinction here being made is, however, similar to Brunswik's conception of two kinds of perceptual achievements, the seeing of "perspectives" and the seeing of "constants" (15).

[2]The center of clear perception corresponds, of course, to the fovea of the eye — that area of the retina best equipped anatomically for discrimination of fine detail and on which is projected an image of the object toward which the eye is pointed.

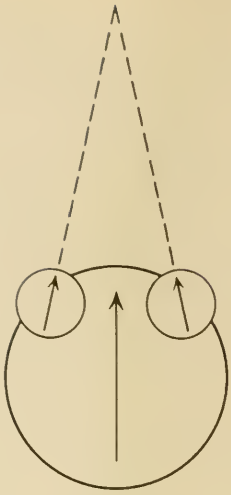
FIGURE 6. The Types of Eye Movement



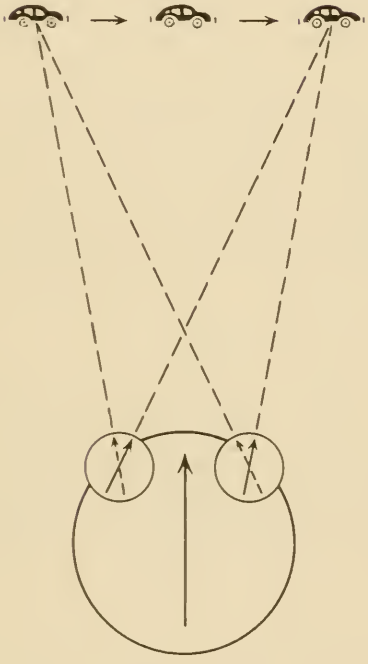
The "primary position" of the eyes in the head



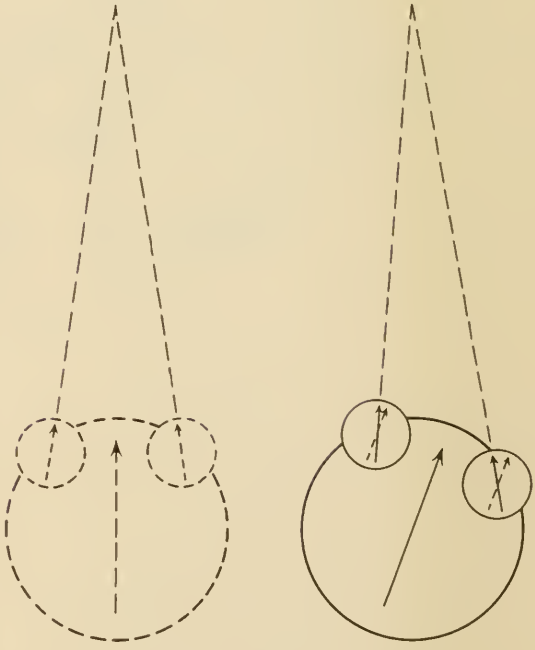
A saccadic eye-movement



Convergence of the eyes



A pursuit eye-movement



Compensatory movement of the eyes in the head (the result of either a voluntary turning of the head or a passive rotation of the head)

you will stand in the middle of the room, close one eye, and turn around or walk in circles, you can observe the way in which these boundaries sweep over the walls as your head turns.

One of the most obvious characteristics of the visual world is its stability. The world does not rotate as you turn around (you would become badly disoriented if it did) nor does it shoot from side to side or up and down as you shift your fixation from one object to another. This fact is so obvious that most of us take it as a matter of course and do not realize that there is any need for explanation. And yet it is really a very astonishing fact. Things possess a direction-from-here not with respect to the margins of the visual field but with respect to a fixed visual world – an external frame of reference which seems unexplainable on the basis of the retinal picture. Try the following experiment: with one eye closed, select some prominent object and then look alternately toward a point just to the right of it and another point just to the left of it. The object will not seem to move. Try as you will to see it as a patch of color which goes shooting from the right to the left side of your visual field, you will probably have only indifferent success. You may be able to see it as displaced from one side to the other of your field, if you concentrate on the boundaries, but you will not see motion. Next, fixate the object and put your finger at the outer corner of your open eye so that you can feel the eyeball under the lid. Press on the eyeball just enough to move it and release it alternately. This time you will see the object move unmistakably. The visual world as a whole is not stable but moves back and forth. In both these situations the same thing has occurred – a displacement of the retinal image across the retina proper – but there has been quite a different result in perception. This result must be due to the difference between the two kinds of eye movement, natural and artificial.

During the natural eye movements of scanning, the visual world and even the colored surfaces of the visual field appear not to move. But there is another type of natural eye movement, the pursuit movement, in which it makes a difference whether the world or the field is attended to. Hold a pencil in front of your eyes, fixate it, and move it slowly from right to left. Looking at the situation as objectively as possible, the motion that you see tends to be concentrated in the pencil rather than in the world behind it. But if you now continue to fixate the pencil but attend to the background, the motion of the latter becomes more obvious. Seen pictorially, or as a field, the illusion of a moving environment is fairly compelling.

The Location of After-Images

Another way of demonstrating the directional stability of the visual world despite movements of the eyes, and at the same time showing that there is another directional system for vision with respect to the eyes themselves, is to observe the location of after-images. Nearly everyone has seen negative after-images and noted that they behave like "spots before the eyes" or other so-called entoptic phenomena. Such phenomena are forms of localized retinal stimulation but, since

they are not projected by light from outside the eye, they are not displaced on the retina when the eye moves. For the same reason they do not disappear when the eyes are closed. When the eyes are open, they appear to be superposed on the objects of the visual world but not to be objects themselves, and they have a filmy insubstantial look.

The fact is that after-images, unlike objects, do jump about when we scan the environment. Their direction-from-here is given with reference to the center of clear vision and the boundaries of the field. You cannot "look away from" an intense negative after-image, and it will reappear wherever you fix your eyes.

After-images, therefore, are localized with reference to the visual *field*. Insofar as they have a visible location, it is in this field. Objects, on the other hand, are located in the visual *world*, which possesses its own independent directional system. Hence, if one attends to the visual *field* in the intervals between movements of the eyes an object (as a patch of color) appears to be displaced, whereas if what you are attending to is the *visual world* the after-image appears to be displaced.

The Apparent Size and Distance of After-Images

After-images are localized in the visual field, not in the visual world, with respect to up or down and right or left. How are they localized with respect to distance? How far away do they look? As everyone knows who has observed an after-image with his eyes open, it appears to be superposed on whatever surface one happens to be looking at and to be at the distance of that surface. In this respect, therefore, after-images do have a certain kind of location in the visual world. They seem to attach themselves to surfaces if there are any surfaces present.

This fact has a very interesting corollary, which will have a special significance when we come to consider the perceived size of objects in the visual world. The apparent size of an after-image becomes greater when one fixates a more distant surface. The seen size is very nearly proportional to the seen distance, a relationship known as Emmert's Law. It suggests that the impression of size must be closely linked to the impression of distance for, of course, the size of the after-excitation on the retina of the eye does not change.

If, on the other hand, one observes an after-image against one's closed eyes, or in absolute darkness, or against the cloudless sky, it seems to float in what might be called an indefinite space. It does not seem to have any precise distance and likewise no precise size.

The Effect of the Posture of the Head and Body

There is still another effect of the observer's movement on his visual field which does not hold true for his visual world. If you tilt your head 90°, or lie down on your side, the patchwork within your field rotates and you may be able to see the physically vertical lines of the room as possessing a kind of horizontal quality. They now extend from right to left instead of up and down. Considering the room *objectively* from this position, however, it is obvious that the room is still upright and that the lines where the

walls meet are still aligned with the up and down of gravity.

The room as a picture may appear as if it had been tilted over on its side, but the room as a room is still upright. The pictorial direction "downwards" goes one way but the objective direction goes quite another way, and the former has the quality of being illusory. The physical vertical of gravity, we may conclude, is somehow implicit in all such tilted visual fields (when they result from a voluntarily tilted or reclining posture) and it is never quite lost. Consequently, no matter how we lie or sit, the visual meaning of how to stand "up" can always be depended on. The direction of "up" in the visual world is aligned with the direction of gravity (42).

There are, it is true, a number of situations in which this sense of the gravitational vertical for the visual world is temporarily lost, and there are diseases – usually of the organs of equilibrium in the inner ear – in which it is permanently impaired. In some flying maneuvers, in amusement park devices, in a special type of vertigo, and in a number of experimental situations (42) the visual world and the visual field cannot be distinguished from one another and some illusory frame of reference – a non-gravitational vertical – may then dominate perception. The experience is disconcerting and unpleasant. It is in these situations that one loses equilibrium.

In the activities of ordinary behavior we may infer that there is a visual vertical-and-horizontal frame of reference which is linked to gravity and is presumably mediated by the muscle sense and the inner ear. It serves to keep the visual world upright and aligned with gravity. But there are also other systems or frames of reference, linked to the boundaries of the field of view, or to the axes of the head or body, or to the lines of the visual field, which may be in conflict with the physical axes and which would then give a visual field not aligned with gravity. Usually, but not in all conditions, such a field has an illusory quality.

The Apparent Size and Shape of Objects

We now come to the differences between the visual world and the visual field with respect to depth. These differences are not so easy to observe as some of those already described, but they are more important for our central problem. The field has been said to have a "pictorial" quality. A picture is something that can be defined by mathematics and optics. The essential physical fact about a picture is that it consists of a projection of objects in three dimensions on a plane of two dimensions. Insofar as the field of view can be seen as a picture, therefore, it will have the characteristics of a projection. Keeping this fact in mind, let us compare the appearance of the visual field with the visual world.

In one sense of the words "to see", objects are seen to decrease in size as they become more distant. In another sense, however, they remain constant in size, whatever their distance. There are transitional stages of seeing between these two extremes, stages which depend on the conditions of observation as well as upon the attitude of the observer, but the fact is that constancy of size tends

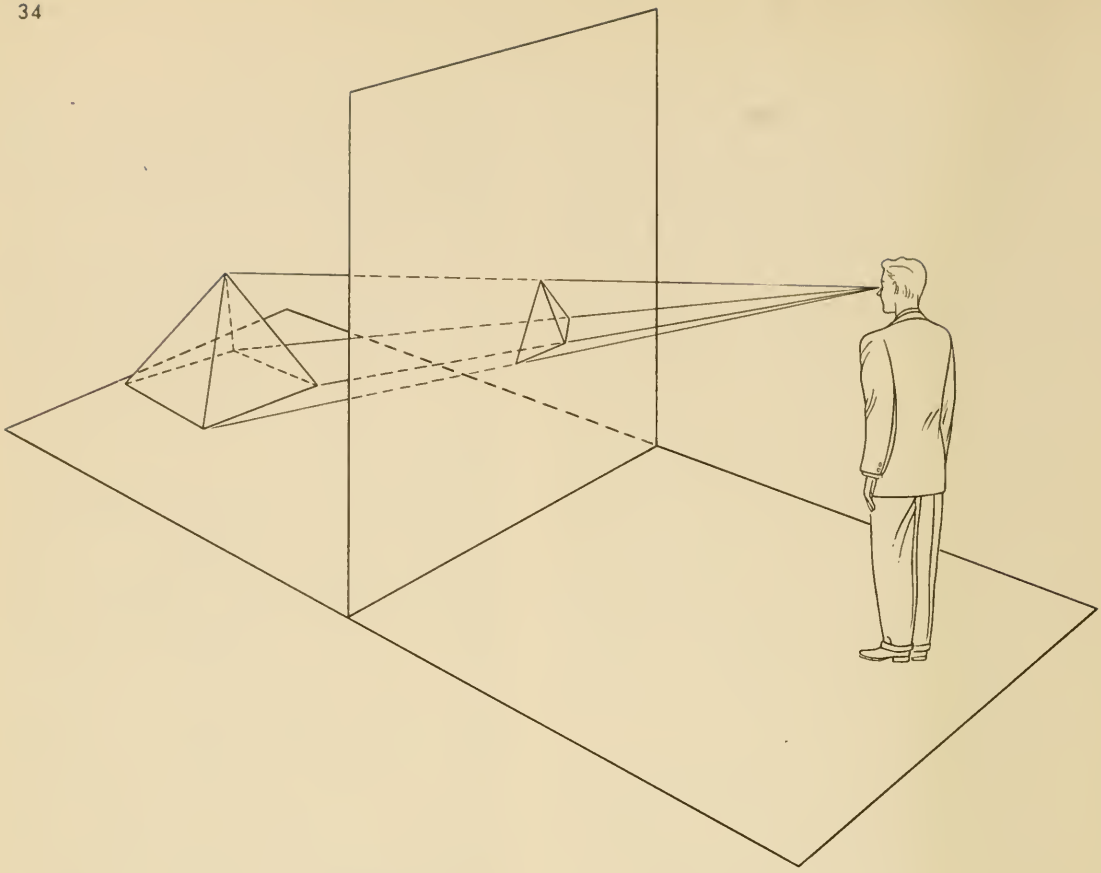


FIGURE 7. An Object Projected on a Plane

to be preserved under the natural conditions of attending to the visual world. Under these conditions an object is seen at a visibly determinate distance.

The same thing is true of the shape of objects. Whatever the orientation of an object to the line of regard, whether we see it from the front, the side, or the top, if the conditions for observation are adequate, it will have the same shape. Now there are two meanings for the word shape. In this context, we mean the shape which an object possesses in three dimensions and which is defined by its surfaces. We shall call this its "depth shape." There is also a more common meaning of the term, the shape which an object possesses when projected on a plane. This is its shape as a silhouette, or the shape which is defined by the outlines or contour. This is its "projected shape." That shape of an object which remains constant from whatever direction it is viewed is its depth shape. That shape which changes with the angle of view – the "aspect" of the object as we say – is its projected shape. It is obviously important to specify which of these meanings is being employed when one talks about shape, and a good deal of confusion has resulted from not doing so. The visual world contains depth shapes, whereas the visual field contains projected

shapes. As you walk about in a room you can, first of all, observe that objects do not change shape in the first sense of the term, and secondly, you may be able to note that the projected shapes do change, especially if you fixate an object as you walk. The only kind of an object whose projected shape would not change in such circumstances would be a perfect sphere.

It so happens that most of the controlled observations of this phenomenon carried out by psychologists have been made with flat objects (whose depth shapes are approximately the same as their projected shapes when the latter are seen from directly in front). Such objects, unlike most, have a unique orientation in which they are best viewed. The example frequently given is a dinner plate. When the conclusion is reached that the "shape" of such a stimulus object tends to remain constant no matter what its angle to the line of regard, it is not clear whether the observer means its depth shape or its projected shape viewed head on. In the case of the dinner plate the former is a solid disk, bent into a rim around the edge, and it is perceived as such in any orientation. The latter is an abstract geometrical circle – a special kind of projected shape. Strictly speaking, it is the former that remains constant. When you simply ask an observer what the "apparent" shape of the dinner plate is, without specifying that you mean its depth shape, there is room for argument as to whether the shape is a circle or some kind of an ellipse. The conventional statement that a dinner plate always looks circular is inexact. What it always does look like is a dinner plate – the three-dimensional shape in the visual world.

The Apparent Convergence of Parallel Lines

All of us have observed the fact of linear perspective at one time or another – the fact that equidistant edges of man-made structures appear to get closer together, after a fashion, as they recede in the distance. When looking down a highway or railroad tracks the effect is strong; when looking at a building or observing the interior of a room it is less obvious and may be difficult to note. Even in the case of the railroad tracks, however, two observers may differ in describing what they see. One will report that the parallel lines definitely converge as they go off toward the horizon; another will insist that the rails do not converge since they are visibly equidistant. Each scene is perfectly clear to each observer, but they are contradictory to each other. Now this fact does not in the least prove that each observer creates the visual scene in his own fashion and that we all have private worlds. It suggests only that there may be two kinds of seeing. Perhaps both observers are correct, but are simply using the verb "to see" with different meanings.

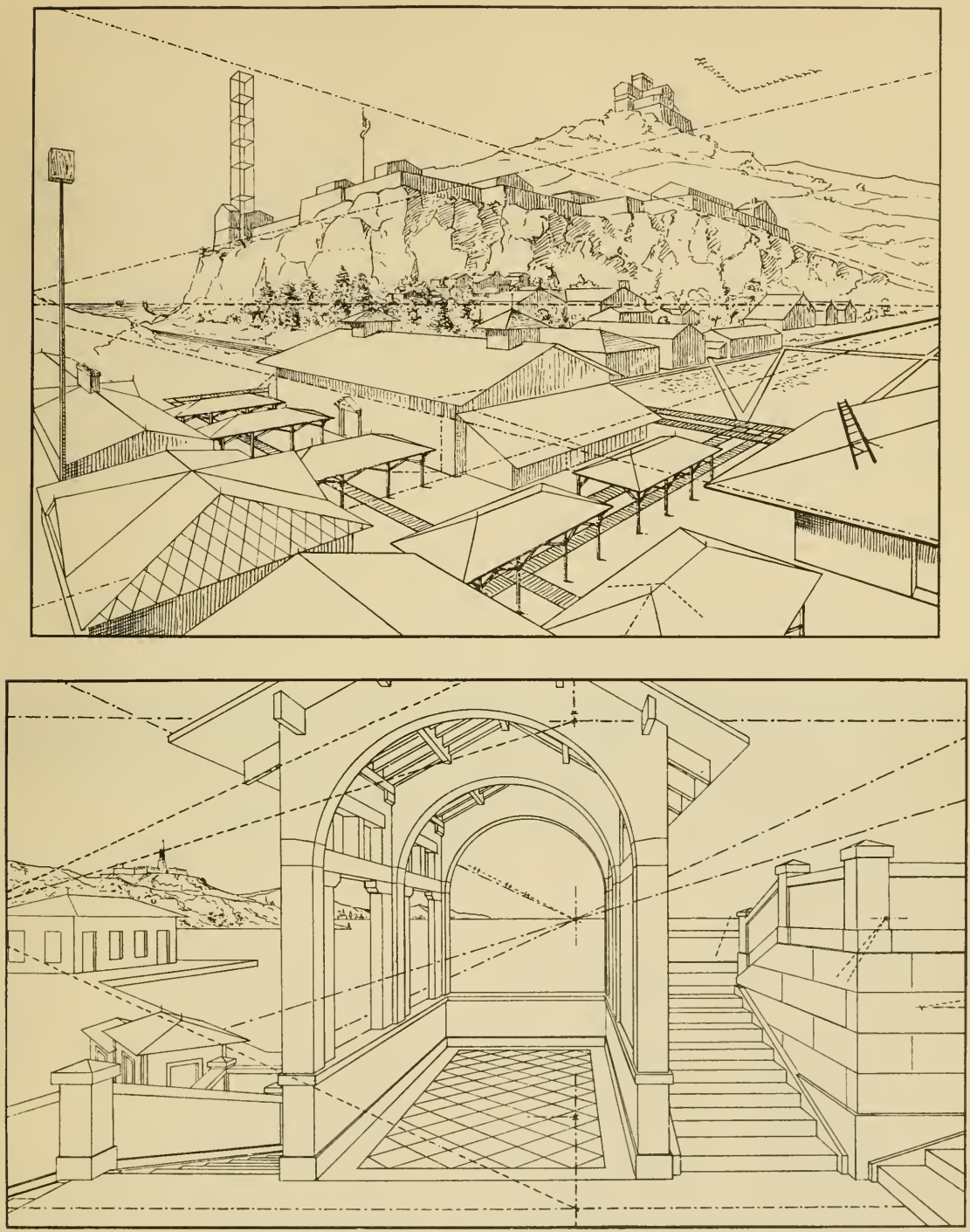
If you lay a sheet of paper on a table in front of you and then look at its right and left hand edges, you will probably not be able to see them as converging. Close one eye and try it again. Unless you have been trained to visualize things in perspective, the sides of the paper will still tend to remain stubbornly parallel. But if you take a pencil in either hand and, with one eye still closed, hold them perpendicular to your line of sight and then align

(By Life photographer Herbert Gehr. Copyright Time, Inc.)

FIGURE 8. Convergence of Parallel Lines to a Vanishing Point

them with the edges of the paper, you will be surprised to see how much the edges do converge when projected on an imaginary plane in front of you. Still holding the pencils in position, try to visualize the lines of the pencils projecting upward until they intersect. They meet at a point exactly at eye level, that is, on the horizon. If you note where this point is superposed on the nearest wall, you can see that it is where the wall would be cut by the horizon of the terrain outside.

You may now have a clearer conception of the visual field as approximating a plane projection. On such a projection parallel lines *do* meet – not at "infinity" but at eye-level, if they are parallel to the ground. On the other hand it should be clear that Euclid was also correct in his postulate that parallel lines do *not* meet. Euclid's proposition applied to the visual world. The observer who saw the railroad tracks as continuously equidistant was aware of the environment as a Euclidean scene, not



(From W. R. Ware, Modern Perspective *(N.Y., Macmillan, 1900)*

FIGURE 9. Two Scenes in Perspective

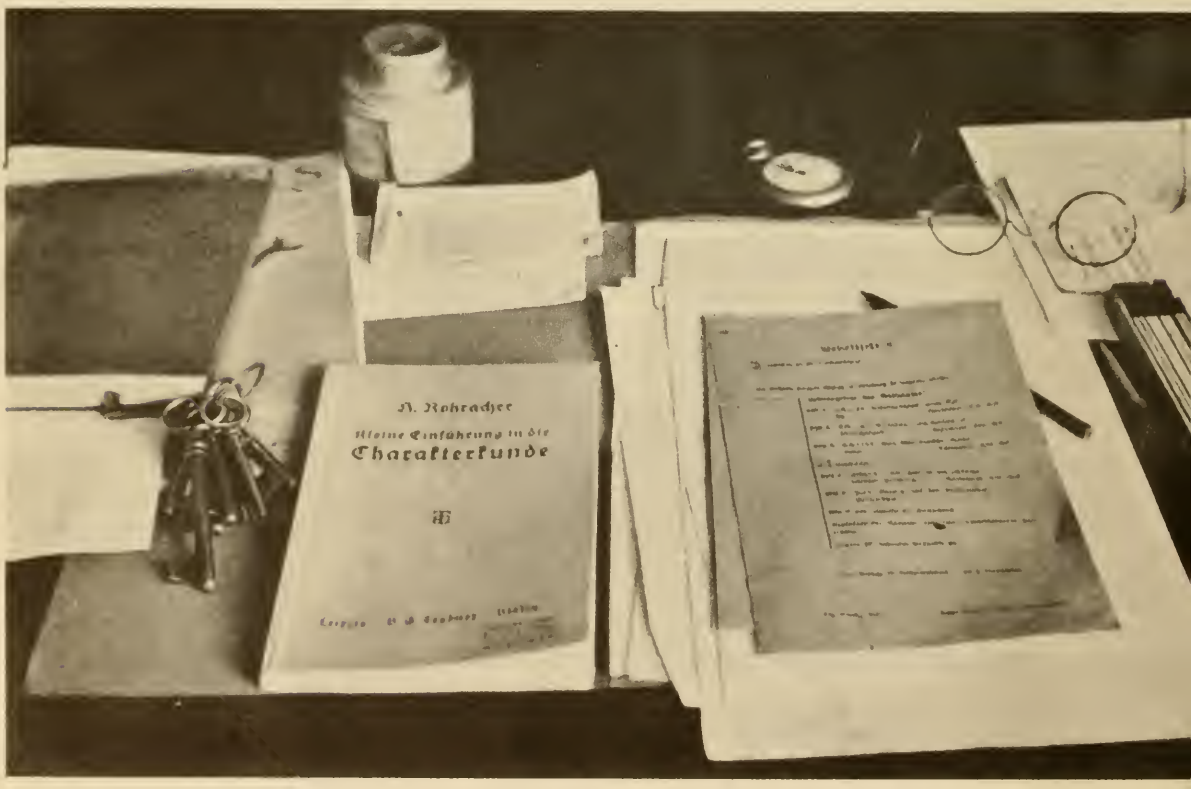
as a scene in perspective. He saw it as a locomotive engineer would, not as a painter. And if the rails appear even slightly convergent to an engineer, it is time to apply the emergency brakes.

The "Eclipsing" of Forms

When we are attending to the visual world and our eyes move over the environment, the points of fixation are the objects in it. These are the elements which arouse our interest and affect our behavior. We do not attend to the spaces between the objects – the gaps or background – and we are almost unaware of their existence. But a little attention to the visual field shows that these interspaces are just as truly parts of it as the areas representing objects. In the field as a projection, the background is not different from the objects in the compelling way it is when you observe the world. The interspaces, like the objects, are areas of color, and the field therefore approximates the appear-

FIGURE 10. Inattention to Interspaces in Ordinary Perception

Can you see two pencils in the photograph? The hidden pencil has an uninterrupted contour, and you might suppose, therefore, that it would be easier to see than the visible pencil. Its contour however is nearly all "used up" by the two pamphlets, and the pencil becomes part of the background, merely an "interspace." *(The photograph was devised by W. Metzger and reproduced in* Gesetze des Sehens. *Frankfurt am Main: Kramer, 1936)*



ance of an irregular patchwork. Considerable study has been devoted to the consequences of this phenomenon by Gestalt psychologists, beginning with the work of Rubin on what he called figure and ground (91).

Natural visual scenes, however, do not divide up neatly into figures and background. In most of them it is a relative matter whether a given area be regarded as a figure or as a background. One object may be the background for another nearer object, and another larger object may be the background for the first. The largest of all "objects" – the object which is literally *fundamental* to the perception of space and the most comprehensive of backgrounds, as we shall try to show later on – is the terrain. Consider now how this phenomenon of relative backgrounds is related to the visual field and to the visual world. In the field, the area corresponding to one object may be diminished by an area corresponding to another object which lies in front of it. Seen as a field, with the head and eyes fixed, one area can be described as *eclipsing* the other, to use an astronomical term. Seen as a world, however, one object lies *in front of* another. In Figure 11, for instance, some areas appear to be in front of others; some do not appear superposed at all; and in some a slight change in the common contour reverses the suggestion of depth. A possible explanation of this will be given at the end of Chapter 7.

Presumably there are transitional stages between the extreme cases of adjacent areas and superimposed areas, and a number of factors play a part in determining how these will be seen. It is clear that even

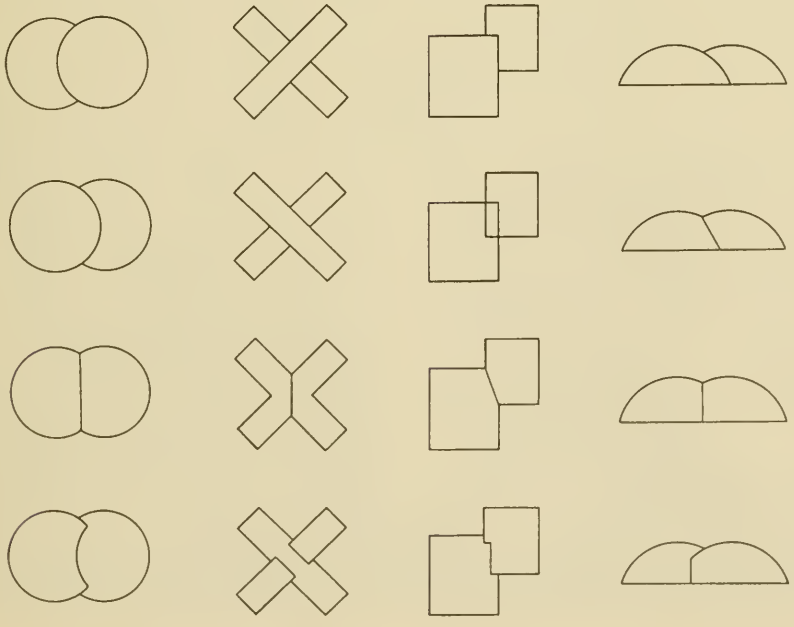


FIGURE 11. One Object in Front of Another

projected shapes like geometrical forms tend to be seen in superposition under the influence of these factors. In a visual world of the sort provided by rooms, streets, and countryside, the actual fact is that we *see one object behind another.* Koffka once argued convincingly that there was truth in the statement that he could see his desk-top extending uninterrupted beneath the book which lay upon it (67), although this statement seemed to violate accepted principles of vision. If the statement seems dubious, the reader may try it for himself.

An illustration of how both kinds of seeing may be obtained from the same stimulus situation is provided in Figure 12. You see in the picture a crowd of individuals, each anatomically complete. These objects, however, as projected shapes are rather thoroughly eclipsed, how much so you may judge by turning the picture upside down and fixating it. What you now see is a fairly good example of a visual field.

Figure 13 exemplifies the same thing. The surfaces in the perception appear to slant, recede, and lie behind one another in a space of three dimensions, although the patchwork of light and dark areas is two-dimensional.

The Visual Field during Movement of the Observer.

It has been emphasized that in the ordinary vision of everyday life any long continued fixation of the eyes is a rarity. It is equally rare to perceive the environment with the head motionless. If the observer is not involved in some kind of locomotion he is at least moving his head from time to time as he changes his posture. To remain motionless for any length of time is a difficult and unnatural achievement. How does this influence visual perception?

Every movement of the head produces a *deformation* of the visual field. This effect is not a sweeping shift such as occurs when the eyes alone move, but is rather a change in the pattern of projected shapes, somewhat analogous to the shifts and distortions of one's image in amusement-park mirrors. If you fixate a nearby object with one eye and move your head from side to side you can observe the way in which the edges in your field move across the surfaces behind them. The superposition of one object on another is unmistakable. If you stand up and walk from side to side, the projected shapes of objects are transformed, as we noted earlier. These are actually only incomplete descriptions of a much more general phenomenon which we will discuss later (Chapter 7). But they serve to suggest that the visual field is ordinarily alive with motion. This motion is not the absolute displacement which goes with eye movements (that is generally invisible in any case), but the kind of relative displacement which goes with head movements.

It is hardly necessary to point out that the visual world is not distorted in any such fashion as this when we move about in the environment. We have already noted that objects remain constant despite changes in the observer's viewing position. It now becomes evident that visual

FIGURE 12. Objects Seen Behind Others

FIGURE 13. Surfaces in Three Dimensions

space in general remains equally constant when we move about. If it did not the driver of an automobile would face a very strange situation.

The Awareness of Distance

One more characteristic of the visual field should be noted. It is never flat, like a surface on which a picture is painted or projected; that is, it is never wholly depthless. Nor is it lacking in the character of being *outside* of us, in externality. Nevertheless, it has *less* of these qualities than the visual world. The depth of the visual world is ordinarily just as visible as its breadth and its height. But you can reduce the depth somewhat by closing one eye. You can reduce it still more by fixating a point and maintaining prolonged fixation. It is lessened further if you then attend not to that point but to the hazy margins of the field and the pattern of shapes there. It is also reduced by tricks such as looking at the environment under your arms or between your legs, so as to invert the field. The impression of distance never quite vanishes, but the facts suggest that you *might* be able to see a depthless field if you had enough practice. Clear and indubitable distance is a characteristic only of the visual world.

Summary

What we have called the *pictorial* quality of the visual field has now been described in a number of ways. The field differs from a literal picture in some very important respects, of course, but the term will serve for purposes of description. Pictorial seeing, then, differs astonishingly from ordinary objective seeing. The field is bounded whereas the world is not. The field can change in its direction-from-here but the world does not. The field is oriented with reference to its margins, the world with reference to gravity. The field is a scene in perspective while the world is Euclidean. Objects in the world have depth-shape and are seen behind one another while the forms in the field approximate being depthless. In the field, these shapes are deformed during locomotion, as is the whole field itself, whereas in the world everything remains constant and it is the observer who moves.

It has the ring of familiarity to say that the field is *sensed* whereas the world is *perceived.* These terms, however, imply the traditional theory examined in the last chapter. It is also plausible to say that although the visual field is *seen* the visual world is only *known.* But this also involves a doctrine of perception which is debatable. The aim of this chapter is to describe the facts, not to explain them. Descriptively, the visual field always seems a little illusory. There is always the sense that one can bring back the world whenever one wishes. There can surely be agreement that the visual world is marvelously well adapted to be the conscious accompaniment of behavior, while the field is not. If we adjusted our actions to some of the peculiarities of the visual field, we should go badly astray; thus when, because of fog or darkness, the environment is not seen as a visual world but only as some kind of a vague visual field, we proceed cautiously.

The reader who is acquainted with

psychological theory will realize that the distinction between the visual field and the visual world is a *substitute* for the traditional distinction between visual sensation and visual perception. Is it not possible to relinquish the latter distinction with all its theoretical implications in favor of a description of our experience-when-we-introspect, that is, the field, and our experience-when-we-do-not, that is, the world?

The Problem of the Visual World

The task of stating generally what we see when we say that we perceive the environment has turned out to be neither short nor simple. But this lengthy exercise in introspection has served a purpose. It leads to a better understanding of the problem with which we started. The problem can now be put this way: How can we account for the perception of the visual *world?* For no theory of anything less than the visual world will be complete. The visual field, as the next chapter will show, is a reasonably close correlate of the retinal image. Therefore, the explanation of pictorial seeing is possible on traditional lines. The theories of vision, generally speaking, have been theories of the visual field, but this type of explanation is insufficient. What is required is a theory of *objective* seeing.

The conception of a clear and accurate visual world as the end-product of perception is unorthodox. The science of vision, almost from its beginning, has emphasized the errors and inadequacies of vision whereas this conception of the visual world has emphasized just the opposite. It may strike the reader as naive to assume that visual perception corresponds to its object when everybody knows how misleading perception can sometimes be. We may not legitimately *assume* the correspondence of perceptions to physical objects: that would indeed be naive. But on the other hand, we may and should consider what correspondence there is, for this is what needs explanation. The discrepancies between percepts and objects are not difficult to understand; what we need to understand is why there are so *few* discrepancies. That is the real mystery and the really important problem.

4

The Formation of Retinal Images

The Sequence of Events in Vision . . . . The Stimulus Variables for Vision . . . . Copies and Correlates on the Retina . . . . The Retinal Image and the Excitation of the Retinal Mosaic . . . . The Retinal Excitation as an Anatomical Pattern and as an Ordinal Pattern . . . . Visual Experience and the Retinal Pattern of Excitation

If we are to understand visual perception we must begin where perception begins: with physical objects, light, and the eye. There is no doubt about the fact that all vision, both the pictorial kind and the objective kind, is dependent on light rays and on the formation of images within the eyes. The discovery of how light behaves and how images are produced – the laws of optics – is one of the most brilliant chapters in the history of science. A culminating accomplishment of these discoveries was the publication of a famous treatise on physiological optics in 1866 by Hermann Ludwig Ferdinand von Helmholtz. The intricate and precise series of events in the physical world and in the human eye which make seeing possible were so accurately described by Helmholtz that there have been few men who could make any important additions to our knowledge about them for 80 years. As a consequence, nearly everybody has some idea of the nature of the optical process or at least a comfortable feeling that it is known by experts.

The popular idea of the optical process is that a picture is formed on the retina of each eye. Everybody knows what a picture is; hardly anything could be more familiar. It is therefore easy to rest content with no more of an explanation than that, or simply to assume that the retinal picture is transmitted to the mind. The fallacy of this explanation for perception, if not already evident, will become clear later on. Rather than examine it now, it would be more useful first to examine the way in which the retinal picture is produced.

The Sequence of Events in Vision

How is the material environment projected as an image, and how can this image enable us to see? The mechanisms by which animals obtain an image are ex-

tremely useful since images enable them to discriminate among objects at a distance instead of only those objects in contact with the body. By means of images things can be approached or avoided and the organism can get about in the environment without collisions. The basic facts of physiological optics are fairly complex. An effort will be made to simplify the technical explanations, however, and to take a fresh approach to the facts. Only those facts will be considered which contribute to an understanding of perception. Treating the problem in this way, the sequence of events can be divided into a number of stages, and we shall take these up one by one.

1. *The Array of Physical Surfaces.* The material world, as we all know, is made up of solids, liquids, and gases. In actual fact, a great part of it consists of earth, water, and air. The first two of these – but not the last – possess *surfaces.* The most common surfaces are those between solids and air. Of secondary importance are those between water and air, and these surfaces, incidentally, are almost invariably horizontal. Surfaces are extremely important for our perception of the world because obviously they determine what we know as objects or things. The surfaces of objects reflect light, if they are illuminated, and this fact is the original basis for visual perception. Generally speaking, air does not reflect light but transmits it; most objects reflect light but do not transmit it.

2. *Differential Reflection of Light from Surfaces.* The surfaces of the material world differ with respect to their structure and composition, both physically and chemically. Depending on how the object is put together (of cells, crystals, and so on) and what it is made of (its chemical substance) it will reflect more or less of the light falling on it. and it will also reflect relatively more of one wave-length or more of another. This differential reflection is the physical fact referred to when we speak of surfaces as having brightness and color. In addition to these simple differences in reflectivity there are a great many other complex differences produced by the structure of a surface. We have names for these denoting the sensory quality but not the physical character of the surface – such names as, shiny, rough, textured, and pebbled. One thing is certain – that a particular kind of surface reflects light in a particular kind of way.

3. *Transmission of Light to the Eye.* The light reflected from the surfaces of the world radiates freely through air but not through other surfaces, most of which are, as we say, opaque. The light can be considered *analytically* to consist of rays which travel in straight lines. Any given point in the open air, therefore, will be the juncture of rays from every surface of the material world which is not eclipsed by another surface at that point. If an eye is stationed at such a point, light from a wide array of objects and surfaces will fall on the cornea and pass through the pupil, although this light is only a minute portion of the sum of all the light being radiated from the surfaces of the world. Only the rays intercepted by the eye are relevant to vision. For a pair of human eyes taken together, the array of surfaces represented in the incoming cone

of rays extends about 180° horizontally and 150° vertically. This is what we mean by the field of view. Actually, however, this state of affairs is only momentary. Any one such cone of rays gives place to another overlapping cone as the eyes move from one point of fixation to another. The comparison already suggested is that the eyes play over the environment like a searchlight, with the difference that they absorb light from the constant flux of rays in the air about them instead of emitting it.

4. *The Projection of the World as an Image.* The cone of light rays which pass through the pupil of the eye forms an image on its rearward surface, the retina. Behind this statement lies a long and complex story. The exact nature of an optical image and the way in which it occurs depends upon the properties of light rays and their refraction, or the bending of their paths, in transparent substances. The forepart of the eye is an exceptional kind of solid substance which transmits the light which falls on it instead of reflecting it. The behavior of light passing through such substances, when they are of certain regular shapes called lenses, is to produce a convergence, instead of a continued radiation,

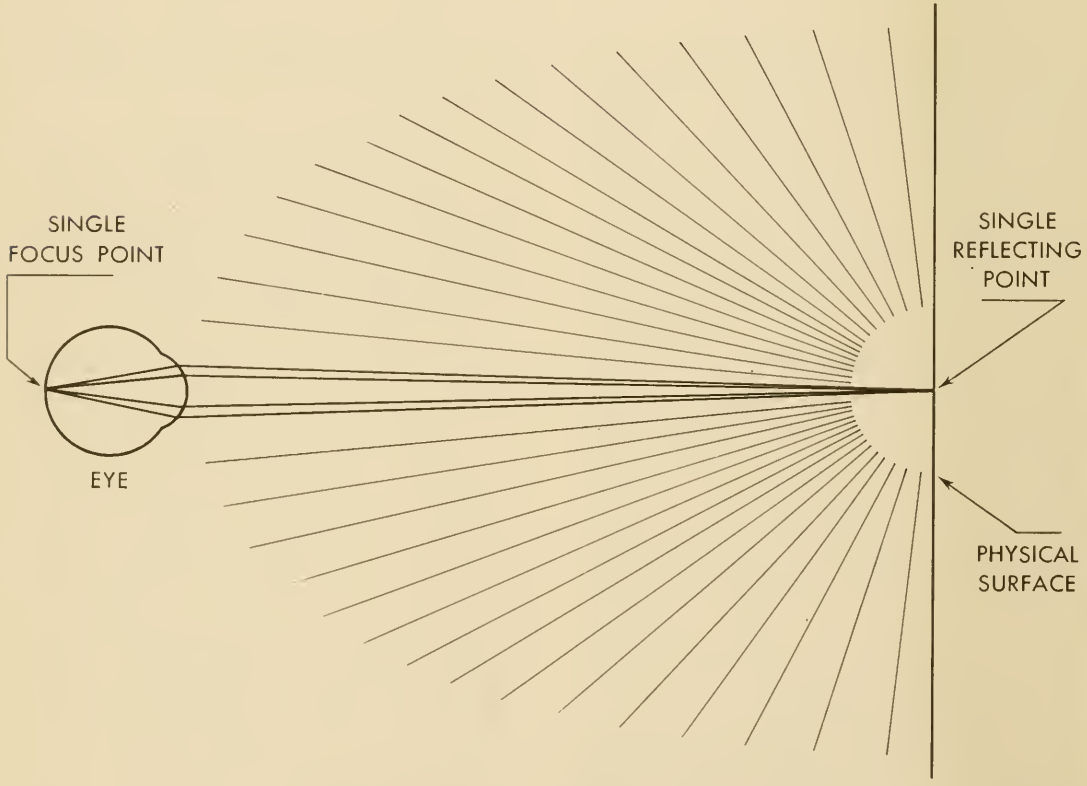


FIGURE 14. The Theoretical Reflecting Point and its Corresponding Focus Point

of the light. Every point from which rays originate is then represented by a corresponding point behind the lens called the point of focus. As a result, there exists a correspondence between reflecting points and focus points, each to each, such that the character of the light reflected at each external point is duplicated at the corresponding focus point, as Figure 14 illustrates. The rays of light which pass into the eye from a single point constitute what is known as a focused pencil of light. In theory there are an infinite number of reflecting points on a given surface, and the same is true of the corresponding focus points in the image of that surface. The total of all these focus points is the image. The cornea and lens of the eye have been shown to produce a very satisfactory image as thus defined.

The proof that light can be considered, for image-forming purposes, to consist of rays is given by the fundamental experiment of the pinhole camera. If light is made to pass through a very small hole it behaves like straight lines intersecting at a point. This is what makes light projective, in the geometrical sense of the term, and defines a projective correspondence. The essential fact about the optical image, however, is that it is a

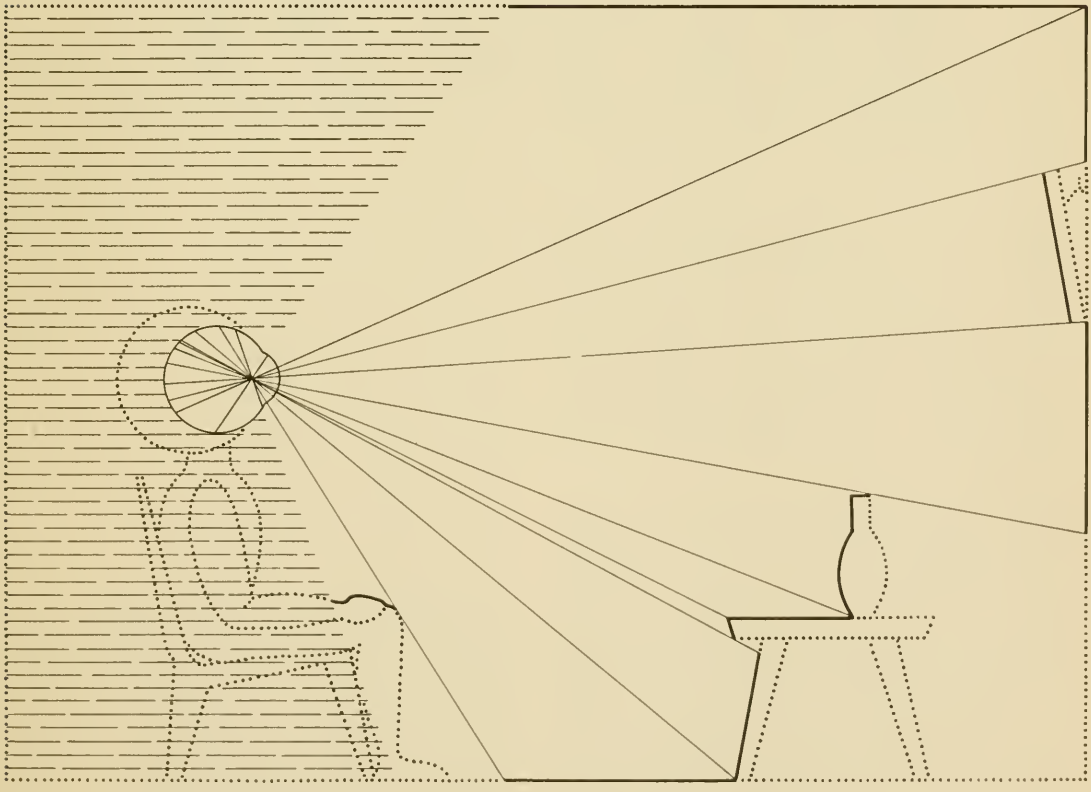


FIGURE 15. The Optical Projection of a Room

geometrical projection of the array of surfaces whose reflected light reaches the pupil. An illustration of this projection is given in Figure 15. The observer is looking at the wall of the room in front of him, and his visual field includes part of his body at one extreme and part of the ceiling at the other. The surfaces from which light reaches the eye are drawn in solid lines; all others are given in dotted lines. The environment outside the visual field is indicated by shading. The straight lines from the surfaces to the retina represent focused pencils of light which form the image on the retina. The only rays of light actually shown in the drawing are those from the edges of physical surfaces which, it will be argued, have a special significance for perception. The areas of physical surfaces correspond to areas of the image. Not all the surfaces of the world at different distances will be in perfect focus at the same time, it is true, but with a normal eye there is very little blurring. Although the retinal image is inverted and the order within the image therefore reversed, it nevertheless corresponds to the physical world as a projection. The assumption here (and throughout this book) is that for certain purposes we may treat the retinal image as if it were a two-dimensional pinhole image. It is important to note that this is not the kind of image defined by physical optics and used in the design of optical instruments, for this latter is three-dimensional.

The formation of an image on the retina can be observed directly. If the excised eye of an albino rabbit is fixed into a hole in a card and pointed toward a scene, by holding it in front of one's own eye, one can actually see the inverted image on the curved rearward surface, looking something like a miniature photographic transparency. It is this demonstration which has led to the theory that the retinal image is a "picture."

5. *The Mosaic of Retinal Elements.* The surface of the retina on which the image is projected is composed principally of extremely minute cells which contain photosensitive substances. Like the substances used in photographic emulsions, these are capable of reacting differentially to the energy and wave length of light. They are superior to any photographic emulsion, however, since they are self-renewing and capable therefore of registering the image continuously. Although the television camera can register an image continuously and in this respect is more like the retina, its mechanism is quite different. The cells of the retina are of two types, rods and cones. The cells are distributed much more thickly in the center of the retina than in the periphery, grading off from a density of 160,000 per square millimeter at the fovea to a much sparser distribution outward from the fovea.[1] Both rods and cones are wholly absent in a small peripheral area where the nerve bundle has its exit from the eye.

[1]The distribution of rods is quite different from the distribution of cones. The fovea contains only cones whereas in the peripheral retina rods predominate over cones. The decrease in visual acuity toward the periphery is probably related to the decrease in density of receptor cells and to the fact that the foveal cones have individual neurons while the peripheral rods are grouped with a single neuron. Even at the outer edge of the retina there appear to be more than a thousand cells per square millimeter.

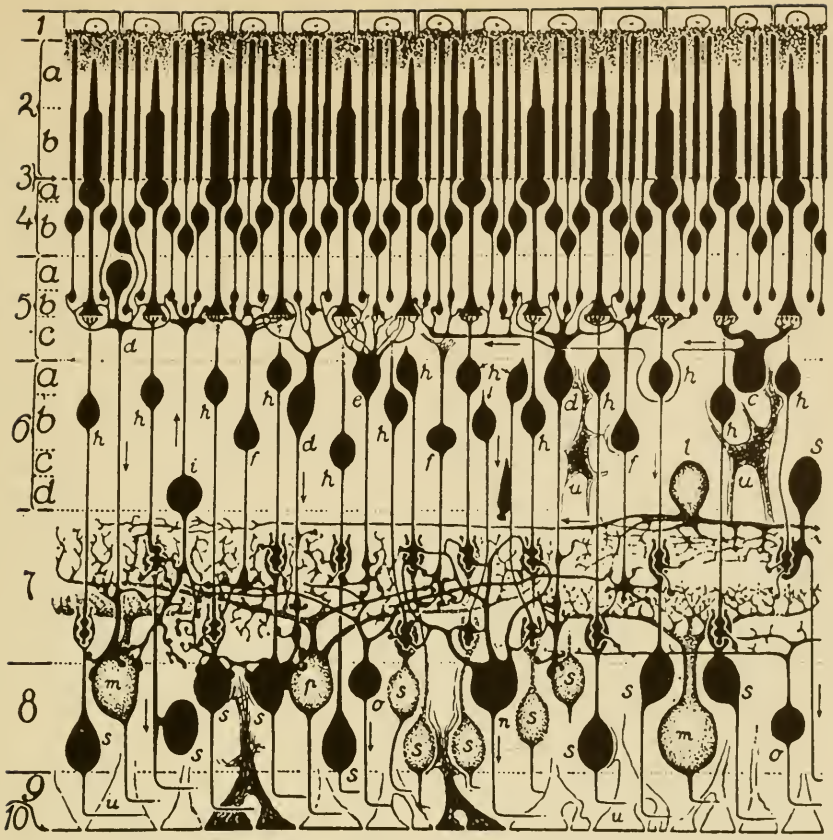


FIGURE 16. Scheme of the Retina

Note the different types of cells (neurons) and the variety of relationships and interconnections of their tips (synapses). The lower edge of the figure represents the layers lying toward the interior of the eyeball; the top edge shows the long, narrow *rods* and the shorter, thicker *cones*, which point outward away from the interior of the eyeball. Light admitted into the eye through pupil, lens, etc., has to pass through layers *10*, *9*, *8* etc., before falling upon the sensitive tips of rods and cones in layers *2a* and *1*. *(From S. L. Polyak,* The Retina, *University of Chicago Press, 1941. By permission of the publishers.)*

The nature of their differential reactions to light has been studied for many years without, until recently, any close approach to an understanding of it. The type of photo-chemical reaction which corresponds to wave length is particularly puzzling, as the various theories of color vision bear witness. One fact is certain, however, that the elements making up the retinal mosaic do react specifically to the character of the light focused on them, and therefore indirectly to the character of the surfaces from which the light was reflected. Another fact is equally certain, that they are connected with the individual fibers of the optic nerve, although not in a perfectly one-to-one fashion.

6. *The Anatomy of the Optic Nerve.* The rodlike and conelike cells of the retina, when they are stimulated by light,

initiate nerve impulses in the neurons which make up the sheaf of fibers, a quarter of a million or so in number, which we call the optic nerve. So far as the evidence goes, these nerve impulses are excited independently of one another and travel their paths separately. Very little more than this is known about them. The anatomical connections of the nerve fibers can, it is true, be traced. Some of them connect with centers in the brain governing the movements of the eyes, the regulation of the size of the pupil, and the accommodation of the lens. By far the largest part, however, connect with an area on the surface of the occipital lobes of the brain. The excitation of this cortical area is probably essential to all vision in man, for destruction of it produces blindness just as much as would injury to the eye or severing of the optic nerve. It has been assumed that the connection of the retina with this visual area was an exact point-to-point relationship and it was possible to infer that therefore the retinal image was projected on the brain in the same way that the physical world is projected on the retina. But the anatomical facts provide only a puzzling and very incomplete support for this assumption (6). Lateral connections exist among adjacent fibers both in the retina itself and at later stages in the tract between retina and brain. The amount of overlap is such that the "image" on the brain (if there were such a thing) would be very blurred. In all probability, it should not be thought of as an image, and even less, as a literal picture. It is an event composed not of light, but of nerve-cell discharges, and if a surgeon exposed the brain to view, there would be nothing to see.

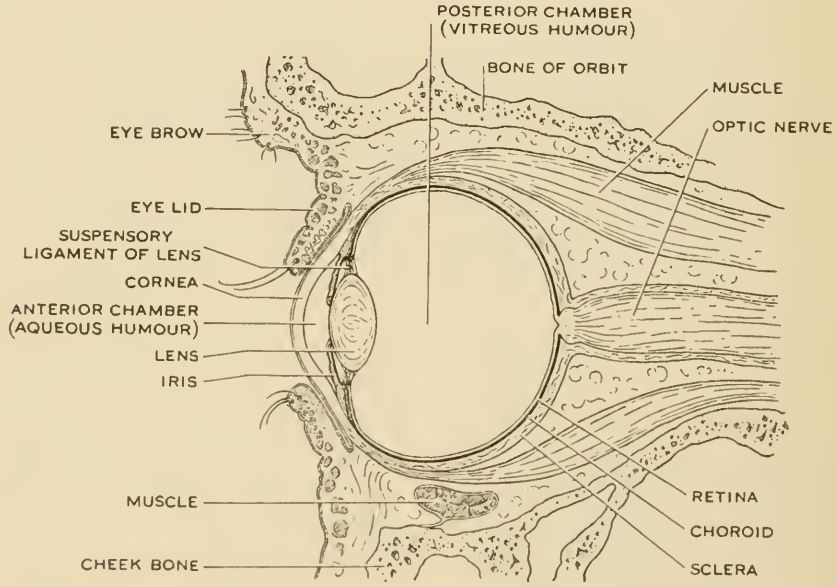


FIGURE 17. A Diagram of the Human Eye

(After H. W. Haggard, Man and His Body *(Harper & Brothers, 1938)*

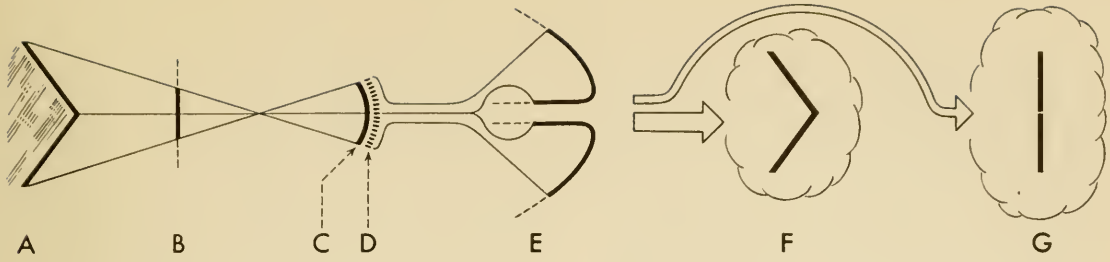


FIGURE 18. The Sequence of Transformations in the Process of Visual Perception

A. The physical environment: a wedge-shaped physical object, reflecting light. B. A "picture" of the physical environment: a plane projection of the light reflected from the physical object. C. The retinal image (the proximal stimulus for vision): a curved projection of the light reflected from the physical object. D. The pattern of excitation: a mosaic of photosensitive receptors. E. The brain process: a bifurcated and oddly-shaped projection of excitations on the rear surfaces of the hemispheres. F. The visual world, or phenomenal experience: the experience of a wedge-shaped object. G. The visual field, or the color-sensations obtained by introspection: the impression of two flat patches of color adjacent to one another.

Sensitive electrodes placed upon the cortex might pick up regions of high and low activity with gradients or contours between them, but this observation has not been conclusively made. The fact is that no one yet has an adequate conception of it.

7. *The Unknown Activity Producing Vision.* All that is certain about the last stage of visual perception can be put into a few words. There are unquestionably neural processes at the occipital surface of the brain. These processes arouse still others. They are almost unknown.[2] Nevertheless, these unknown events are the sole basis of our visual experience of the world. This experience is both elaborate and exact. So this discussion has come around again to the same problem with which it started.

The Stimulus Variables for Vision

Fortunately, it is not necessary to understand the events within the nervous system (stages 6 and 7) in order to be able to make a scientific attack on the problem of perception. One can by-pass the nervous system and jump from the retinal image directly to the perceptual experience. One can, in other words, seek to establish an empirical corres-

[2] There are, to be sure, a number of established facts about the processes within the brain which correlate with visual perception, and more are continually being established. The facts are, however, as yet puzzling and incoherent, or at least they seem so to the writer. The study of visual brain-processes is being pursued by Kohler (69) and others (66, 50). The possibility of reaching principles of explanation at the psychophysical level without knowing the principles at the psychophysiological level has already been referred to in Chapter I.

The processes referred to above may not even be wholly cortical. Certain mammals are known to have some residual sub-cortical vision persisting after the extirpation of the whole occipital cortex. Whether the same condition is true in any degree for man is not yet certain.

pondence between the stimulus and its conscious resultant. This is what psychologists have been doing with color stimuli, sound stimuli, and others for more than a hundred years, and the results of this endeavor have produced the most securely established body of facts known to psychology. It is traditionally known as psychophysics. A retinal spot of light having a wave length of 760 millionths of a millimeter yields a visual spot of the quality red; an airwave of 256 double vibrations per second in the ear produces the quality of middle C. Experiences generally have a specific relationship to the stimuli which arouse them, as indeed they must if the experiencer is to adapt his behavior to his environment. This is the principle of psychophysical correspondence. Little as we may comprehend its physiological and nervous basis, the rule is that variations in stimuli are co-ordered with variations in the character of the perceptions. Musical tones, for example, are related to the frequency of air vibrations in much the same way that the letters of the alphabet are related to the number series from 1 to 26. This rule has never failed of verification for stimuli which can be ordered in physical terms, or in other words, for stimulus *variables*. Let us now analyse the retinal image to see what kinds of stimulus variables are included in it.

The classical stimulus variables for vision are, of course, the physical variations of light itself. These are its wave length or frequency and its energy or intensity. By combinations of these, and by mixtures of wave lengths, all the experienced qualities of pure color and brightness – of color as such – can be accounted for. If our environment consisted of nothing more than a homogeneous sea of light, without surfaces or objects, then all our visual experiences could be specified in terms of these variables with nothing left over. Each retinal point would be stimulated in the same way as every other retinal point. But obviously our visual world consists of more than this. The stimulus situation for a typical environment is diagrammed in Figure 15. Not merely colors but surfaces and edges are projected in the retinal image. There must exist, therefore, a second type of stimulus variable in the image. The locus of the classical stimulus variables is the single spot of light, since each focused pencil of light may possess its own unique combination of wave lengths and intensity. But surfaces and edges are not related to this kind of variation *within* the spots of the image; they are related instead to variations *among* the different spots of the image.

The facts of the situation are represented in Figure 15. The array of physical surfaces whose reflecting points are duplicated in the image is shown in solid lines. As projected, these surfaces border on one another. In other words the edges of these surfaces correspond to abrupt changes in the energy and wave length character of the light spots composing the image. For example, if the floor is dark brown and the table is light gray, the color stimulation in the image will shift accordingly along the margin between the two parts of the image. The image, therefore, is made up of areas of different light-character, and it is the transitions between these areas which give rise to visual lines

or contours. In the illustration, the pencils of light from an edge to its corresponding border in the image are represented by straight lines.

Not only the edges, however, but also the areas of the physical surfaces have specific representation in the image. The physical irregularities of the surface, both its gross composition and also its minute structure (if the lens of the eye is accommodated for that surface), are projected as correlated irregularities in the image. A tiny depression of the surface, for example, is focused as a dark spot, a slight protuberance as a minute high light, and an array of such surface-elements as an array of dark and light spots. This type of stimulation gives rise to what we will call the quality of visual *texture*. It seems probable from the evidence of Metzger (81) and others that texture is what makes a surface perceptible as a surface instead of as mere insubstantial areal color. It may be noted that physical things like clear sky, dense fog, and regions of complete darkness such as the mouth of a cave do not reflect light as a surface does, do not possess texture, and are not seen as surfaces.

A typical retinal image, then, contains two fundamental types of stimulus variation, one in the character of the focused light at any point and another in the relation of these light-points to one another. The first is the classical variation in stimulus quality and intensity; the second is variation in what has loosely been termed stimulus "distribution" or "pattern." An image is an *arrangement* of color-points, and it may vary either in the color of the points or in their arrangement. Terms like distribution, pattern, and arrangement are not very exact, it must be acknowledged, and an effort will be made in Chapter 5 to be more specific about this variable. It is evident, however, that the kind of arrangement we are talking about is simply that of adjacent order on the retinal mosaic. A transitional arrangement of color-points yields a line or contour. An alternating or scattered arrangement of color-points, so far as we know, yields a surface. An array of homogeneous color-points, all identical (which is not an arrangement at all), yields pure insubstantial color (61).

Copies and Correlates on the Retina

The foregoing short survey of optics points to a conclusion. The image is an arrangement of focused light on a physical surface of two dimensions which is specific to an array of reflected light from physical objects and surfaces in three dimensions. Since the reflected light is specific to the objects and surfaces themselves, the image is also specific to them. Geometrically, we say that the image is a *projection* of the world. The conclusion is that the image is *not* a replica of the world. If taken seriously this conclusion has far-reaching implications.

Unfortunately, the word "image" has more than one meaning. It may refer to an effigy or copy – the "graven image" of the Bible – or it may refer to the projected arrangement of light as just described – the image of physiological optics. The two meanings of the term are easily confused and there may be intermediate meanings between them. But the retinal

image is unquestionably a projection rather than a facsimile. Everybody knows that objects themselves do not get into the eye. Neither do small replicas of things get into the eye, although this was a reasonable theory of perception before the discovery of the nature of light. Nothing gets into the eye but radiant energy. Only because it is focused is it specifically related to the object. The object therefore does not have a *copy* in the image but a *correlate*. The fact is that the optical image does not have to be like its object to make vision possible.

The chief source of misunderstanding here is the assumption that the retinal image is a *picture.* It might be argued that even if the image is not a replica of the environment it is at least a representation of it. The apparent simplicity of this pictorial analogy for vision makes us reluctant to give it up, scientists as much as anybody else. But a picture as a representation of something is nothing if it is not presented to an eye. An unseen picture is only an arrangement of pigment spots, if it is a painting, or an arrangement of metallic grains of silver, if it is a photograph. It is simply a part of the material world which has to be seen, like anything else. If the retinal image were really a picture there would have to be another eye behind the eye with which to see it. The notion that we see our retinal images is based on some such idea as a little seer sitting in the brain and looking at them. The question which then arises is how *he* can see. The retinal image should not be thought of as a picture or a representation but as a physical arrangement on a two dimensional surface.

The correspondence between the world and the optical image need not be that between a thing and its copy; it need only be that between a material quality and its correlate. There is no counterpart in the image of that physico-chemical character which gives a surface its particular hue, but there is a correlate, wave length. There is not a counterpart in the image of the physical microstructure which gives a surface its texture but it does have a stimulus-correlate, as will be evident. Above all, there is no copy in the image for the shape of an object in three dimensions, or what has been called its depth-shape, and this is something which all genuine objects possess. Similarly there is no copy in the image of the solidity and distance of the environment in general, but there must be some correlates for these variables, or we could not see them. Finally, the size-relations of the objects in the environment and the interspaces between them, following as they do the laws of Euclid and not the laws of perspective, are not copied in the image. But for these and the other features of the world there must be some basis in stimulation, however complex. This basis remains to be discovered.

There could, theoretically, exist a material environment for which the retinal image *would* be almost a duplicate. It would consist of a large picture at right angles to the line of sight and filling the entire field of view. The hypothetical "picture plane" which is posited by the perspective draftsman at a fixed position in front of the eye would define such a

picture. It has been illustrated in Figure 7. The lines, areas, shapes, and sizes in the picture would then be duplicated in the image, the principal difference being that the image would be upside down relative to the picture. Such an environment, of course, never existed. But most of the experimental research on visual perception in psychological laboratories has been performed with stimulus objects of just this kind, drawn or exposed on a plane surface. And the assumption that the fundamental kind of seeing is "pictorial seeing," or the perception of a depthless world, is also consistent with the duplication theory of the retinal image. Perhaps these facts provide the explanation of its persistent hold on our thinking about the perceptual process.

The Retinal Image and the Excitation of the Retinal Mosaic

The step between the formation of an image on the retina and the excitation of the mosaic of rods and cones, stages 4 and 5, is one which must be kept in mind if our reasoning is not to go astray. It is easy to assume that the retinal image and the retinal excitation are the same thing. But the former, clearly, is a matter of physics while the latter is a matter of physiology. The image is an arrangement of light-points while the excitation is an arrangement of discharging nervous elements. These individual points of the image, it may be noted, together with the rays of light which explain the correspondence to the world, are pure geometrical fictions introduced for purposes of analysis, whereas the individual spots of the excitation-pattern are anatomical facts. The light composing the image of a uniform environment is equally dense over its whole area, being evenly distributed, whereas the pattern of excitation is most dense in the center of the retina where the cones are concentrated and least at the periphery where the cells are thinly distributed. Above all, since the image is an event in the light-flux of the physical world, it has reference to the world and is fixed in relation to it. It keeps a constant alignment with gravity, for instance, when the head is tilted and the retina rotated. The retinal excitation on the other hand, having its reference to the retina itself, is a pattern composed of retinal elements which remain the same only so long as the eye does not move. When the eye moves, the image is transposed on the surface of the retina and consequently there is a shift in the pattern of excited elements in the anatomical mosaic of cells.

The Retinal Excitation as an Anatomical Pattern and as an Ordinal Pattern

Another troublesome question now arises: how can the retinal image be transposed and still retain its equivalence as a stimulus for vision? The pattern of excitation, it must be remembered, is a set of units of finite size, correlative with but by no means exactly duplicating the image. This spot-pattern is about as close to the immediate basis of visual perception as our present knowledge will take us, so it needs to be examined with care. From one point of view it is composed of nerve cells in an *anatomical* relation to one another. But from another point of

view it is composed of nervous excitations in a purely *ordinal* relation to one another.

An analogy may be useful here. Everyone has seen electric signs, consisting of a mosaic of light bulbs wired to a corresponding set of contacts, of such a sort that any number of patterns can be lighted up on the same sign. The retina is more like this than like a photographic film, which is only good for one exposure. In some of these signs the patterns of lighted bulbs can be transposed on the mosaic so that words or figures will move across it. Now a stationary pattern can be defined by specifying the contacts of the bulbs which light up. This would be an anatomical pattern. But a moving pattern can best be defined by specifying not contacts but the *adjacent order* of contacts. This would be what was called above an ordinal pattern. In other words the pattern can legitimately be thought of as either an arrangement of bulbs or as an arrangement of lights without regard to bulbs. Likewise the pattern of excitation on the retina might be defined either in terms of the units of the mosaic or the units of excitation as such. The former pattern is embodied in units which have an anatomical meaning; the latter pattern is not embodied in this way but nevertheless it is definable in terms which have mathematical meaning. The units of excitation maintain a constant ordered relation to one another when the retinal image is transposed even though the anatomical units do not. The ordinal pattern, therefore, is preserved when the eye moves although the anatomical pattern undergoes a complete rearrangement. It seems possible that the organism can react to an ordinal pattern as well as to an anatomical pattern. The television camera can register a purely ordinal and transposable pattern – why not the eye of a living animal? An attempt to deal with the ordinal type of stimulation more exactly will be made in Chapter 5.

For the present, it is enough to emphasize the fact that the *identity* of a given point-stimulus in the eye depends not at all on the anatomical point of the retina stimulated but entirely on the position of that point relative to other points of stimulation. *A given spot of light in a given retinal image is the same spot at different instants of time when its position relative to the order of spots is determined, not when its position relative to the retina is determined.*

This fundamental fact, that a spot of light is a stimulus for perception by virtue of its ordinal location and not by virtue of its anatomical location, can be illustrated by experiment. Even in the case of two spots in an otherwise blank field of view, the principle holds. Let a, b, and c represent three separated retinal points in a line. If points a and b are stimulated briefly and a moment later points b and c are stimulated in the same way, an apparent movement will occur. What the observer sees is a pair of spots which move toward the right. The identity of the anatomical point b with itself is not sensed. The spot at a moves to b and the spot at b moves to c.

$$a^1 \qquad b^1{}_2 \qquad c_2$$

Points a^1 and b^1 flash on first. Points b_2 and c_2 flash on a moment later.

The relation of one spot on the right of another is what is sensed; the relation of either spot to the retina is not apprehended at all. This becomes wholly compelling when the number of spots is increased and when the "form" of the spots is easily seen. The experiment was carried out and elaborated by Ternus early in the program of the Gestalt theorists (104, translated in 32) and it illustrates their main contribution to the science of vision – the doctrine of the "transposable" Gestalt and the abstract conception that the "parts" of a perception exist only in relation to a "whole."

Visual Experience and the Retinal Pattern of Excitation

We can now consider once more the two kinds of seeing described in the previous chapter, the visual field and the visual world, and we can assess their correspondence or non-correspondence with the retinal stimulation. In several obvious ways the visual field corresponds with the anatomical pattern of excitation. It is finely differentiated at the center of clear vision and becomes progressively less determinate away from the center; this reflects the dense distribution of cells at the retinal center, the fovea, and their thinning out toward the periphery. It has boundaries which can be plotted; so also does the mosaic of rods and cones, and these boundaries agree. An object which passes out of the visual field corresponds to an object which ceases to project rays on sensitive elements. The field can easily be accounted for in these respects. But how about the visual world? It does not have boundaries and it is more nearly clear in all its parts. Only the suggestion of an explanation of these facts can be given here and the answer is therefore incomplete and tentative. We can be fairly certain, however, that the visual world is dependent on eye movements and is not seen as the result of a single fixation or a momentary visual field. It must correspond, therefore, to *successive* patterns of excitations on the retina, united perhaps by a kind of immediate memory. These patterns will overlap one another anatomically as the eye moves, and the basis for the visual world, therefore, must be what has been called the *ordinal* pattern of excitation rather than the *anatomical* pattern. If it be assumed that there *is* an ordinal pattern which keeps its integrity during eye movements, then it is possible for any part of it to be brought to the center of the anatomical mosaic and registered in fine detail. A complex of this sort, over time, would be both uniformly differentiated and unbounded, and might therefore provide a basis for the perception of the visual world (Chapter 8).

There is also the question why the visual field shifts during eye movements but the visual world does not. It was made clear in Chapter 3 that the visual fields before and after a change of fixation are different. The first array of color patches is by no means the same as the second. This fact is parallel to the rearrangement of the anatomical pattern, and implies that the field corresponds specifically to that pattern. The visual world, however, is not rearranged after a change of fixation; the objects which are seen appear to be the same before and after the eye movement. The stimulus complex to

which it corresponds must therefore be based on the ordinal pattern of excitation which is transposable over the retinal mosaic.

Will these hypotheses also explain the appearance of after-images? If a negative after-image is a patch of excitation in the *anatomical* pattern, that is, one which is fixed on the mosaic and not displaced when the eye moves, it should be seen in a fixed position with reference to the visual field. It should, in other words, appear wherever we turn our eyes, and this is what it does. If, moreover, the visual world is a correlate of the *ordinal* pattern, then the after-image should appear to vary its location with reference to the world. It should, in other words, be superposed on different object surfaces as the eye moves, and so it is. The patch of excitation corresponding to an after-image and the patch of excitation corresponding to a transposable object are both results of stimulation, but the difference between them is highly significant for the theory of vision. The former reminds us that stimulation of a mosaic must always necessarily have reference to the mosaic. The latter reminds us that stimulation of a mosaic must with equal necessity consist of a mathematical order.

For all these characteristics of the visual field and the visual world, the anatomical pattern and the ordinal pattern provide correlates. But there remains the most important of all the differences between the visual field and the visual world, namely the three dimensional character of the world. The psychological fact that the visual field tends to be made up of projected shapes rather than depth-shapes, and that it tends to have an appearance consistent with the laws of perspective rather than the laws of Euclid, is in agreement with the physical fact that the retinal stimulation is a projection. But how are the depth shapes themselves to be explained, and why does the horizon really look like the world at a very great distance and not like a line at which sizes vanish and where parallels meet? The eclipsing of projected shapes in the visual field is consistent with the projected character of the retinal stimulation, but what about the "seeing behind" impression which we get with objects in the world? How about the obvious but puzzling fact that the world, and the field too, are external? The stubborn fact is that we see and get around in a world which stretches from here to there and this fact remains in need of explanation. The effort to find such an explanation will be the principal concern of the next few chapters.

5

A Psychophysical Theory of Perception

Abstract Space and the World of the Flier . . . . Stimulus Correlates . . . . The Hypothesis of Ordinal Stimulation . . . . Retinal Gradients as Variables of Stimulation—Texture . . . . The Cues for Distance as Stimulation Gradients . . . . The Concept of Gradient . . . . The Concept of Psychophysical Correspondence . . . . Summary

At the beginning of World War II, the theoretical problem of space perception became a practical problem almost overnight. The skills of aviation began to be of vital interest to millions of individuals. The abstract question of how one can see a third dimension based only on a pair of retinal images extended in two dimensions became very concrete and important to the man who was required to *get about* in the third dimension. If the visual world of the airplane pilot were not in fairly close correspondence with the material world on which he had to land his airplane such as a carrier-deck, the practical consequences could be disastrous. The theories of space perception, therefore, became of more than academic interest in the rapidly developing field of aviation psychology.

But the fact was that all the evidence from the laboratories and all the theories of ingenious men had little practical application to the problem of flying. The theory of the binocular and the monocular cues for depth, perfected eighty years before by Helmholtz, could explain how a pilot might see one point as nearer than another point. But the pilot was not looking at points of color in a visual field; he was typically looking at the ground, the horizon, the landing field, the direction of his glide, not to mention several instruments, and visualizing a space of air and terrain in which he himself was moving – very fast and possibly in a cold sweat.

Abstract Space and the World of the Flier

The space in which the pilot flies is not the abstract space of theories, nor the lines and figures of the stereoscope, nor the space of the usual laboratory apparatus for studying depth perception. It does not consist of objects at varying empty distances. It consists chiefly of

one basic object, a continuous surface of fundamental importance – the ground. A pilot who cannot see the ground or sea is apt to lose touch with reality in his flying. A visual field of blue sky, or fog, or total darkness yields an indeterminate space which is the nearest thing to no space at all. Only a substitute for the ground and its horizon in the form of instruments will permit him to maintain the level flight of the airplane under such conditions and to proceed from one place to another. The spatial situation which needs to be analysed, therefore, must involve the ground and everything that it implies. Instead of calling it a space it would be better to call it a world.

The conception of an empty space of three dimensions was a conception of philosophers and physicists. It was appropriate for the analysis of the abstract world of events defined by Newton. It was and still is of enormous value for analysis in the physical sciences. But the fact that it simplifies such problems does not make it the best starting point for the problem of visual perception. Space, time, points, and instants are useful terms, but not the terms with which to start the analysis of how we see, for no one has ever seen them.[1]

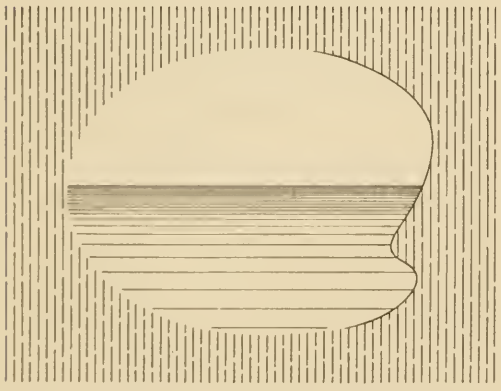


FIGURE 19. The Typical Visual Field of One Eye

The world with a ground under it – the visual world of surfaces and edges – is not only the kind of world in which the pilot flies; it is the prototype of the world in which we all live. In it, one can stand and move about. It conditions and provides support for motor activity. A ground is necessary for bodily equilibrium and posture, for kinesthesis and locomotion, and indirectly for all behavior which depends on these adjustments.

An out-of-doors world is one in which the lower portion of the visual field (corresponding to the upper portion of each retinal image) is invariably filled by a projection of the terrain. The upper portion of the visual field is usually filled with a projection of the sky. Between the upper and lower portions is the skyline, high or low as the observer looks down or up, but always cutting the normal visual field in a horizontal section. This is the kind of world in which our primitive ancestors lived. It was also the environ-

[1]The theories of space-perception which flourished in the 19th century were all theories of abstract, empty space. The experiments concerned lines and points in an indeterminate visual field, as seen in a stereoscope or a depth-perception apparatus. The theories and the experiments alike may be characterized as *geometrical*. They were great intellectual achievements (the theories of the "horopter" are an example), but they will not be considered at present. The theory of disparate retinal images as it applies to surfaces rather than abstract points will be restated in the next chapter, and the conception of geometrical space will be treated in Chapter 10.

ment in which took place the evolution of visual perception in *their* ancestors. During the millions of years in which some unknown animal species evolved into our human species, land and sky were the constant visual stimuli to which the eyes and brain responded. In the typical *indoors* world of civilized man, a ceiling and walls take the place of the horizon and sky, but the floor is still an equivalent to the ground. This basic surface is the background for the objects to which we normally give attention and, as we learned in Chapter 3, its horizontal axis is implicit in every visual field whatever the posture of the body may be. It is little noticed, but on the average and over the ages it must have determined the fundamental pattern of retinal images for all or most terrestrial animals.[2]

The classical theories of space perception conceived the third dimension to be a line extending outward from the eye. Space was therefore empty between the eye and the object fixated. The perceived distance of this object seemed to be what needed explanation, and the explanation was sought in the consequences of the possession of two eyes. It would have been better to seek an explanation of the sensory continuum of distance as such which, once visible, determines how distant *all* the objects within it are. But this explanation was impossible so long as the continuum of distance was conceived as the third dimension. The solution of the difficulty is to recognize that the continuum of distance depends on a determinate *surface* which extends away from the observer in the third dimension. Such a surface is projected as an image which is spread out on the retina, not confined to a point.

Figure 20 illustrates the two formulations of the problem. The points $A,B,C,$ and D are not discriminable on the retina. Distance along this line may be a fact of geometry but it is not one of vision. The points $W,X,Y,$ and Z at corresponding distances *are* discriminable on the retina. They represent the image of an extended surface, the points being, for example, highlights on the surface. It may be noted that the retinal spots become progressively closer together as the distance increases. What kind of a theory, we may now ask, is implied by this latter formulation of the problem? How is a surface seen?

Stimulus Correlates

The first place to look for an explanation is obviously the retinal image. If, contrary to past teaching, there *are* exact concomitant variations in the image for the important features of the visual world a psychophysical theory will be possible. The image, according to the evidence in Chapter 4, is a good correlate (but not a copy) of the physical environment. It may also prove to be a good correlate of perception, despite an entrenched opinion to the contrary. The retinal image, it is true, is not much to look at when one

[2] So universally is the ground taken to be the background of objects that the mere location of one patch of color *above* another in the visual field tends to make it appear more distant. Height in the visual field can be a genuine clue to distance only if upright posture, a level ground, and a tendency for objects to be *on* the ground are assumed. This point will come up in later chapters.

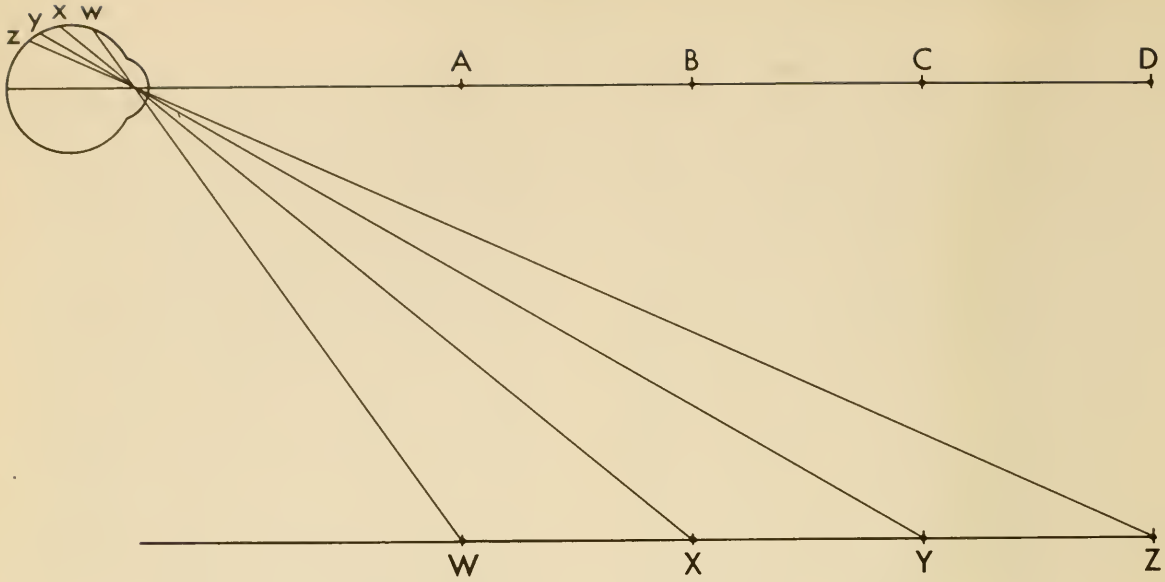


FIGURE 20. Two Formulations of the Problem of Distance Perception

compares it with the elaborate reality of the visual world, but the fact is that it is not something to be *looked at;* it is a *stimulus*. The question is not how much it *resembles* the visual world but whether it contains enough variations to account for all the features of the visual world.

If we can analyse the retinal image for its stimulus variations, we shall open up the possibility of experimental control of these variations. Given a means of producing them, an experimenter, and an observer, it can be determined whether the variations are or are not in psychophysical correspondence with the observer's perceptions. This is the method by which the sensory capacities, so called, of men and animals have been determined. The test is simple: does a specific variation in the observer's experience (or behavior) correspond to a variation of the physical stimulus? Although this experiment has seldom been applied to what are traditionally called perceptions, it can and should be performed (40).[3]

It is, of course, a departure from tradition to conceive that a surface, or an outline, or the depth of a surface should have a specific stimulus. The stimulus for vision, we are accustomed to think, is simply light energy. But such a definition reflects the preoccupation of nearly a hundred years of research on color vision and light-dark discrimination, the outcome of which still leaves us ignorant of the vision employed in everyday living. The higher animals do not simply react to the

[3]Students of psychology will recall, in this connection, that the Gestalt theory denied any one-to-one correspondence between the stimulation of receptors and the experience which resulted. The assumption of such a fixed correspondence was called the "constancy hypothesis." It seemed to be untenable since everyday visual experience was so demonstrably unlike its retinal image. The aim of this chapter is to reassert the constancy hypothesis on the basis of a broader conception of stimulation.

direction of light as a plant does; they have a specialized neural surface, the retina, on which their environment is projected by means of focused light. As a result they can react *indirectly* to the environment itself, and the point would be missed by insisting that they are actually reacting only to light. In what respects is this *projected* light a stimulus?

The Hypothesis of Ordinal Stimulation

Before attempting to answer this question it would be useful to agree on just what the term "stimulus" is to mean, for it is a much misused word. Let us assume that a stimulus is a type of variable physical energy falling within certain ranges of variation (the limits being called absolute thresholds) which excites a receptor or a set of receptors differentially. If it does not release physiological activity in a receptor-mechanism it is not a stimulus. As the energy varies successively, the excitation varies concomitantly in some specific way. This is a strict definition of a stimulus. For our present purposes, as applied to the retina, we wish to extend the term to mean also a *simultaneous* variation over the set of receptors, or a differential excitation of different receptors, and the order of such a variation. For the extended meaning the term *ordinal stimulation* will be used. "Ordinal" simply refers to order or succession. This is what has usually, but inaccurately, been called pattern stimulation.

In this book, the term "stimulus" will always refer to the light energy on the retina, never to the object from which light is reflected. The term "stimulus-object" will never be used, since it can serve as a cloak for ignorance. The distinction of Heider (51) and Koffka (67) between the "distant" stimulus (the object) and the "proximal" stimulus (the image) is illuminating just because it implies, and just so long as it implies, that the latter stimulates the organism. The term "stimulus situation" likewise will never be used since the situation does not exist in the retina any more than the object does, and the question is how both are seen.

How can we specify the *order* of visual stimulation? The retinal image as a physical event may be treated as an infinite series of geometrical points or as a finite number of minute areas of arbitrary size. The latter is the more useful assumption for our purpose. In either case the image can vary in two fundamental ways: first in the character or "color" of the focused light at a given spot, and second in the distribution of these spots, or their geometrical relations to one another. This second variable is the one that makes all the difficulty. It is not easy to deal with the complexities of distribution or arrangement in a mathematically precise way. Nevertheless this variable is the one on which everyday perception depends. Let us assume, as a start, that organisms can react specifically to the *order* of the light-spots as well as to the character of the light in each spot. How an organism can do so, we do not know – that is another question. But if it can react differently to a spot-sequence such as "black-gray-white" from the way it does to the sequence "white-gray-black," then the order is the effective fact

and it would be legitimate to invent for it the term *ordinal stimulation*.[4]

The spots or elements of the image, in this analysis, are to be understood as arbitrary units of area. Like the points of geometry and optics they have only a logical existence. They are necessary to a mathematical treatment of retinal stimulation, but they are not to be considered the elements of visual experience, or the sensory units of perception. They are analytical fictions, and they do not add up to a visual field any more than the geometrical points employed in the operations of a surveyor add up to the surface of the earth which he surveys. It should be emphasized that the fundamental variations in light energy and in order or arrangement which constitute the retinal image are both abstractions.

The prevailing assumption about pattern vision has always been that the ocular mechanism enables the organism to respond to a specific set of *ray-directions* (112, p. 85). Each hypothetical light ray was supposed to be an individual stimulus. It has been argued, however, that light rays are analytic fictions. Furthermore, the homogeneous total field experiment demonstrates that when every ray has the same wave length and intensity there is no perception, and this experiment implies that the organism *cannot* respond to ray-directions as such. What the retina does respond to is a differential intensity in adjacent order over the retina. The necessary condition for pattern vision is an *inhomogeneity* of the set of hypothetical rays, not the rays themselves. The ray-direction theory of the stimulus, the point-theory of objective space, and the local-sign theory of subjective space all collapse together if this implication is correct and require a thorough reformulation.

Considering the retinal image as an array of small adjacent areas of different radiant energy, let us try to state the kinds of order into which the elements might fall. For the sake of simplicity we may consider a hypothetical case in which there are only two levels of light intensity in the image and no differences of wave-length. An element may be relatively "light" or "dark," but nothing more. If the former it may be indicated by the letter *l* and if the latter by *d*. The simplest of all orders would then be *lllllll* or *ddddddd*. All the elements of the order are the same. This is what Koffka has called homogeneous retinal stimulation (67, p. 110). In a two dimensional array it is the stimulus correlate of visual fields like the sky, absolute darkness, or the "film-colors." The experience is one

[4] As will appear in a later chapter, there is evidence that organisms can react specifically to a *successive* order of stimulation of the same spot as well as to an *adjacent* order of stimulation of different spots. Both kinds of order are present in retinal stimulation. The *successive* order "black-gray-white" yields a lightening effect and the reverse a darkening effect; the *adjacent* order yields a patterning effect. A co-variance of successive and adjacent order seems to be the essential condition for visual motion. For the present we are concerned only with adjacent order.

The term *order* is often used by philosophers and artists in a very inclusive sense. It may mean form, pattern, arrangement, position, direction, and even magnitude or distance. The term is here used, however, in an exact and literal sense to refer to that characteristic which numbers have of making a sequence which is not the same in one direction as in the other.

of pure areal color, seen at an indeterminate distance.

A second type of order would be one containing a single step or jump, such as *lllldddd* or *ddddllll*. This order may occur along one or both dimensions of an array of elements and, when it does, what we call lines or discontinuous areas will appear in the visual field. Presumably this type of order is a stimulus correlate

llll dddd dddd llll llll llll dddd dddd
llll dddd dddd llll llll llll dddd dddd
llll dddd dddd llll dddd dddd llll llll
llll dddd dddd llll dddd dddd llll llll

for the margins or outlines which are the necessary conditions for seeing figures and shapes.[5]

A third type of order would be one similar to *llddllddl*, which contains a cyclical or alternating change. It is a reasonable hypothesis that when such an order is found in both dimensions of an array of elements there will occur the visual quality of texture, and that this is the stimulus correlate of a visual *surface*. The varieties of texture in experience are innumerable, of course, but the varieties of a cyclical order of elements could be equally enormous in number. With only the two elements *l* and *d*, there are many repetitive sequences possible; when all the levels of intensity and wave length are taken into account the variety of cycles become incalculable. The assumption is that the microstructure or texture of a visual surface is the phenomenal correlate of *some* repetitive type of retinal stimulation. If physical surfaces have regular structures peculiar to them, as wood, cloth, or earth have, the regularity will be projected in a focused image, and this repetitive character of the stimulation, in turn, may well be the basis for the perception of a surface.[6]

The three kinds of order just defined are hypothetical stimuli for pure visual extent, for outlines, and for surfaces in the abstract. But we need to account for surfaces as they are seen in three dimen-

[5] It must be remembered that we are describing what first happens on the retina, not what might happen at later stages in the physiological process of seeing. The occurrence of a margin or outline in perception is determined primarily by the step from light to dark but also of course by the subsequent events in the optic nervous system. The latter may be guessed at from such phenomena as brightness contrast at a border and the inhibition of one border in perception by an adjacent border or a succeeding one, all of which suggest some kind of a process of "contour-building." The significant experiments on this problem are described by Bartley (6) in his chapter on visual contour. The stages intermediate between a true contour and a shadow penumbra have been studied by MacLeod (78) together with the accompanying effects (contrast or constancy) on the areas separated by the contour or penumbra. Here also there is presumably some kind of interaction between adjacent areas.

[6] A striking illustration of this point has been suggested by Dr. Leonard Carmichael. Many of the great classical painters, especially those Dutch painters who worked with magnifying glasses and the finest of brushes, could simulate velvet, satin, the texture of flower-petals, and even the peculiar sheen of a drop of water on the flower by the precise arrangement of spots of pigment. The microstructure of the paint was quite different from the microstructure of the real fabric, the real petal, or the real water-drop. What the painter could reproduce was *the microstructure of the light reflected from these surfaces*. Qualities of lustre, softness, hardness, wetness, and the like are very clear in these paintings. The analysis of visual texture will be carried further in the next chapter.

sions, and for the background surface best exemplified by the terrain. For that it is necessary to consider a *serial* or *progressive* order of elements in the retinal image, or *gradients* as forms of stimulation.

Retinal Gradients as Variables of Stimulation – Texture

Consider for a moment the physical environment from which light is reflected and which is projected on the retina. The problem of distance perception has been reduced to the question of how we can see surfaces parallel to the line of sight (Figure 20). These will be called *longitudinal* surfaces to distinguish them from *frontal* surfaces, which are perpendicular to the line of sight. The former are best exemplified by the ground; the latter are characteristic of objects. The surfaces of the physical environment and its parts are either longitudinal, frontal, or somewhere between these two extremes.

In Figure 21 the material surface *AB* is a longitudinal surface, and the surface *BC* is a frontal surface. In the retinal image *ab*, there exists a gradient of texture from coarse to fine, whereas in the retinal image *bc* no such gradient occurs, and the texture is uniform throughout. The diagram may be conceived either as a cross-sectional view from the side (*AB* is a floor or the ground), or from above (*AB* is a wall to the right of the observer). The slant of a surface is something that we can see, and the surfaces of the visual world are in fairly good agree-

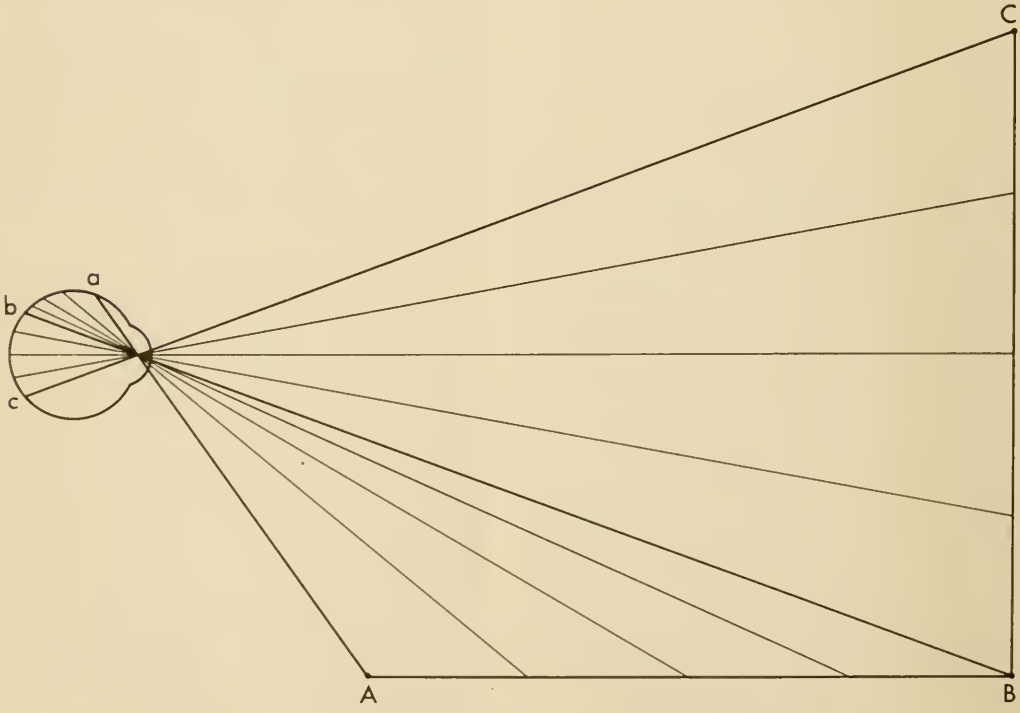


FIGURE 21. The Optical Projection of a Longitudinal and a Frontal Surface

ment with the surfaces of the physical environment with respect to slant. Moreover, as everybody knows, a photograph or a painting can serve as a good substitute for a physical environment in yielding a picture-world with surfaces which seem to slant or confront us just about as they did in the original. The picture surface is flat, but we have all learned to neglect that impression and to see an array of longitudinal, transverse and slanting surfaces which make up the "space" of the picture.

The retinal stimulus-variable which makes possible the perception of a longitudinal surface must be a continuous change of some sort in the image of that surface. To the distance of the physical surface at successive points there must correspond a variation in the image at the projected points. Then, as the image differs progressively from point to point, the perceived surface can differ correspondingly in its distance or depth. There must, in other words, be a stimulation *gradient.*

We can now define a fourth type of order among the elements of a retinal image. It would be a serial change in the length of the cycles of a repetitive order. An example might be *ddddlllldddllldddlldl.* If a repetitive order is the stimulus for visual texture, this would constitute a *gradient of the density of texture.*

We know from ordinary experience that the texture of different surfaces may vary from coarse to fine. The various grades of sandpaper used by carpenters differ in just this respect. Figure 22 shows the same texture in various grades of density. When the image of a *single* surface varies progressively in this way, it may be that *the gradient of density is an adequate stimulus for the impression of continuous distance.*

In order to verify this hypothesis a program of experiments would be necessary, and a beginning on such a program

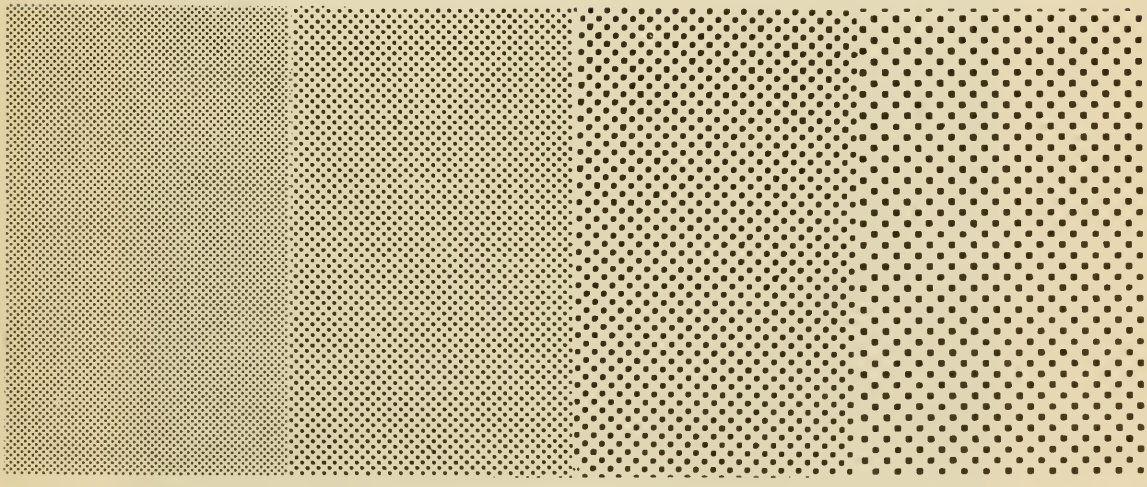


FIGURE 22. Grades of Texture

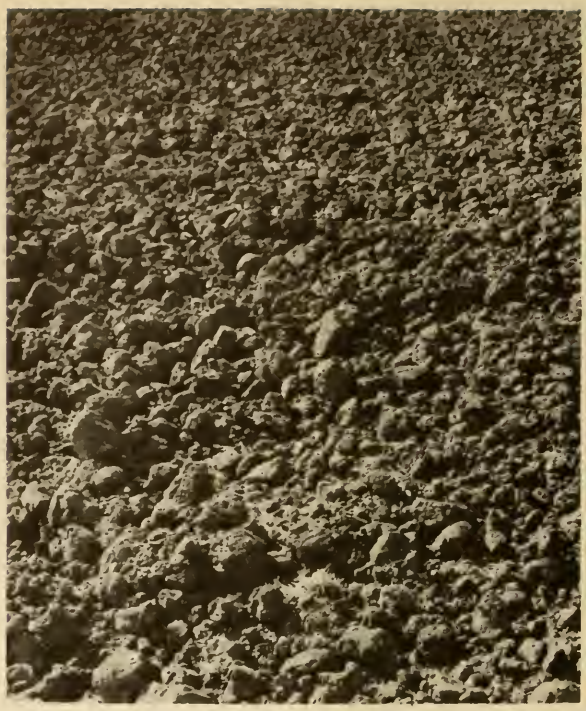


FIGURE 23. Gradients of Natural Texture and the Resulting Impression of Continuous Distance

FIGURE 24. Texture-Perspective and the Impression of Receding Surfaces

Courtesy of Professor R. B. MacLeod

will be described in subsequent chapters. The hypothesis can be illustrated in a preliminary way by examining pictures with respect to the impression of depth and distance which they yield. Figure 23 shows two examples of textured surfaces found in a natural environment in which a gradient from coarse to fine runs from the bottom to the top of the picture. Although the elements of the texture in the two cases are of different shape and mean size, the *gradients* in both pictures are similar. In both pictures there appears a continuous increase in the visible distance of the surface. The impression of a ground extending away from the observer is fairly compelling. In Figure 24, many different gradients of texture-density are combined to yield a complex scene. Half a dozen different kinds of texture are visible in the photograph.

These photographs represent surfaces which are familiar in everyday vision. Although the gradient of texture is the only noticeable variation to be discovered in them, they are *interpreted* by most observers from cues present in the picture and are given a meaning. The meanings usually assigned to the upper pictures are a ploughed field and a field of growing alfalfa, which are correct. It is possible to suppose that the interpretation is the cause of the depth-impression. Such would be the explanation given by an empirical theory of space-perception. Figure 25, however, was constructed artificially out of line-segments, with a gradient of lines and gaps decreasing toward the top of the picture. The impression of a surface extended in distance is clear in this picture as well as

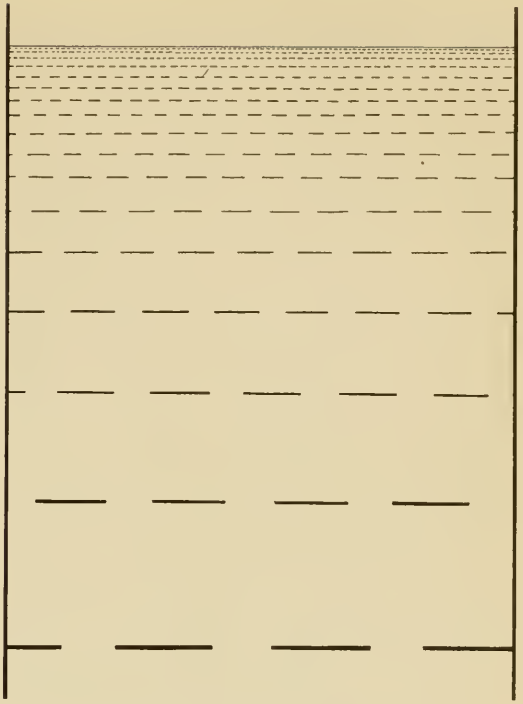


FIGURE 25. A Gradient of Artificial Texture and the Impression of Continuous Distance

in the others. This result suggests that the gradation of texture elements, not the familiarity of elements, is the principal cause of a depth-impression. The last picture may also be interpreted as a level terrain extending off to the horizon, but there are no actual cues for such a meaning, and we may conclude that the impression of distance is an immediate process, while the interpretation follows upon it. "Immediate process" does not imply an innate intuition of distance; it only implies that the impression of distance may have a definable stimulus just as the so-called "sensations" have.

The line segments of Figure 25 were not drawn so as to fall one above the other in straight lines converging to the

horizon, but were instead offset. Aligning them would have induced the familiar appearance of linear perspective. The gradient of texture is not the same thing as ordinary perspective, although the two are united by underlying principles as will be shown. The projected size of things in the environment does decrease as their distance increases from the observer and as their size approaches zero or "vanishes" at the horizon. In this respect the texture of a longitudinal surface and the perspective of objects are alike. But the former leads us to a general phenomenon, of which the latter is only a special case.

The artificial texture of Figure 25 might have been drawn with the line segments at the bottom of the picture twice as long as they are and the line segments above also twice as long all the way up to the level where they diminish to zero. In other words the horizontal dimensions, but not the vertical ones, could have been proportionally increased. The resulting impression of distance on a surface, however, would have remained as strong as before. The only change would have been a faster *rate* of decrease of the line segments from the bottom to the top of the picture, or a larger *angle of convergence* of the theoretical lines connecting their ends (linear perspective). So long as the elements approach a vanishing level at the top of such a picture, the impression of a sort of disembodied terrain is the result. An increase in the gross size of the lines suggests an impression either of larger texture elements or of viewing the terrain from a lower position, down near the ground (Figure 61, Chapter 7). In perceiving distance on a real terrain, similarly, the gross size of the texture elements on the retina will vary depending on whether they are predominantly sand, grass, bushes, or trees, and also on whether the observer is flying an airplane, perched on a telephone pole, standing, or sitting on the ground. Whatever their size may be, however, they diminish to zero in a gradient up the visual field.

The hypothesis implies that a gradient of texture in the visual field corresponds to distance in the material environment on the one hand, and to distance in the visual world on the other. If true, this principle should apply not only to distance-perception on the ground, in aerial and out-of-doors space, but also to distance-perception in the civilized spaces of rooms and other man-made surfaces. In order to apply the principle, we need to remember the types of surfaces already distinguished: longitudinal, frontal, and slanting, with respect to the line of sight. Gradients of texture on man-made surfaces may decrease, but the texture does not diminish to a zero limit as that of the terrain does. On such bounded surfaces, the rate of the gradient is a function of the slant of the surface. A gradient of texture may decrease rapidly, slowly, or not at all, and these are the three respective conditions for a longitudinal, a slanting, or a frontal surface. The texture of a surface faced directly does not change from coarse to fine, and correspondingly an unchanging texture gives the impression of a frontal surface. When there is any gradient of texture, it may decrease upwards, from left to right, right to left, or downwards, and these are the four respec-

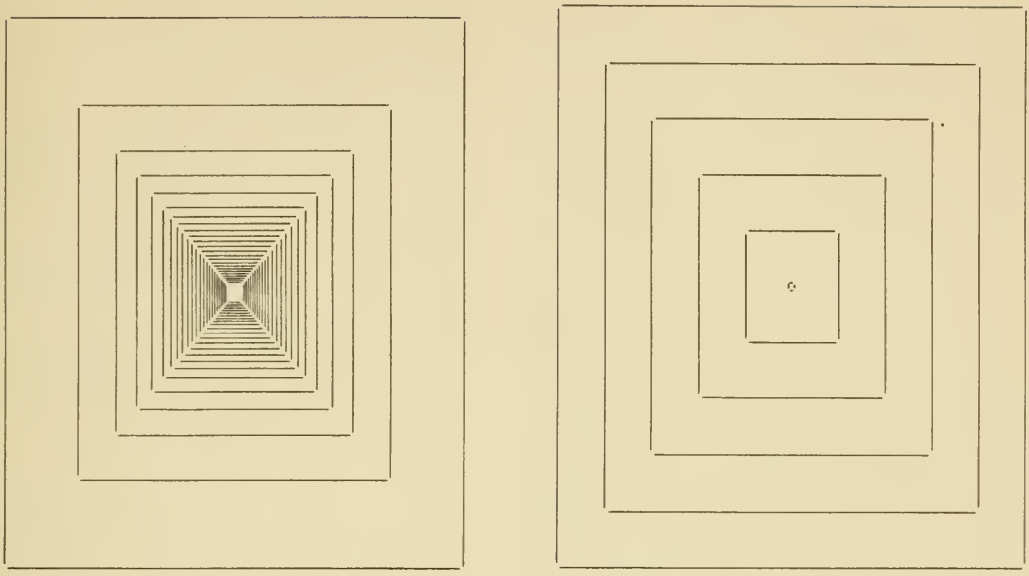


FIGURE 26. A Gradient of Density in Four Different Directions, as Compared with an Even Density

tive conditions for a floor, a left-hand wall, a right-hand wall, and a ceiling. These rules are illustrated in Figure 26, where an artificial gradient has been constructed in each of these four directions. It can be compared with the similar figure beside it which lacks any gradient and where the surface, insofar as a surface is represented, appears to lie in the plane of the picture.

If the slant of any plane surface, such as a floor or wall, has a unique gradient of texture, then the changing slant of a curved surface or one with edges, such as an object possesses, should have a unique change of the gradient of texture. It therefore seems possible that a change of gradient may be a stimulus for the impression of depth and relief in an object.

The Cues for Distance as Stimulation Gradients.

The historical origins of the traditional cues for distance have already been discussed in Chapter 2. The accepted list of these criteria or signs usually includes the following:

1. Linear perspective.

2. The apparent size of objects whose real size is known.

3. The relative apparent motion of objects as the observer moves his head. This is often called motion parallax.

4. The covering of a far object by a near one, or the superposition of one contour on another produced when one object "eclipses" another.

5. The change in color of distant objects, to which is sometimes appended the loss of sharp outline and detail. This is called aerial perspective.

6. The degree of upward angular location of the object in the visual field, the ground and skyline being necessarily implied as the background.

7. The relative brightness of the object. This has been conceived by some writers to be an inverse indicator of its distance; optically, however, this is based on a misconception. It is sometimes mistakenly assumed that the more distant an object is in the ordinary environment the lower will be the intensity of its retinal image, but this principle applies only to point-sources, not to reflecting surfaces.

8. The relation of the lighted to the shadowed areas of an object, or shading. This is an indicator or sign, not of distance but of the depth or relief of a single object.

The "secondary" signs listed above have traditionally been considered less important than the "primary" signs of distance and depth listed below:

9. The disparity of the binocular images of the object as a cue to its depth, and the relative disparities (crossed and uncrossed) of different objects as cues to their relative depth.

10. The degree of convergence of the eyes on a fixated object, the convergence being inversely related to its distance.

11. The degree of accommodation of the lens for a fixated object necessary to maximize the definition of the image.

Since we have now reformulated the problem of distance and suggested a theory of texture gradients, these factors in depth perception must be re-examined. For when they are considered as variables of perception rather than as facts of knowledge – once we understand that they apply to an array of objects in the visual field rather than to a single object – they may all prove to be gradients of stimulation, or related to such gradients.

Linear perspective, for example, might be a special result of the decrease in size of figures in the visual field from the lower margin to the horizon.[7] Motion parallax, as seen from a train window, might be a special result of the gradients of deformation which fill the visual field when the observer moves. Superposition of one shape on another is best understood by analyzing the outline between them, and this outline may prove to involve a step separating two continuous gradients. Aerial perspective is, for the most part, a simple gradient of hue in the visual field, a gradient running toward the violet end of the spectrum. The shading on a curved surface is obviously a gradient, as every artist knows. It would be surprising but significant if retinal disparity, like the other signs of depth, could be defined as a gradient of stimulation – not of the single retina, it is true, but a gradient of the theoretically combined images of the two retinas. A visual field obtained with both eyes open, as we shall see, always contains a gradi-

[7] Linear perspective is also a geometrical technique for drawing the edges of straight-sided objects, but the two meanings should not be confused.

ent of "double images." Finally, the cue of upward height in the visual field becomes intelligible in the light of a "ground-theory" instead of an "air-theory" of visual space, inasmuch as the ground embodies and is a precondition for the gradients mentioned.

These possibilities for a general theory of visual gradients are, at this stage, merely claims. They may appear plausible after they are developed. If they are to be convincing they will have to be demonstrated, and the attempt to do so will be made in the next two chapters.

The Concept of Gradient

The word *gradient* means nothing more complex than an increase or decrease of something along a given axis or dimension. As such it is related to the plots or curves of analytical geometry. The gradient of a railroad or highway, for example, is its change of altitude with distance. This change may be positive or negative or zero (the last being a *level* gradient), and it may also be rapid or slow, corresponding to a steep or a moderate gradient. The gradient may itself change, as the slope of a road does in hilly country. When the change is very abrupt – if the slope of a road should end at a cliff – it is properly thought of not as a gradient but as a step or discontinuity. These concepts appear to be admirably adapted for describing the retinal image, since both gradients and steps of stimulation can be found within it.[8]

According to the evidence of C.M. Child and his students (22), all living tissue is characterized by *physiological* gradients. Along the axes of an organism, from head to tail, from front to back, or from the apex to the base of a limb, there exist gradients of metabolism, excitability, and growth. Now these gradients of activity are not merely spontaneous self-generated phenomena but are also reactions of the living cells to their environment. Although conditioned in part by the genes within each cell, these reactions are primarily determined by differentials of temperature, light, chemical concentration, or electrical activity – that is to say, by gradients of these kinds of energy. The proposal that the light-sensitive cells of the retinal mosaic and the neural tissue in the brain connected with them can react to gradients of stimulation, therefore, is not without analogy in other kinds of organic tissue. The special application

---

[8] In Bartley's work on vision (6), he has used the term gradient to refer to a change in the luminous intensity of stimulation at a border within the retinal image. He is thinking of a microscopic shadow-edge as it falls on the mosaic of retinal cells which can be considered a gradient since the change must be distributed over the width of a number of cells. This is what was called a *step* above. It might be termed a *microgradient* as distinguished from a *macrogradient.* Visual contours, visual acuity, and the elements of visual texture all seem to involve microgradients of luminous intensity. The cues for the depth or slant of a surface, on the other hand, seem to involve macrogradients over a considerable dimension of the retina. A gradient of the density of texture would be one case. A gradient of shading in the hollows of a surface or the shading toward the unlighted side of a curved object would be another. The penumbra of a shadow is such a gradient according to MacLeod (78), and he has demonstrated with "artificial penumbrae" the different effects of a steep gradient as compared with a gentle gradient.

of a gradient theory to a *sensory surface* such as the retina or the skin has not yet been attempted, however, and the ways in which neural tissue may differ from other kinds of tissue in this respect is not known. Is it suggestive for the psychologist's problem of the physiological correlate of *visual* form that biologists have found the concept of gradients useful in understanding the development of *organic* form?

Child has pointed out (22, p. 275) that physiological gradients may overlap and combine geometrically within the organism to yield what could be termed a physiological "field." The concept of a field of forces, in analogy with a gravitational or electrical field, has found application not only in physiology but also to problems of visual perception and of goal-directed behavior (68, 74). Child argues, however, that in physiology the field-concept as such is vague, and that a field theory is useless without analysis. Such analysis can only be carried out in terms of the definable and measurable gradients which constitute the field. The writer agrees with Child in this criticism. Field theory in psychology, as practiced by Gestalt psychologists, is not always rigorous or precise. Assuming that a field is determined by its gradients, an analysis of

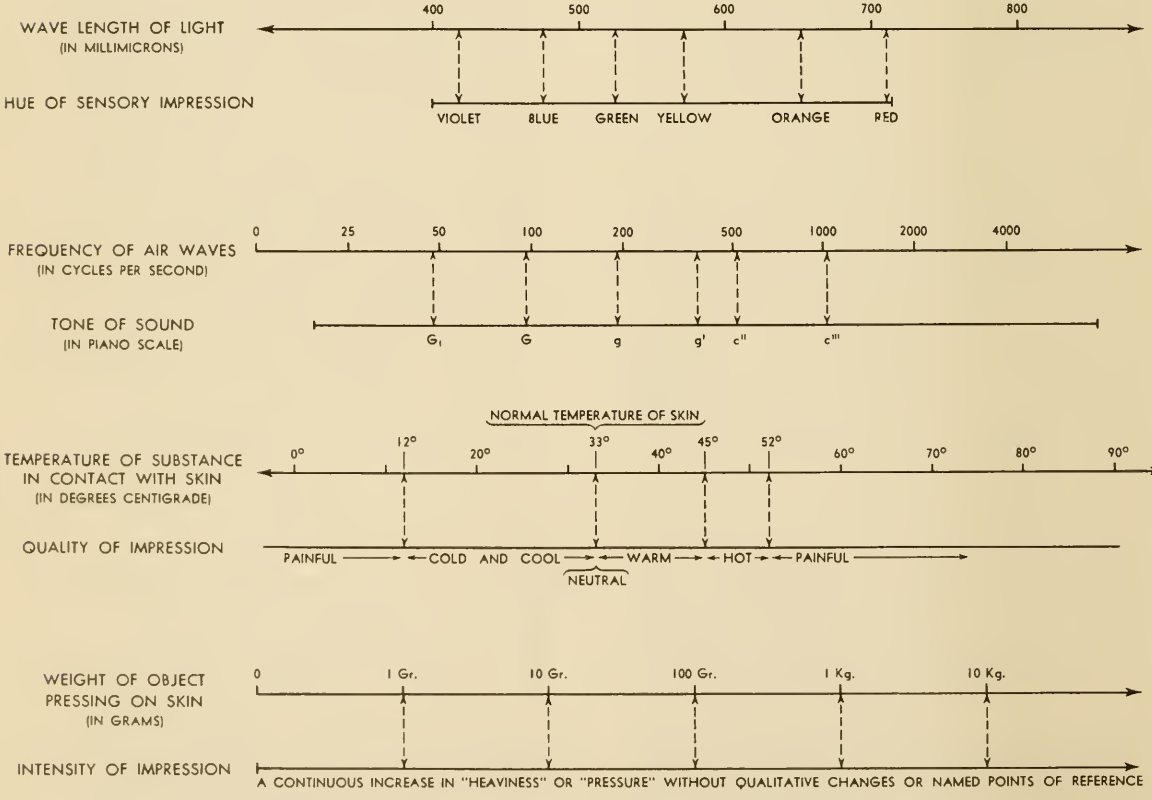


FIGURE 27. Examples of Psychophysical Correspondence

the stimulation gradients involved in perceptual (and possibly in behavioral) fields would probably be more profitable than further attempts to discover the laws of field-phenomena as such.

The Concept of Psychophysical Correspondence.

The correspondence of the variables of perception to the variables of stimulation is exemplified in Figure 27. Four pairs of such variables are given. The lower line of each pair represents a variable of experience. Each line is to be regarded as continuous. Points on the upper line represent *possibilities* or *instances* of stimulation, and points on the lower line represent *descriptions* or *judgments* of the ensuing sensory impression, but these "points" are not isolable. They cannot be thought of as stimuli and sensations respectively; the points are simply *numbers in a serial order.* The variable of physics and the variable of experience in each case are in a one-to-one correspondence. In terms of the geometrical model, for every point on the upper line there is one and only one point on the lower line. The lines (or numbers) need not be conceived as scales possessing units of length. For present purposes they are continuous series merely. An introduction to the problem of *scaling* the variables of physics and of experience – the "dimensions of consciousness" – is to be found in Stevens (101), and the background of this problem is given by Boring (10).[9]

A few pairs of corresponding points are indicated by dotted lines. It is noteworthy that, for some variables like temperature and others like weight, the correspondence of sensory qualities to their physical variables may be *shifted* by adaptation. For example, after holding the hand in warm water, a stimulus which formerly felt warm now feels neutral and a stimulus which formerly felt neutral now feels cold. The correspondence has been displaced, *but it is still a specific and regular correspondence* (38). It is reasonable to suppose that the spatial qualities of the world, as well as the "sensory" qualities illustrated above, may undergo a shift in their correspondence to stimulation without a destruction of the correspondence. Something of this sort probably occurs in the process of getting adapted to eyeglasses.

Summary

A theory of visual space perception has now been outlined. Its strength or weakness can be estimated better if its postulates are made clear. It may be useful, therefore, to summarize the theory in a series of propositions.

1. It was assumed that the fundamental condition for seeing a visual world is an

---

[9] Boring has also discussed the seeming paradox (exemplified by auditory "volume") that there may exist more dimensions of sensory experience than there are simple dimensions of the physical stimulus (11). The difficulty is resolved if one defines the dimensions of the stimulus as those variables of stimulation, *however mathematically complex,* with which the variations in discriminative response prove to be correlated as the outcome of a psychophysical experiment.

array of physical surfaces reflecting light and projected on the retina. This is in contrast with the usual assumption that the problem of perception should start from the geometrical characteristics of abstract "space."

2. In any environment, these surfaces are of two extreme types, frontal and longitudinal. A frontal surface is one transverse to the line of sight, and a longitudinal surface is one parallel with the line of sight.

3. The perception of depth, distance, or the so-called third dimension, is reducible to the problem of the perception of longitudinal surfaces. When no surface is present in perception because of homogeneous retinal stimulation, distance is indeterminate. Although the ground is the main longitudinal surface, the walls and ceilings of man-made environments constitute three other geometrical types.

4. The general condition for the perception of a surface is the type of ordinal stimulation which yields texture.

5. The general condition for the perception of an edge, and hence for the perception of a *bounded* surface in the visual field, is the type of ordinal stimulation consisting of an abrupt transition. The simplest and best understood kind of retinal transition is one of brightness.

6. The perception of an object in depth is reducible to the problem of the changing slant of a curved surface or the differing slants of a bent surface. In either case the problem is similar to that of how we see a longitudinal surface.

7. The general condition for the perception of a longitudinal or slanted surface is a kind of ordinal stimulation called a gradient. The gradient of texture has been described, and it has been suggested that gradients dependent on outlines, a gradient of retinal disparity, a gradient of shading, a deformation gradient when the observer moves, and possibly others, all have the function of stimulus-correlates for the impression of distance on a surface.

Conclusion

The correspondence of the visual *field* to the total retinal image is an anatomical point-for-point correspondence which is not hard to understand. The correspondence of the visual *world* to the total retinal image is an ordinal correspondence which is more difficult to analyse and specify. But the latter correspondence *is no less literal and exact,* we may believe, than the former, and it is clear that the way to determine it is to find the obscure variations of the projected image which yield coordinate variations in perception.

6

The Stimulus Variables for Visual Depth and Distance -- Momentary Stimulation

The Stimulus Gradients of the Density of Texture and the size of Objects . . . . The Depth-Shape of Objects—Gradients of Texture and Grades of Illumination . . . . The Stimulus Gradient of Binocular Disparity Between Images . . . . Definition or Resolution in the Retinal Image . . . . The Gradients of Aerial Perspective . . . . Summary

Consider once more the way in which the physical world is projected on the retina of the eye, remembering that a typical scene consists of the *ground* and of *objects* (and remembering also that the image is inverted). Near objects will be imaged large and high up on the retina. Far objects will be imaged smaller and lower down. Very far objects will be imaged so small that their size approaches a vanishing point. At the line where the earth ceases and the sky begins, the separate images of objects become indiscriminable. This is the *fundamental* world for vision. The rule is that those parts of the world just under a man's nose are projected large and those parts at a distance are projected small. In the visual field, the patches and spots of color are gross and far apart at the lower boundary of the field and become progressively smaller and denser upwards towards the horizon.

Let us make two assumptions about the typical physical world of ground and objects and assert that, first, objects tend to be in contact with the ground instead of up in the air and that, second, they tend to be distributed over the ground with an even scatter. The first assertion will probably meet with no objections. As for the second, it can be proved that the physical spacing between many kinds of things tends to be regular. The principle holds for grass in a meadow, for trees in a forest, for the boards in a floor, and the patterns of a carpet. Above all, it holds excellently for a special type of object possessing the greatest importance for vision – the single element of the texture of a surface. The grain or structure of a surface is made up of units of one kind or another which are repeated over the entire surface. These units are characteristic of the physical substance in ques-

tion. They determine an array of light reflections from the surface which can be focused as an image. The number of these units which can be counted in any single square inch (or yard or mile) tends to be the same as the number in any other square unit of measure, and this is what is meant by an even spacing of the units.

It is of course true that the spacing of either particles on a surface or things on the ground is seldom perfectly regular except in the case of mechanically processed surfaces or man-made layouts of things. Tiled pavements, planted fields, telegraph poles, and railway tracks are examples of exact spacing. In this special case, as we shall emphasize, the retinal image and the resulting perception can be fitted to a set of special rules, but this fact must not distract us from considering the importance of natural distributions.[1]

Our hypothesis is that *the basis of the so-called perception of space is the projection of its objects and elements as an image, and the consequent gradual change of size and density in the image as the objects and elements recede from the observer.* Whenever the observer moves his head there will also occur a gradual change of motility in the image as the corresponding objects and elements recede. If both eyes are functioning there will be a gradual change in the disparity of the elements of one image relative to the other as the corresponding objects recede. There may be still other changes in the retinal image corresponding to the physical recession of the environment, but they all presuppose a textured image. The gradient of size and density, therefore, is a necessary correlate of recession. This condensation of the image cannot be eliminated by holding the head motionless or by closing one eye. It has a special status, and consequently it is this gradient which should receive first consideration.

The purpose of this chapter and the next will be to consider, one by one, these various so-called cues for distance perception when they are reformulated as gradients of the retinal image. In this chapter, for the sake of simplicity, the retinal image we shall refer to is the image obtained by a motionless observer with his eyes fixed straight ahead. In the following chapter we shall go on to consider the image obtained by a moving observer, and only then attempt to understand the image of an observer who scans his environment – the image which samples a different cone of light-rays from one moment to the next and which accordingly registers in succession different sectors of the physical world.

The Stimulus Gradients of the Density of Texture and the Size of Objects

Illustrations of a gradient of texture have already been given in Figures 23 and

[1] The structure of substances at microscopic and sub-microscopic levels of size is studied by crystallography and physical chemistry. The structure of the universe on the scale of miles and light-years is studied by astronomy. But the structure of the world on the scale of millimeters and meters – the textures of surfaces and the distributions of objects – is so familiar that it has been very little studied. The eye and the retina are adapted to register the structure of the world only at this range of sizes. Finer and grosser structure can be "seen" only by the use of special devices such as microscopes and telescopes.

25 in the preceding chapter. What is now required is a closer study of this little understood form of visual stimulation. A method of isolating and varying texture is needed. Its implied relation to linear perspective should also be explored, together with its relation to line or contour. But first we need to consider how a gradient of texture on the retina can be artificially produced and hence experimentally controlled.

The Relation of a Picture-Plane to its Retinal Projection. The obvious method of constructing a gradient of texture would be to draw it on paper and present it to the observer's eye in the position of the picture-plane, so that it will produce a retinal gradient of texture. This is the method employed in the exploratory experiments described later in this chapter. Is it a legitimate method? The retinal image of a given object is not a plane projection of that object, such as a drawing or a photograph would be, inasmuch as the retina is a curved surface. The retinal image, moreover, is inverted relative to the plane projection; it is not actually a picture. One might suppose, therefore, that the existence of a gradient in the retinal image could be verified only by getting at the retinal image itself in some way and measuring it.

Actually, the measurement of a retinal gradient is not necessary, and the measurement of its corresponding plane projection may be substituted for it without error. As Figure 28 illustrates, they are in a perfect point-to-point correspondence with one another, and the one arrangement may be mathematically transformed into the other at any time if the dimensions of the retina and the picture are known. It is obviously much more convenient to specify the retinal distribution of light on a picture-plane than it would be to specify it on the curved surface of the retina itself. Moreover, the plane projection is more readily compared with the experienced visual field and hence is easier to conceptualize. We will therefore speak of the plane gradient as a visual stimulus, recognizing that the picture gives no more than a convenient substitute for the retinal gradient.

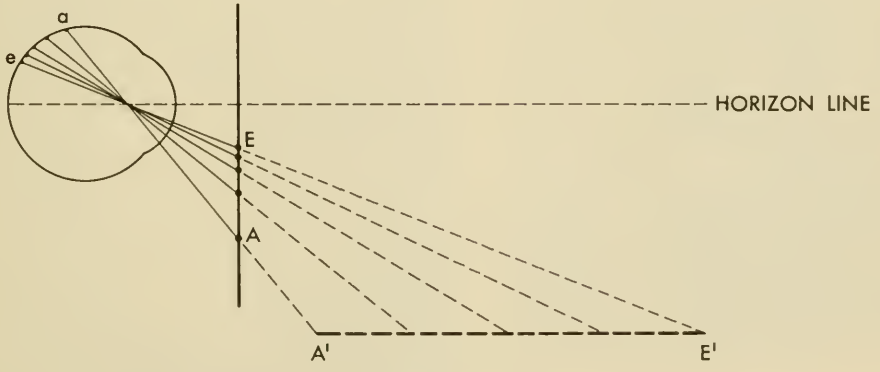


FIGURE 28. A Picture-Gradient, its Retinal Gradient, and the Longitudinal Surface Corresponding to Both

The method of drawing the elements of a texture gradient on the picture plane is undoubtedly a crude one. Techniques of photography, photoengraving, and of controlled optical distortion offer other methods of producing and varying texture which ought to be exploited.

The Elements and Gaps of Visual Texture. There are, as theatrical magicians know, some physical surfaces which are invisible. The test for whether a physical surface does or does not possess visual texture is whether the surface can or cannot be brought into focus by a lens, that is, whether an image of the surface can be produced. It makes no difference, theoretically, whether one uses an eye or a camera for this test. Perfectly flat transparent surfaces or perfect reflecting surfaces, such as glass, cannot be focused on by a camera or accommodated for by an eye in the absence of any highlights or luster. Neither can *any* surface when its illumination is sufficiently low. It is no more possible to get an optical image of a sheet of plate glass or a large mirror (if highlights are absent and the edges of the surface are not in the field) than it is to get an optical image of the cloudless sky or the interior of a completely blacked-out room. Surfaces of this type are fortunately not the surfaces on which we walk and sit and which characterize the objects of our visual world.

Ordinary surfaces are rarely both physically smooth and chemically homogeneous, like plate-glass. If the surface is rough, it has crests and troughs. A piece of cloth, a ploughed field, or a hilly terrain seen from the air are all alike in this respect except for the difference in magnitude and shape of the typical crest and trough. If the surface is smooth but not chemically homogeneous – if it is composed of different substances – the reflectivities of the different particles are likely to differ. An example would be polished granite, or any conglomerate material. In either event, whether the reflecting particles are structural or chemical or both, they will reflect light differentially and the image of the surface will consist in an array of cyclical changes in light energy which we experience as variations in brightness or hue. The optical image, it must be remembered, implies a correspondence between two sets of abstractions, reflecting-points and focus-points, such that the character of the light at the former is duplicated at the latter, point for point. The structural and chemical cycles of the surface, therefore, are projected on the retina as cycles of color in corresponding order.[2]

These cycles, we suppose, constitute the stimulus for visual texture. Both the cycles and the resulting texture can be of many different types, such as the rippled surface of water, the complex roughness of a plaster surface, or the regular units of a grating or a tiled pavement. The type of texture to be tried out in experiments should be as simple as possible and at

[2]The assumption is that a texture can be analysed by plotting it in two dimensions, that is, by specifying the alternations or repetitions of light stimulation along two axes. This is what is meant by cycles. Admittedly this assumption needs mathematical study. A texture cannot be analysed conveniently in terms of lines, nets, grids, or other patterns with which the writer is familiar because these are themselves special cases of texture.

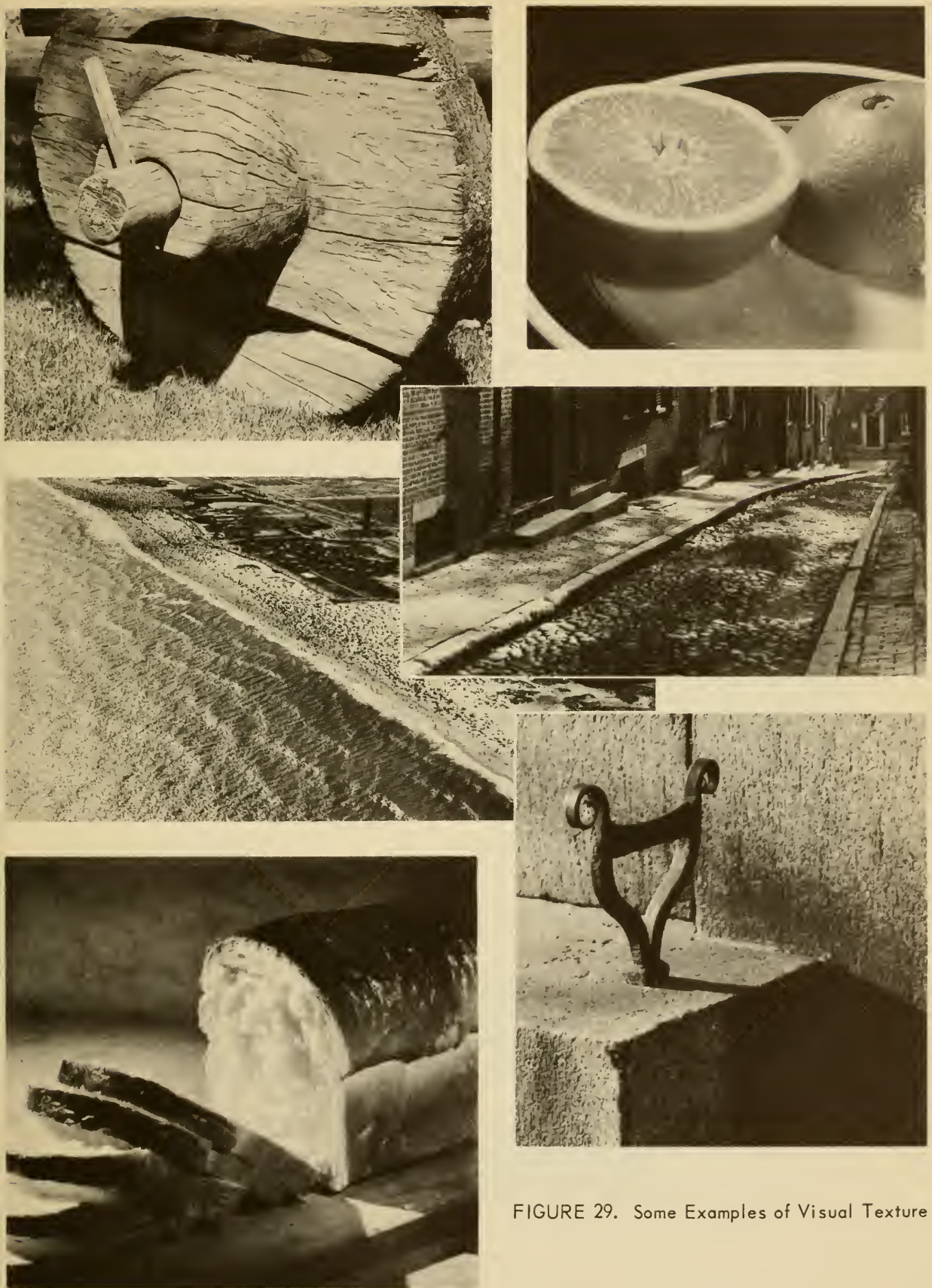


FIGURE 29. Some Examples of Visual Texture

the same time unfamiliar or abstract. The kind of surface which can vaguely be termed "spotted" might be the best to begin with. Its texture could be defined more accurately as consisting of elements and gaps. The elements correspond to the spots and the gaps to the areas between them. If the elements are black and the gaps white, the texture can be artificially constructed with a pen and ink. The cyclical character of texture can be preserved by making the size of the elements a constant ratio of the size of the gaps. Several types of abstract texture can be produced by arranging the elements and gaps in any desired gradient on the picture-plane with respect to their size. We can then determine whether the gradient produces a corresponding impression of a longitudinal surface in experience.

The Method of Drawing Gradients. The projection of a longitudinal surface on a picture-plane is obtained by perspective geometry. The principles governing such a projection are illustrated in Figure 30, which is itself a perspective drawing.

The surface projected in Figure 30 is a series of adjacent squares, like a sidewalk, extending from the eye of the observer to the horizon. We may think of the squares as unit areas of surface. The projection is a series of adjacent trapezoids diminishing gradually to a point, so that the whole constitutes a triangle. Considering the width of each projected square one notes that it decreases upward on the picture evenly on the scale of the picture. This linear decrease in the width of pictorial things is what makes possible the techniques of drawing guide-lines which converge to a vanishing point. What is popularly known as "linear perspective" is comprised by this fact. Considering the height of each trapezoid, however, one notes that it decreases upward on the picture with a gradient which

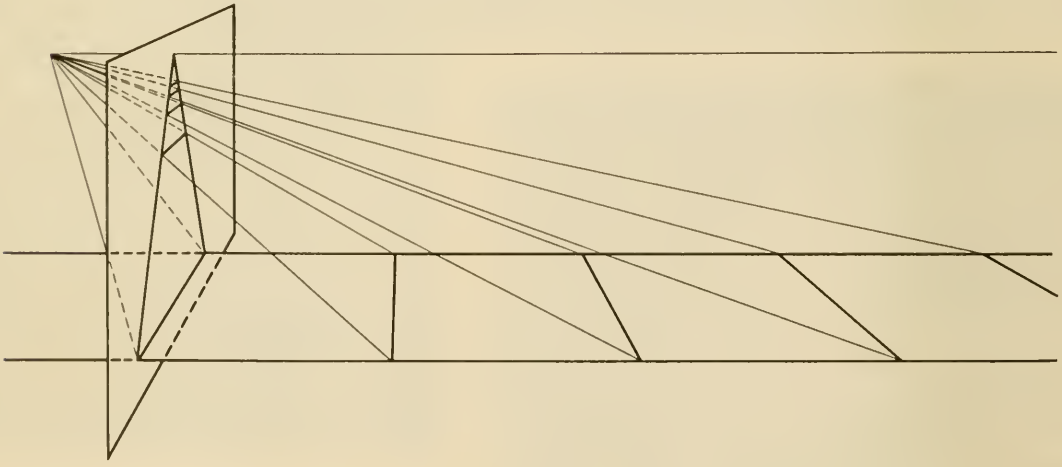


FIGURE 30. The Principles of Perspective for Square Units of the Ground

is not linear but negatively accelerated. The projected height of flat areas, sometimes called foreshortening, can be drawn only by a more complex technique than that of lines to a single vanishing point, a technique less familiar but equally important in the practice of perspective. These two techniques enable the experimenter to construct on paper a gradient of "flat" elements and gaps of texture which will bear a specific relationship to a material surface in the third dimension.

The decreasing projected size of a unit object on the ground with increasing distance from the observer is subject to two rules, one applying to the frontal dimension of the unit object and the other to its longitudinal dimension. The first rule is simply the Law of the Visual Angle, which states that a frontal dimension is projected as a size *(S)* which is the reciprocal of the distance *(D)*. Algebraically this means that S is proportional to $\frac{1}{D}$. Geometrically these frontal dimensions project so as to give a perfectly linear decrease up the plane of the projection, and the ends may therefore be joined by straight lines to a vanishing point, as illustrated. The second rule states that a longitudinal dimension of the unit object is projected as an altitude *(A)* which is a negatively accelerated function of the distance *(D)*. Algebraically it comes out that A is proportional to $\frac{1}{D^2}$. The longitudinal dimension is the one which is said to be foreshortened in the projection, that is, it is the dimension which is compressed relative to the frontal dimension. The successive altitudes of unit objects can be constructed geometrically by the method of diagonals or a variant of it (116), which requires the use of more than one vanishing point and is therefore not illustrated in Figure 30.

The novel feature of this use of the perspective techniques is that they are here applied to the texture of a surface instead of to the edges of a surface, as is usual. Conventional perspective is the perspective of the edges and boundaries of things and when a surface is to be represented only the outlines of the sur-

FIGURE 31. The Perspective of Objects is the Same in Principle as the Perspective of Texture

Gendreau

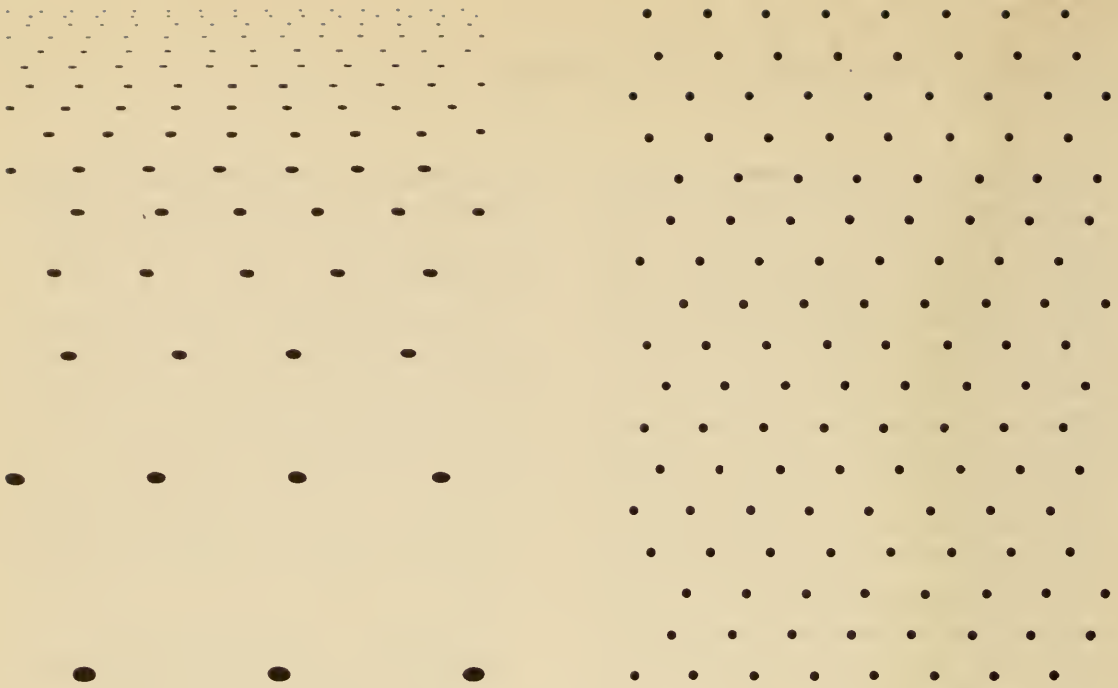
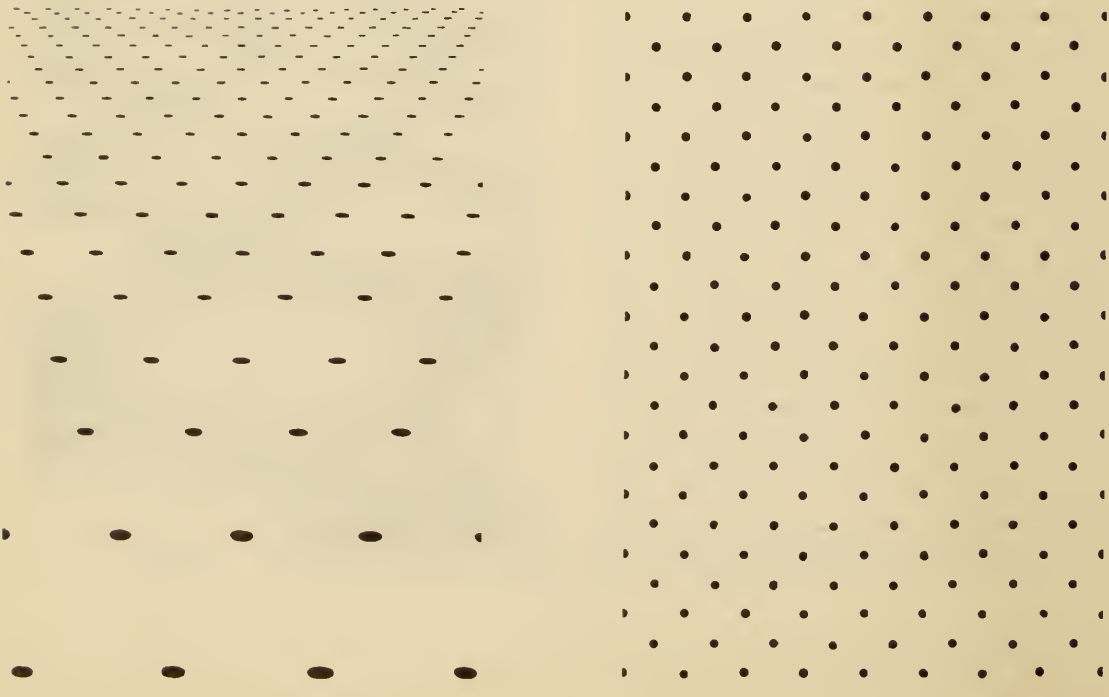


FIGURE 32. Spot-Distributions which Yield Impressions of a Longitudinal and a Frontal Surface

FIGURE 33. Spot Distributions Based on Perfectly Regular Spacing of the Elements



face are drawn. We require a perspective of visual areas.

Texture Gradients and the Impression of Distance. Figure 32 consists of two different arrays of spots and gaps. In the left-hand array, the width of the gaps is governed by a linear decrease upwards and the height of the gaps by a negatively accelerated decrease. In the right hand array there is no decrease; the gradient is zero. The dimensions of the spots are determined by the same rules as for the gaps but, in the drawing, these dimensions are only approximate. For both pictures the presumption of a perfectly equal spacing of the spots on the corresponding objective surface would produce vertical and diagonal alignments in the pictures, similar to the alignments observed in planted fields or wall-paper patterns. Such alignments would produce the familiar appearance of perspective (Figure 33). In order to demonstrate that this special type of perspective is not necessary in a texture gradient, a slight irregularity of spacing has been presumed such as to destroy vertical and diagonal alignments. This was achieved in the drawing by offsetting each successive row of spots to the right or left in a random fashion. The horizontal alignment of spots was left undisturbed by this expedient.

The left-hand array of spots gives an impression of continuous receding space, while the right-hand array does not. The space is, moreover, one on which the observer looks down from above. It is what might be called a ground scene. If the pictures are inverted, the sensory impression of distance remains but it is a distance looked at from below – a ceiling scene. The retinal gradient is reversed in the latter situation. There will be other instances later in which the reversal of a gradient yields the impression of the inversion of a surface.

According to the theory formulated, the distribution on the left should produce a longitudinal surface and that on the right a frontal surface. With respect to being a surface this prediction is not wholly confirmed, since for many observers the gradient of spots suggests an array of objects on an invisible ground. Interpretations such as "lily-pads on the surface of a pond" or "the heads of people in a crowd" or "disembodied cabbage-patch" are frequent. To common sense, an array of objects seems to have nothing in common with a surface, and a stimulus for the first could hardly also be a stimulus for the second. Nevertheless a transition must be possible between an array of objects and an array of texture elements. It is not always clear whether a given scene is one or the other. A forest, for example, appears to consist of trees to the observer nearby but a surface to the observer at a distance or the flier at an altitude. The significant fact is that the same laws apply to both.

Figure 32 was composed of flat texture elements of the sort found on a physically smooth surface. They might have been spots on a floor, for instance. Many physical surfaces, however, are composed of elements which do not lie on the surface but project upward from it. A stubble field would be one example. Figure 34 was constructed in the effort to produce a synthetic gradient of texture of this sort.

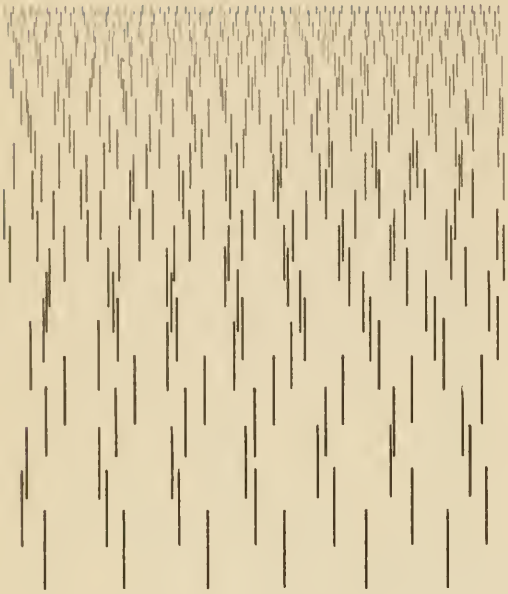


FIGURE 34. A Gradient of Textural Elements which Project Upward from the Surface

The elements, like those of Figure 25 in the last chapter, are line segments but, unlike them, they are vertical on the picture-plane. The height of the elements decreases up the picture in a linear gradient, like the width of the lines in Figure 25. The width of the gaps between the line segments of each row also decreases in a linear gradient. The only non-linear gradient in the drawing is in the height of the gaps, or the mean vertical distance between the bottom of the lines. In this drawing, as in the last, all observers see a plane of distance and also see the lines as perpendicular to this plane.

These drawings, including Figure 25 in the last chapter, scarcely make a beginning at exploring the complex relations between visual texture and space. The artist who knows how to see the various textures of the material world and knows how to reproduce some of them will be able to think of a hundred interrelations and subtleties not even hinted at. The drawings do suggest that the theoretical-experimental approach of the psychologist to such problems may be illuminating, and they tend to verify the hypothesis that a gradient of texture is, in isolation, a stimulus for the impression of continuous distance on a surface.

Gradients of the Spacing between Edges — Straight Lines and Linear Perspective. The impression of distance on the left-hand section of Figure 36 is not surprising since it is common in pictures and photographs — stimuli to which we have all been exposed since infancy. Floors and pavements generally have rectilinear joints between the parts which compose them (boards or tiles, for example) and these joints reflect light in much the same way as do the lines of a drawing. One might think of them as *inlines* as distinguished from *outlines* and they may have the effects of producing a surface in perception just as texture produces a surface. It may be noted in passing that both the inlines and the outlines of a longitudinal surface are projected on a picture-plane by the rules of geometry illustrated earlier. We are, however, deferring all consideration of outlines until we come to the problem of objects.

Surfaces composed wholly of such inlines are represented in these three drawings. When their density increases toward the top of the picture, the impression

FIGURE 35. Two Pictures with Varieties of Texture within the Same Scene

Courtesy of Harris-Seybold Company

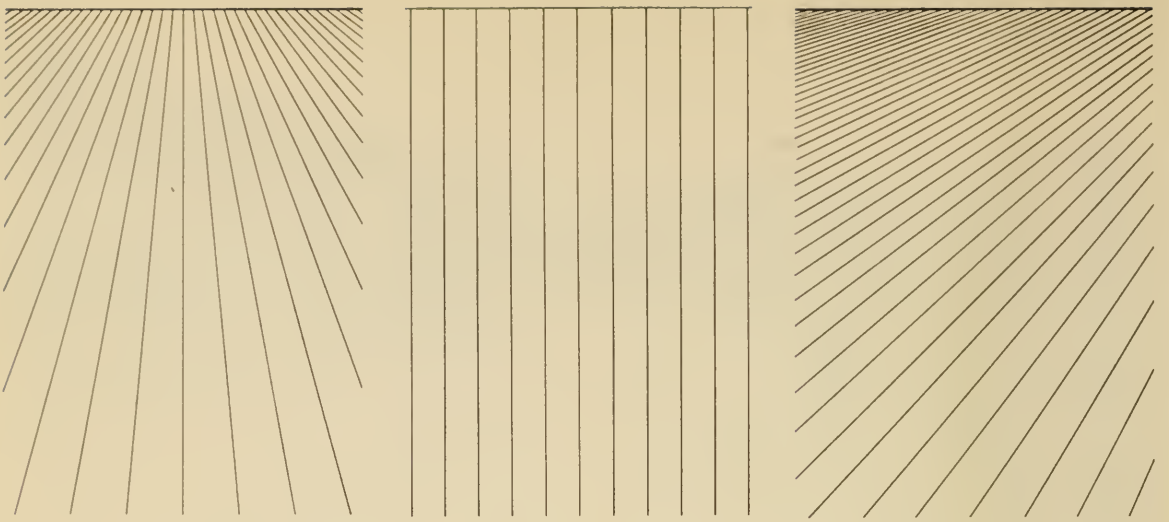


FIGURE 36. The Effect of a Width-Gradient, or Linear Perspective

is that of a longitudinal surface; when the gradient is zero the impression is that of a frontal surface. The meaning most frequently suggested by the drawing is that of a floor made of boards, or (since the texture of the boards is missing) an open grille of some sort. The direction in which the lines run within the visual field makes no difference for the impression of distance; it is the gradient of density which yields the sense of a continuous third dimension. In the right-hand drawing, for instance, although the lines run off toward the right, their density increases up the picture, and the surface we see is the same as in the left-hand drawing.

When the inlines of a surface run horizontally across the visual field, as in viewing a board floor crosswise of the boards, the space between lines decreases on the scale of the picture in the specific gradient appropriate to longitudinal dimensions on a plane surface. Figure 37 consists of three drawings of this sort. That on the left yields the impression of a longitudinal surface like the ground. That in the center yields the impression of a frontal surface. That on the right yields an impression of increasing distance at the top, but the drawing does not have the kind of flat distance which the first has. The first drawing is determined by the specific gradient derived from the perspective of a floor. The second has a zero gradient of density, or the perspective of a wall. The third has an increase of density up the picture-plane, such as to give the perspective of a curved surface which slants upward from the observer and then away from him. Some observers can see just such a curved surface in the drawing.

The fact that a mathematically simple gradient corresponds to a geometrically complex surface in this case should not be considered puzzling. All gradients become mathematically very complex when projected on the retina. The implication to be noted is that a curved physical surface may have just as specific a correlate

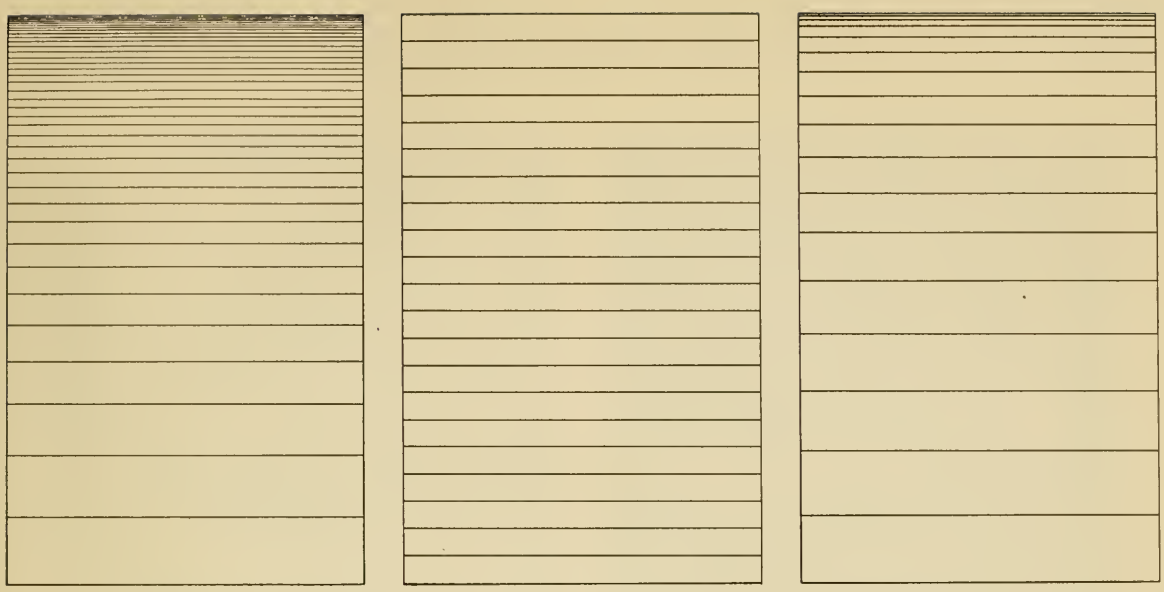


FIGURE 37. Gradients of Vertical Spacing Corresponding to Three Types of Surface

when it is projected as a flat physical surface has.

A flat ground surface composed of geometrically simple forms such as tiles or paving stones can be projected on a plane with none other than straight lines. It is an ideal exercise for the use of the classical rules of linear perspective. All lines either converge to some vanishing point on the horizon or are horizontal and parallel. Figure 38 is an example of such a surface. The units are the triangles formed by the diagonals of a square. There are four sets of parallel lines on

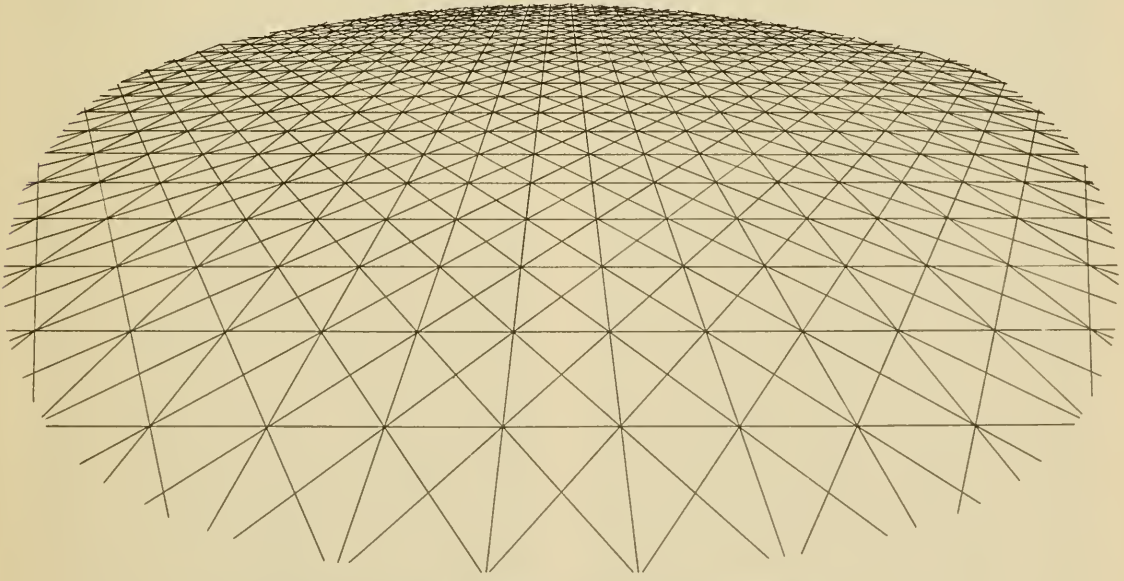


FIGURE 38. The Perspective of a Pavement

the pavement, and they are represented by three sets of lines to three vanishing points, plus one set of horizontal lines. The drawing illustrates the width and height gradients which we have isolated in previous drawings, and gives some indication of the method of constructing them. The combined gradients yield a vivid and compelling perception of continuous distance.

The physical composition of a surface like this differs of course from the physical texture of most natural surfaces. Instead of what we have called elements and gaps it is composed of flat geometrical forms which fit into one another. It nevertheless yields the same type of visual perceptions as a surface of natural texture and it may be considered as simply a special case of a much more general phenomenon. Our conclusion is that a gradient of inlines is a stimulus for continuous distance on a surface as well as a gradient of texture composed of elements and gaps.

The Apparent Size of Familiar Objects as a Cue to their Distance. The classical theory that distance is perceived by a mind seated in the brain and making use of the cues presented to it by a retinal picture is very different from the theory we are now engaged with. While ours conceives an array of texture elements or objects on a substratum, the former conceives an isolated object or two in empty space. Instead of asking "How do we see continuous distance from here in all directions?" the classical theory asks "How do we judge the distance of that object, or the relative distance of those two objects?" Accordingly, it treats the imaged size of a given object not as merely something related to the other sizes in a graded array but as something to be sensed in isolation and then interpreted. A given retinal image as such could give no evidence of the distance of its object, for the object might be either something small and near or something large and far off. Such a confusion would actually occur in the empty space under consideration. Only if the real size of the object were known because the object was familiar could a decision be made between these possibilities. In that event the sensed size could be compared with a remembered size and the distance of the object could be obtained by a kind of unconscious computation. A human figure could be perceived as 100 yards distant rather than 50 yards because it is seen as a 6 foot man and not a 3 foot boy. If the man were a midget, however, the perception might be erroneous. According to this theory the perspective size of an object is only a cue to its distance in the sense that it provides a fact upon which the mind can work. The theory of cues is obviously a very roundabout way of explaining distance perception as compared with our hypothesis, which suggests that gradients of size and density, being in correspondence with physical distance, are stimulus correlates for a continuum of seen distance. The gradient theory accounts for the distance of all objects in the array, rather than the single object on which attention is fixed.

It would be too much to say that the inferring of the distance of a familiar object on the basis of its perspective size never occurs in human perception, for

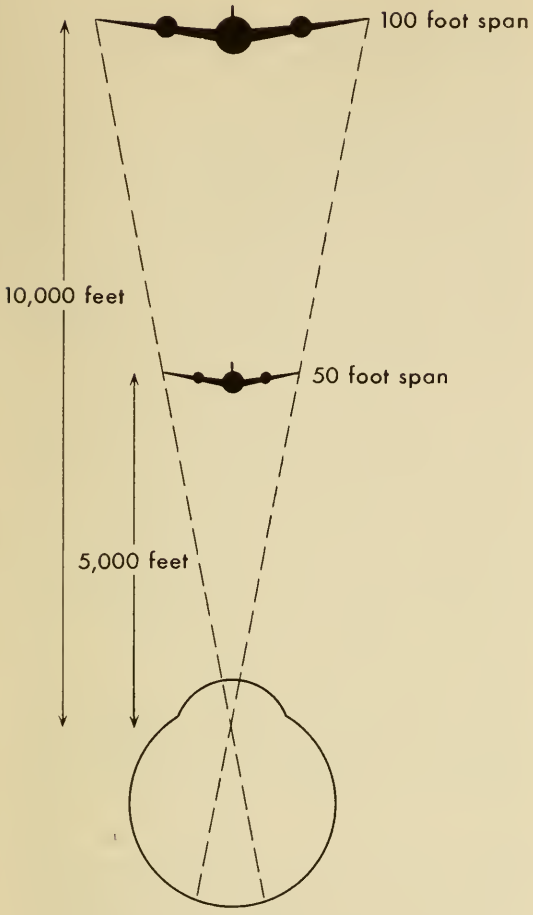


FIGURE 39. Two Airplanes with the Same Retinal Image

this rational kind of judgment is the only one possible in certain circumstances. Airplane spotters, for instance, can be trained to estimate the altitude of the craft they must identify and report, although their error is at best considerable. An airplane seen against the clear sky is not set into a background of continuous distance; it appears in the closest possible approximation to empty space. A plane with a wingspan of 100 feet at an altitude of 10,000 feet will be imaged exactly as large as a plane with a wingspan of 50 feet at 5,000 feet. If the silhouettes of the two were the same, no observer could tell one from the other by vision. But if the shapes differ, and one can be recognized as a large heavy bomber while the other is a smaller attack plane, both being familiar from repeated experience, the former is known – even seen – to be high, while the latter is perceived to be low. This type of distance estimation improves with training, and thousands of observers have had experience in it during the last war. The point to be noted is that it is not the kind of distance perception which occurs in the everyday visual world.

The Depth-Shape of Objects – Gradients of Texture and Grades of Illumination

The problem of the perception of distance is not really separate from the problem of the perception of depth or solidity. Granting that abstract empty space is irrelevant to either problem, the impression of a surface is the basic factor underlying both the experience of space and the experience of objects. Space is the visual background of objects and, when it is determinate, space reduces to the surface or surfaces which comprise the background. Objects are also defined by their surfaces. But the visual surface of an object has some characteristic features which we have not yet considered. It is necessarily either curved or bent in some way (this is the depth-shape of the object) and it is always delimited by a contour (the projected shape of the object). What are the stimulus gradients which might give rise

to a depth-shape, and what is the nature of a contour?

The Density of Texture and the Depth of Objects. The first problem is to account for the perception of what has been variously called relief or modelling of objects in the third dimension. The traditional explanation has been the stereoscopic effect of binocular vision, supplemented by the mental interpretation of the shadows formed on the side of the object away from the light. Little or no attention has been paid to the fact that the texture of the object varies in density according to the laws of frontal and longitudinal surfaces (see page 66). Imaged or projected texture varies with the slant of the surface, and with the facing of the slant – floor or ceiling, right wall or left wall. The analysis of texture-perspective suggests a general formula for bounded surfaces in the third dimension, whatever their inclination to the line of sight. It is that *the gradient of density in a projection of a physical surface bears a fixed relation to the slant and facing of the physical surface projected.* If this principle holds for a surface it holds for portions of that surface, and if the slant of a surface varies continuously or suddenly from point to point the projected gradient likewise varies continuously or suddenly. Here is a possible stimulus basis for the perception of a curved or bent surface.

An illustration has already been given of a gradient in the density of texture which yields a curved surface, Figure 37. What kind of a drawing would be expected to yield a bend or corner of a surface? If we can account for the curves, corners, and planes of objects we shall have gone a long way toward accounting for their depth-shape. A corner is geometrically the intersection of two planes with differing slants. An abrupt variation in slant corresponds to an abrupt change in the gradient of texture-density. Figure 40 represents such a change, employing a width gradient for simplicity in drawing. The change is from one rate of convergence to a slower rate, and the impression is that of a corner concave toward the observer. If the change had been to a faster rate of convergence the impression would have been one convex toward the observer. The locus of the discontinuity also gives the impression of a visual line even though no line is drawn. It may be thought of as a kind of stimulus for a corner.[3]

The contour line, at which one surface eclipses another, is often accompanied by the impression of a jump in depth whereas the inline of a joint or corner is not. The abrupt variation in brightness or color, which accompanies and ordinarily accounts for a contour according to the suggestion in the last chapter, does not account for this stepwise depth effect. Perhaps an abrupt variation in texture, however, can account for it – at least as a contributing factor.

Figure 41 is like Figure 40 in showing a width gradient. Instead of a change in

[3]There are evidently different ways in which a line may be evoked in vision. One should not be too preoccupied with the lines produced by pens and pencils. Besides the outline or contour of an object on its background there are the inlines produced by joints within a surface (Figure 38) and also, as we have just seen, by the corners of a surface.

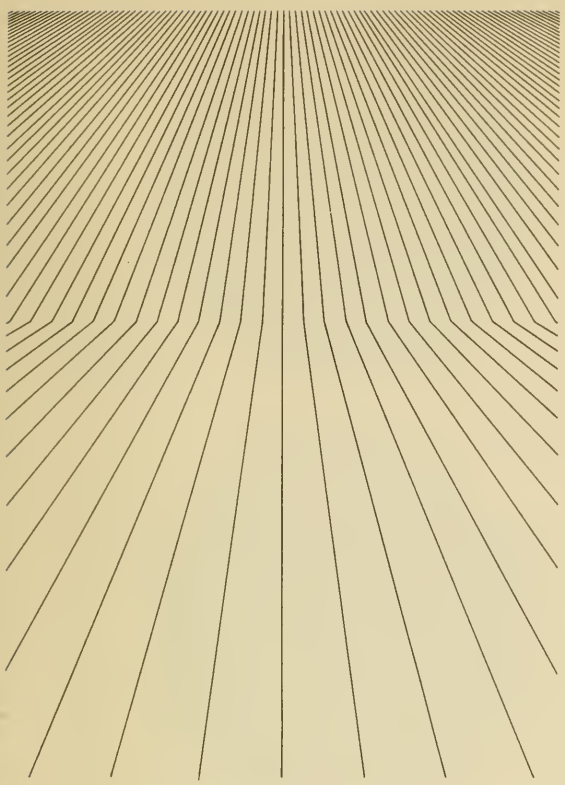


FIGURE 40. The Change of Gradient Corresponding to a Corner

FIGURE 41. The Jump Between Two Gradients Corresponding to an Edge

the gradient of density, however, there is a change in the amount of density with the gradient remaining constant on either side of the change. This discontinuity also gives the impression of a visual line even though no line is drawn. It differs from the previous drawing, however, in that the lower surface now appears in front of the upper surface with a jump in depth between the two. The picture looks like a floor ending at a step, or a plateau with the edge of a cliff and the country beyond. It seems likely that here is a contributing stimulus for the experience of a contour or outline with depth. Both a sudden change in density and a sudden change in the rate at which density changes will apparently produce a visual line, but of different types. A corner and a contour, although both are lines in the visual field, have separate contributions to make to the perception of depth in the visual world, one helping make an object look solid and the other making it stand out from the background. The perception of depth at a contour (particularly when it is enhanced by binocular vision) probably has a great deal to do with the impression that we can see empty space.

This evidence indicates that variations and changes in the density of texture are specific stimulus correlates of the planes and curves of an object, with their various slopes, and of the corners and contours

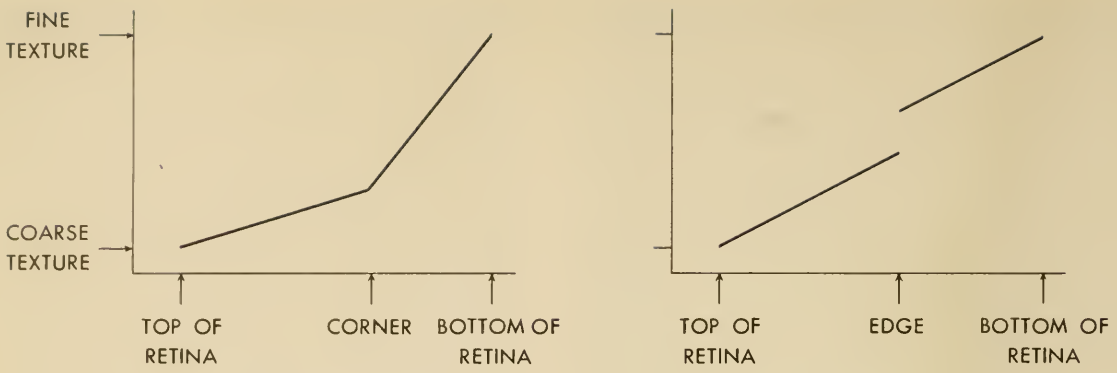


FIGURE 42. Plotted Variations in the Stimulus-Gradients at the Corner of an Object and at the Contour of an Object

of an object which separate it from the background as well. The retinal image is evidently richer in opportunities for depth-perception than the classical theorists realized.

Grades of Illumination and the Modelling of an Object or Surface. Painters have known for centuries that an object is modelled in perception by the light and shade over its surface. In present terminology, this means that the shading of a visual object can give it a depth-shape. Evidently it can do so independently of the density of the object's texture and the perspective of its edges, as Figure 43 illustrates. What is the logic of this effect? Much has been written about shading, for it is easy to produce with a pencil or charcoal, but little or nothing with a view to establishing the psychophysical correspondence, if there is any, between depth-shape and grades of illumination.

Excepting objects which are themselves sources of light, no surface can be seen unless it is illuminated in some degree. Considering the array of surfaces which fill the ordinary visual field with patches of color, some of these will usually have a high illumination and some a low illumination. The principal factor determining the degree to which a physical surface is lighted or shaded is whether it is directly or only indirectly illuminated by the light-source – the sun, let us suppose. A section of surface facing the sun is brighter than a section of surface facing away from the sun. The latter possesses what is called an attached shadow. There are also, of course, cast shadows which are projected on a facing surface by an object which intercepts the light, but these are less important for our problem.

The illumination of a given section of surface, then, is *a function of the orientation of the surface toward or away from the source of light.* A fact to be especially noted is that illumination is not a function of the distance of the surface from the observer. The physical world gets visually denser as it recedes, but it does not get either darker or brighter, and a distant area of uniform terrain is not different from a nearby area in this respect.[4] Grades of

[4] In the writer's opinion, some authorities have misapprehended the relation between brightness and distance, probably because they considered isolated points as the sources of light instead of surfaces. The question will be considered at the end of the next chapter.

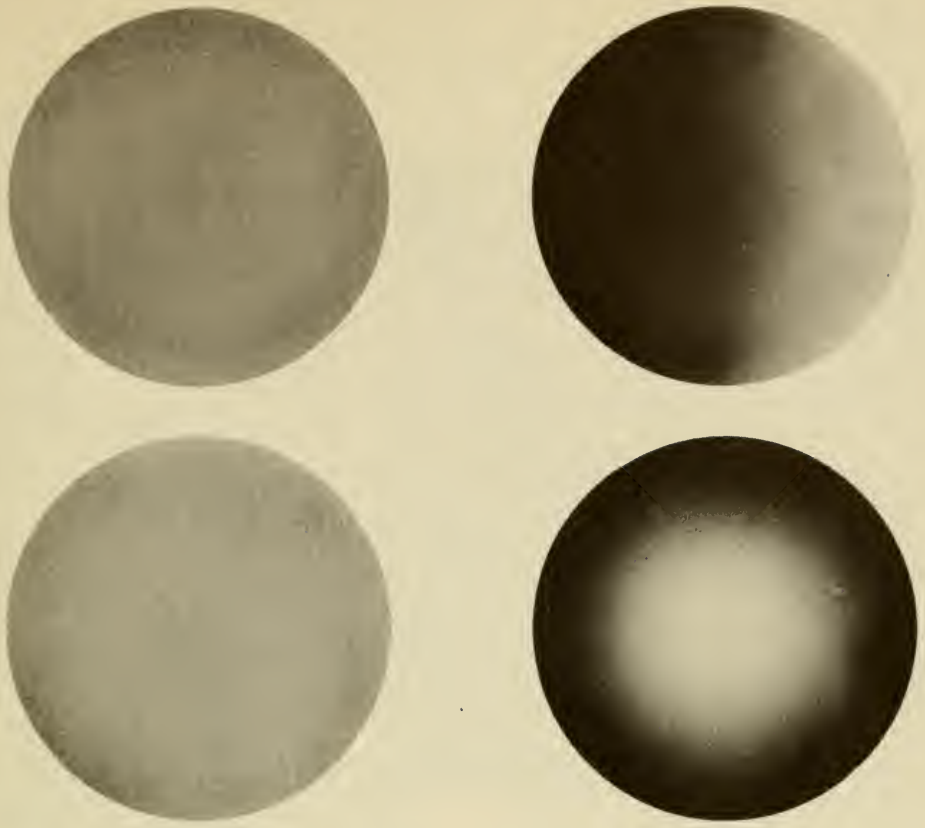


FIGURE 43. Gradients of Shading which Yield Depth

illumination, therefore, are not like the gradients of density, disparity, and motility with which this chapter is principally concerned. They are a correlate not of continuous distance or "space" but only of the depth-shape of objects and surfaces.

The depth-shape of an object seems to be that quality which, apart from its contour or silhouette, makes it object-like. This kind of shape must have as a minimum either a pair of surfaces making a corner or a curved surface. It therefore necessarily involves the opposite facing of adjoining surface areas. The elementary depth-shape (a pure abstraction like all elementary variables) would be a convexity or a concavity, both of which possess opposite faces. We may call the first a protuberance and the second an indentation. The three-dimensional shapes of things, abstractly considered, are made up of protuberances and indentations.

This statement is perfectly consistent with the analysis of three-dimensional surfaces in terms of slope; it only implies that an abrupt or a gradual variation in slope produces a sharp or a rounded junction of surfaces. The junction may be either a bend or a curve.

If one face of a protuberance is lighted the other is necessarily shadowed, and the same is true for an indentation. But the two shapes are distinguished by the order of stimulation. For, obviously, if the sun is in the south, then the southern

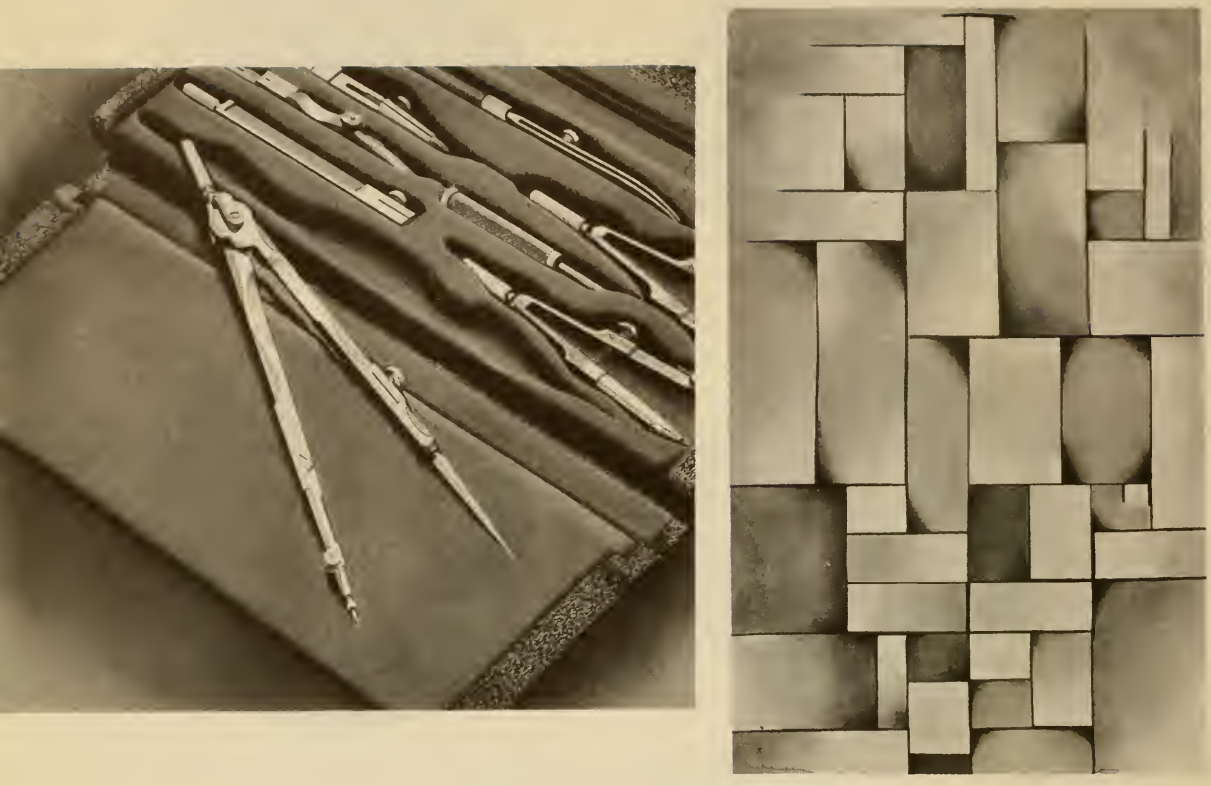


FIGURE 44. Shading which Yields Edged and Curved Surfaces

There is nothing more in either of these drawings than shades of black, gray, and white. Nevertheless we see in each a complex array of physical surfaces in three dimensions. But note that the *transitions* between shades are of different kinds: some are abrupt or "sharp", others are gradual or "rounded". Accordingly, we see *edged* surfaces in some places and *curved* surfaces in others. That the curved protuberances and identations of the instrument-case are produced by gradual transitions between light and dark one may verify by inverting the picture; transitions which formerly went "in" now come "out," or tend to do so whenever the factor of superposition does not inhibit this reversal. The relation of the drawing-instruments to their pockets is no longer precise but ambiguous. This rendering is to be contrasted with the abstract drawing on the right where the transitions have *intentionally* been arranged so as to make the depth-relations equivocal and non-representative, and therefore to make the space of the picture fluctuate in an interesting manner. *(Left: Rendering by Paul Madden. Right: From a painting entitled "Composition", by Van Doesburg. Courtesy of Mrs. Peggy Guggenheim.)*

face of a protuberance but the northern face of an indentation will be the one illuminated. A pair of adjoining lighted and shaded regions on a picture (and presumably on the retina) can yield the impression of a depth-shape. If the transition from light to shade is gradual the shape is a curve; if the transition is sudden the shape will be a corner. When convexities and concavities of a picture

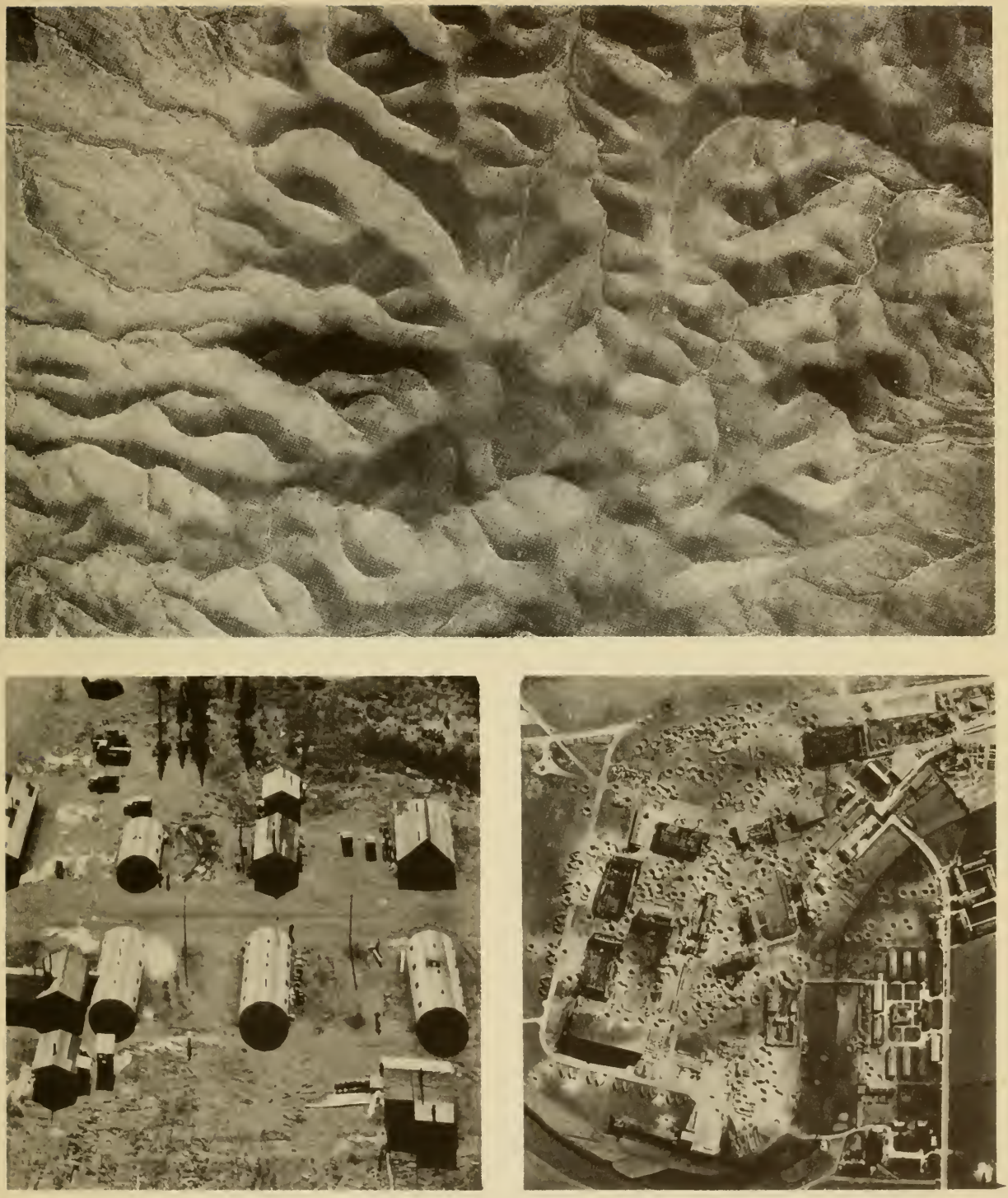


FIGURE 45. Shapes Which Are Reversed when the Picture is Inverted

(a) Hills turn into valleys. (b) Quonset Huts turn into towers. (c) Craters turn into mounds.

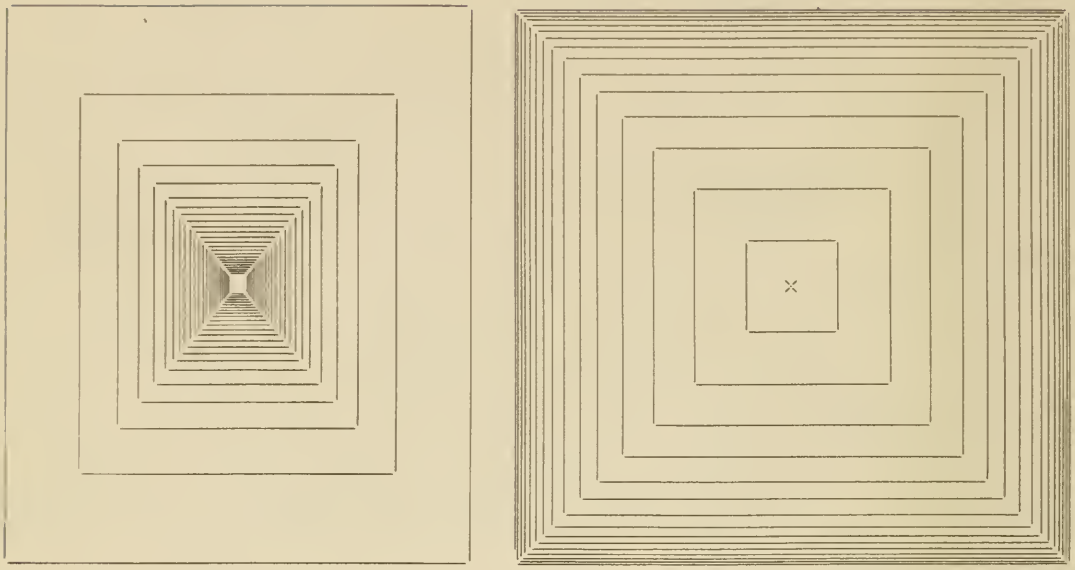


FIGURE 46. Reversal of Shape Due to Reversed Gradients of Density

are produced by light and shade alone, without any significant contribution by changes in the gradient of texture-density, it might be predicted that inverting the picture, which reverses the order of light and shade, would turn each protuberance into an indentation and vice versa. This effect can be observed in Figure 45. The hypothesis which suggests itself is that for any given visual field the order of projected light and shade is a contributing stimulus for an elementary shape in depth. If the order "lighted-shadowed" yields a protuberance in perception, then the order "shadowed-lighted" anywhere in the same field will give an indentation.

Assume that an observer in an average outdoor environment faces east. The sun will be in the south, if he is in the northern hemisphere and if the hour is not too near sunrise or sunset. Then the order shadowed-lighted in his visual field (from left to right) will correspond to a protuberance and the order lighted-shadowed to an indentation. But if he faces about and looks west, the order shadowed-lighted will be an indentation and lighted-shadowed will be a protuberance. The order of shading is a stimulus for depth, therefore, *only in relation to the orientation of the observer to his total visual world* or, more specifically, his orientation to the direction of the illumination. Is it conceivable that the way things face and the way the observer faces are reciprocally interrelated even in stimulation? The fact is

that when the observer turns around the convexities of an object are not converted into concavities. If, however, the observer maintains his orientation and the light on a surface in relief is experimentally reversed in direction without the observer's knowledge the protuberances tend to become indentations just as they do in an inverted picture. These inversions of depth occur, of course, only when gradients of texture are absent or ineffective. The latter gradients are never equivocal with respect to depth (Figure 46).

The reversal of relief in relation to illumination has puzzled and fascinated men for centuries. Convex and concave relief correspond to the coin, and the stamp that made it, or to the wax impression and the seal that produced it, both of which were familiar to the ancients. A modern instance of the phenomenon is encountered in the interpretation of aerial photographs. Boring's *Sensation and Perception in the History of Experimental Psychology* traces the investigation of the problem back to 1786 (12, pp. 266 and 304). Experimental investigation of the phenomenon is still needed.

The Perception of Objects as Such. A visual object in depth may be analysed in terms of several abstract variables which are interrelated with one another: the slope of its surfaces to the observer's line of vision, the orientation of its surfaces to the source of illumination, the corners or curves of its surfaces (either convex or concave), and above all, its outline or contour separating it from the background. To all these features of an object there correspond either variations in the density of texture of the retinal image, or variations in the intensity of the image, or both together. Insofar as the drawings have served as psychophysical experiments to the reader, it may also seem fair to conclude that the retinal stimulus variations yield impressions of depth, of slope, and of surface-shape which correspond to some of these abstract features of an object.

The gradient of texture, we may repeat, is a function of the slant of a physical surface away from the observer and the density of the texture varies with physical distance. Variation in shading, on the other hand, is a function of the physical orientation of the surface to the light source. It varies not with distance but with the curving or bending of the surface relative to the direction of illumination. Even the slightest curve or bend, insufficient to make much difference in a texture gradient, can produce a variation in shading if the direction of the light is favorable. Hence arises the capacity of light and shade to give what artists call relief to a surface, and to supplement the modelling of the surface in three dimensions.

In conclusion, the stimulus variations we have described need mathematical analysis, and the impressions of depth require verification. The mathematical analysis is possible, however, and predictions can be drawn up which should make experiments decisive. The theory suggested is far from being a complete explanation of the perception of material objects. Although retinal correlates have been proposed for shape in depth, for a contour, and for depth at a contour (with additional correlates still to be described), no theory has been ventured

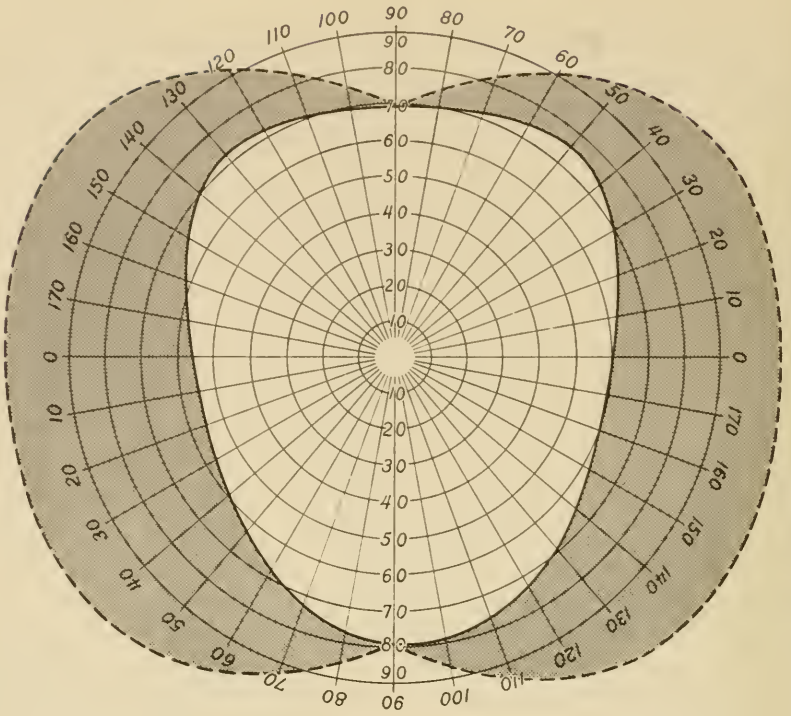


FIGURE 47. The Visual Fields of Each Eye and the Binocular Field, as Measured with a Perimeter

to account for shape without depth, or what is ordinarily called form. This is the problem with which the Gestalt psychologists began. Paradoxical as it may seem, the perception of shape without depth is more difficult to understand than the perception of shape in depth. We shall return to this problem in Chapter 10.

The Stimulus Gradient of Binocular Disparity Between Images

In addition to those differences between the visual field and the visual world which were described in Chapter 2, there is another difference which can be observed only with both eyes open. The visual field of both eyes is usually filled with double images. They are a characteristic of the binocular field (Figure 47) as distinguished from the monocular fields.[5] The double images are not easy to observe, since they

[5]The exact locus of all those points in the visual field which are not doubled will depend on the particular arrangement of objects or surfaces in the environment projected. The locus of all points in empty space which theoretically should not be doubled is quite a different matter. This latter is called the horopter, and a great deal of effort has been expended on this geometrical abstraction. Empirical determinations of the horopter do not agree well with theoretical constructions, however, and it can be argued that the horopter and the theory of corresponding points which it expresses do not clarify the problem of why we see objects in depth. In a surface theory of space perception the horopter is irrelevant. Carr suggests that the horopter is merely a geometrical curiosity (19). The clearest explanation of it in English is given by Troland (111, p. 342).

are mostly peripheral and always out of focus. If, however, one brings a finger into the central part of the field, continuing to fixate the wall of the room, the two fingers become obvious, one on the right and the other on the left. With a little practice in observing other things in the field which one is not focusing on, their doubled character usually becomes visible. In contrast to this appearance of the field one can see that the visual world as one looks at it uncritically is not in the least doubled, but is thoroughly single.

The doubled character of things in the field is a product of binocular vision, as the reader can verify by holding up his finger, closing and then reopening one eye. The monocular field of view appears single, clearly outlined, and "photographic" as compared with the binocular field. If the monocular field appears somewhat less deep than the binocular field, or to lack a special quality of depth, the implication is that the double images are symptomatic of that particular increment or quality of depth.

If you fixate the right index finger at arm's length and then move the left index finger toward your eyes, you can see the disparity of the double finger increase. You can also do the converse, fixating the near finger and moving the far one. The fact is that in the visual field as a whole, objects, edges, or surfaces appear doubled in proportion as they are physically nearer or farther than the point of fixation. In the central portion of the binocular field the disparity of visible objects and elements is a direct function of their distance along the line of sight from the point on which the two eyes converge. It is known that the stimulus for this visible disparity is *an optical disparity of the two retinal images*, each being a projection of the environment from a different position.

Retinal Disparity as a Function of Distance and as a Stimulus for Perception. Figure 48 illustrates this disparity of the two retinal images schematically. The scene is a road down which the observer looks toward the horizon. The image of the left eye is drawn in dashed lines and that of the right eye in solid lines. F indicates the point of fixation. The drawings are upright, as if the images had been projected on a picture-plane, and the displacement in one image relative to the other has been exaggerated. Assuming for the moment that the eyes are fixated on the horizon, which implies that their axes are parallel and without any convergence, it can be noted that the image of the left eye is relatively displaced to the right (or that of the right eye to the left) except at the horizon, and that the amount of disparity is inversely proportional to distance from the observer. This disparity is crossed, the image of the left eye being on the right in the visual field.

When the two images are combined by superposing one on the other, the result represents something like the total field of view of the two eyes. The binocular field, strictly speaking, is only the central portion of the total field, where the two monocular fields overlap, and this has been represented in the drawing. Only in this middle region are there double images. The disparity of the edges of the road increases as the distance of the road from

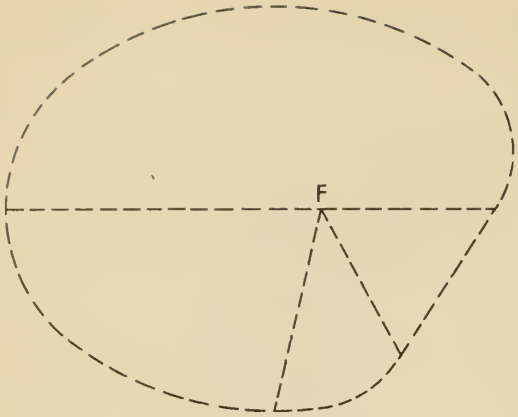


Image of Left Eye

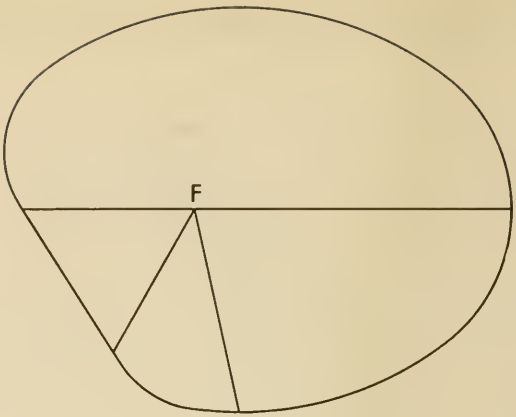
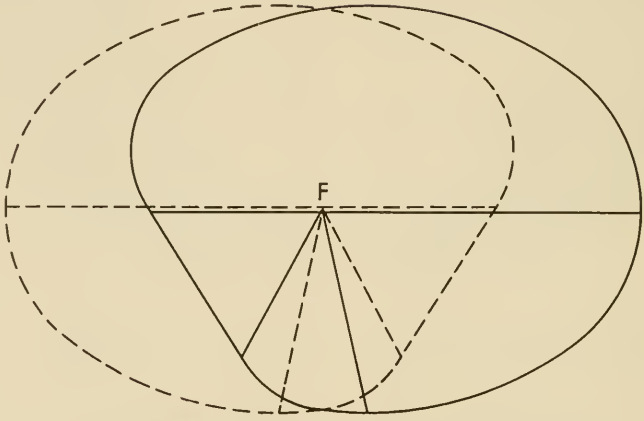
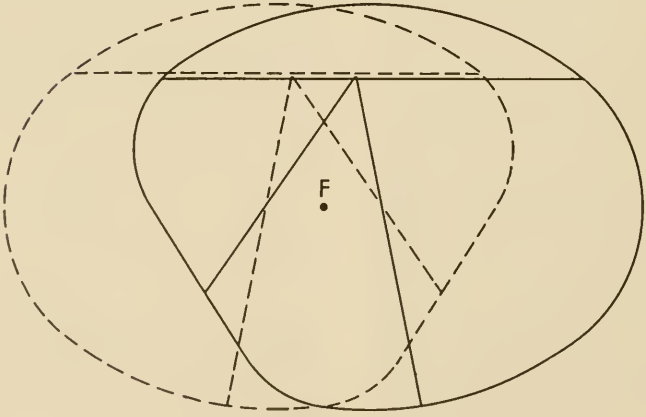


Image of Right Eye



Combined Images when Fixating the Horizon



Combined Images when Fixating a Near Point

FIGURE 48. Schematic Projections of the Retinal Images of the Two Eyes and their Combination, showing Disparity

The images are represented as projections on a picture-plane in front of each eye, not as if seen from behind each eye. Hence they bear some resemblance to the visual field. This expedient makes the disparity easy to visualize, but the relationship of inversion is left out of account. It must be remembered that retinal images as such are never seen by the animal that has them.

the observer decreases. This increasing displacement of the image toward the bottom of the visual field is, of course, a displacement relative to the other image, for there is no other standard with respect to which it could be displaced. It is in fact a gradient of disparity between the binocular images.

The diagrams represent only a single pair of edges in the environment since more could not be drawn without confusion. Actually the two images are usually composed of an array of texture-elements and contours, like a patchwork, each image being somewhat skewed or stretched horizontally relative to the other, as if it were printed on a rubber sheet which was then pulled in such a way as to transform a square into a parallelogram. These disparate images are united in the process of vision to form a single visual field even though they are not geometrically congruent. A fusion occurs between the neurological processes originating in each eye separately, and the visual world which results from this process is seen as a complete unity of objects in three dimensions.

If we no longer assume that the eyes are fixated on the horizon but are converging in some degree toward a spot or object on the ground, the combined images are represented by the lowest drawing in Figure 48. The center of each visual field has now moved downward so that the horizon is higher in the combined field than before. At the level of the point of fixation there is no disparity in the combined image and no double imaging in the binocular field. There is a gradient of disparity below the level of fixation, as before, but now there is also an opposite gradient of disparity above the level of fixation up to the horizon. In the latter gradient the image of the left eye is displaced to the left instead of to the right. This kind of disparity has been called uncrossed to distinguish it from the crossed disparity already described. In ordinary vision, with the eyes moving rapidly from one object of the environment to another, the disparities will shift with each new fixation. In this situation, we may inquire, is there any longer a clear relationship between degree of retinal disparity and distance of the environment?

The correspondence between disparity and distance is seen to be preserved if one takes account of the order of the disparity in the combined images, and considers the stimulus not as a geometrical picture but as an algebraic variable. Beyond the fixation point the projected disparity is a left-right one whereas in front of the fixation point it is a right-left one. The terms right and left are merely conventional; the essential fact is that disparity may vary in opposite directions. Physical distance from the observer to the horizon can be put into correspondence with a scale of minus to plus disparity as well as it can with a scale of minus to zero disparity.

When double images in the visual field are thought of as cues for distance perception and retinal disparity is conceived as falling into two types, a difficulty arises. How can the mind distinguish between a crossed pair of images and an uncrossed one, considering that the one looks like the other so far as any experimenter has been able to observe? In the

face of this difficulty, the fact remains that the impression of distance is not changed by a change in fixation, such as is represented by the two lower pictures of Figure 48. The difficulty is resolved if one treats binocular disparity as a stimulus gradient. A rate of decreasing crossed disparity can be the equivalent of a rate of increasing uncrossed disparity in the same way that the numbers -3, -2, -1, 0 are equivalent in order and rate to the numbers 3,4,5,6. It is a reasonable hypothesis that some graded process in the brain of the observer reacts to the disparity of his binocular images in just this fashion. The rule would be that an impression of increasing depth on a surface is in psychophysical correspondence with any gradient of disparity running in the direction toward uncrossed, and that increasing nearness is in correspondence with any gradient running toward crossed disparity. The rule assumes the conventions of Figure 48 and refers, for convenience, to the picture-projection of each retinal image, not the retinal image itself. It can be generalized to include the inclined surface which we have called a ceiling scene as well as that of a ground scene, and a wall scene looked at from the left as well as one looked at from the right. These types of slanting surface, it will be recalled, are as fundamental to the perception of object-surfaces as they are to the perception of surfaces of the environment.

Figure 49 illustrates these stimuli. A gradient of binocular disparity which runs horizontally across the combined images corresponds to a "wall scene." If the uncrossed disparity of the two views increases to the right, the wall increases in depth to the right. This yields the left hand wall of a corridor, or the left side of an indentation. The pair of pictures at the top have this gradient of disparity, and yield a surface of this sort when they are viewed in a stereoscope. If uncrossed disparity increases to the left, the wall increases in depth to the left, as does the right hand wall of a corridor or the left side of a protuberance. The same pair of pictures, when interchanged right for left, have a reversed gradient of disparity and will yield the latter type of surface in a stereoscope.

A gradient of disparity which runs vertically up or down the combined images corresponds to a surface which slants as a floor or a ceiling does. If the uncrossed disparity of the two views increases upward, as it does in the bottom pair of stereoscopic views, the surface slants into the distance as it goes up (like a floor); if it increases downward, as it would if the bottom pair of stereoscopic views were interchanged, the surface slants into the distance as it goes down (like a ceiling). These are the slants which actually appear when the pair of views is presented in a stereoscope. It may be noted in a stereoscope that, since the lines do not have a gradient of monocular perspective, the floor (or ceiling) gives the appearance of boards which get wider as the distance increases. The same types of surface may equally well be obtained with a pair of views composed of dots or short vertical line-segments.

The aniseikonic glasses devised at the Dartmouth Eye Institute for experiments on stereoscopic distortions of space and recently described by Ames (1) are re-

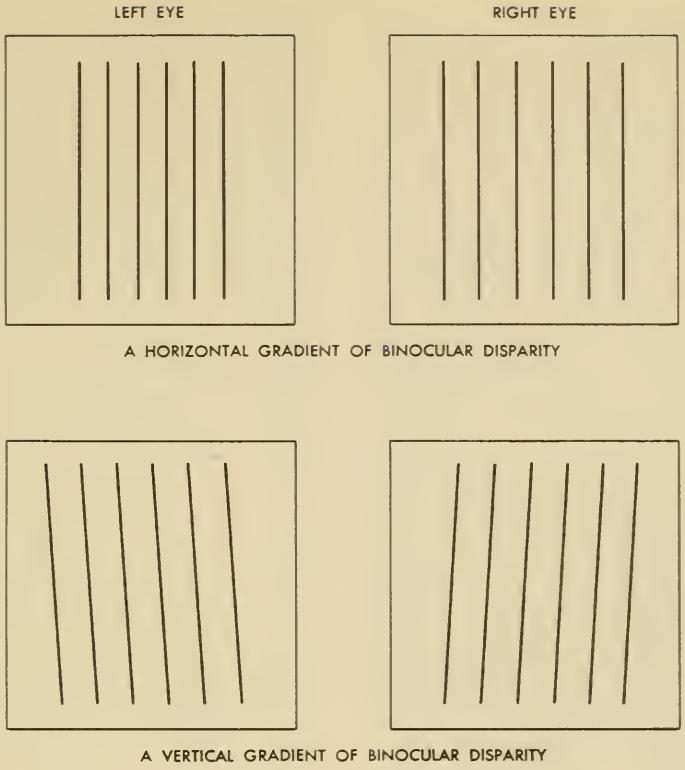


FIGURE 49. Gradients of Binocular Disparity in Relation to the Fundamental Types of Slanting Surface

lated to the present theory. They have the effect of magnifying or stretching the image of one eye along one dimension only or, in other words, of increasing the relative size of one retinal image on a single meridian. The lens employed and the disparity produced have been reported by Ogle (86). These glasses should have the effect of increasing the gradient of disparity over certain physical surfaces and thereby increasing the apparent slant of those surfaces. Allowing for the influence of conflicting gradients of texture and perspective, the changes in the apparent slant of a physical surface which the lenses produce when an observer wears these glasses ought to be predictable from the above rules. The validity of the rules may be tested by their success in making the predictions.

In general a vertical gradient of disparity can be described as a horizontal skew of one image relative to the other. A horizontal gradient of disparity can be described as a horizontal stretching of one image relative to the other. The disparity is always horizontal whereas the gradient may be either vertical or horizontal. These relative distortions of the image are not to be confused with a rota-

tion of one image relative to the other, such as would occur in cyclotorsion of the eyes. An image which has been skewed or stretched has undergone a non-rigid transformation, mathematically speaking, whereas a rotation, like a simple transposition of an image, is a rigid type of change. These two types of distortion are also not to be confused with an enlargement or magnification of one image relative to the other, which is the original meaning assigned to the term aniseikonia at the Dartmouth Eye Institute. A relative enlargement makes the two images disparate in quite a different way from the one defined above: the disparity is not horizontal, it does not fall into a one-dimensional gradient, and it would not occur in natural vision. The effect of relative enlargement, indeed, appears to be equivocal and not completely understood.

The image of the nose, that unnoticed but important feature of every binocular visual field, is actually a crossed double-image. It is, in fact, the ultimate limit of crossed double imagery – the end of the gradient of disparity in the crossed direction. Between the disparity of the nose and the disparity of the next color-patch above it in the visual field – the ground or floor nearby – there is always a very considerable step or jump. A discontinuous step in an otherwise continuous gradient, we have suggested, is the stimulus for the impression of depth at a contour. Between the nose and the next visually adjacent surface, therefore, a considerable appearance of depth should occur. This depth is probably the basis for our impression that we see empty space between ourselves and the nearest object, and for at least part of our impression that we look out upon the world. It is important to account for these impressions since, unexplained, they lend plausibility to the prevalent but sterile conceptions that the perception of depth reduces to the perception of an abstract third dimension and that somehow visual sensations are projected outward from the eye in perception.

Binocular Disparity and the Depth of Objects. Figure 50 illustrates the way in which two projected views of a solid ob-

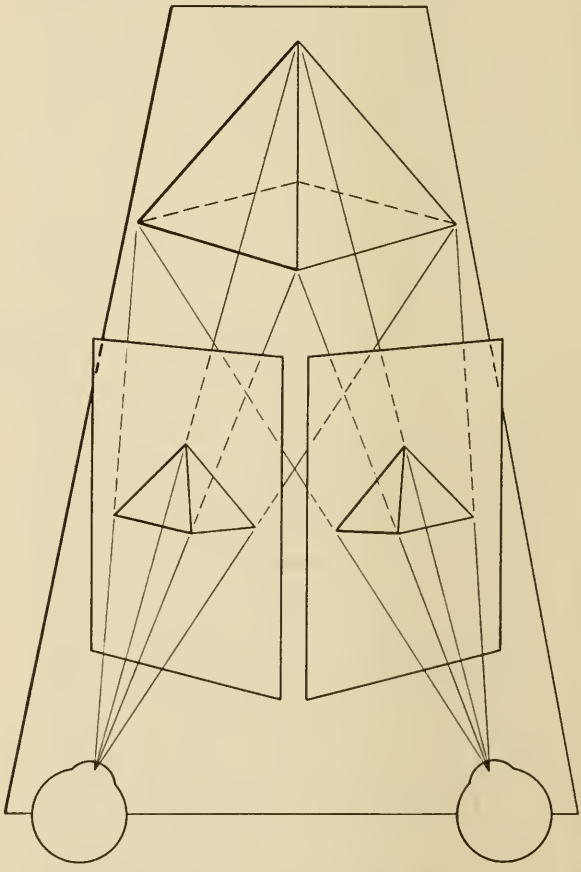


FIGURE 50. The Disparate Views of an Object by the Two Eyes

ject are disparate for the right and the left eye. One projected shape is skewed horizontally relative to the other. When the two projected shapes are fused in binocular vision a depth-shape results, the depth being in correspondence with the amount and direction of skewness. The figure illustrates at the same time the principle of the stereoscope, the invention of which provided Wheatstone with a dramatic proof that binocular disparity was a cause of depth perception. The stereoscope is simply a device which permits the two views of Figure 50 to be projected independently on the two retinas.

The drawing represents the edges and corners of the solid object, as most drawings do, but not its textured surfaces. The gradients of disparity with respect to texture are not evident in the illustration. In the binocular vision of a real object, or when stereoscopic photographs of a natural scene are used, the gradients of disparity on its various surfaces will produce the kinds of slant we have already described and will have the effect of giving the object a shape in depth. The gradient of disparity on each surface area is, of course, concurrent with a gradient of the density of its texture and a gradient of perspective (if the edges are parallel). The change in the gradient of disparity at an edge or a corner, moreover, is consistent with the change in the gradient of texture and with the change in the grade of shading from one plane to another. In ordinary vision the gradients supplement one another, it must be assumed, so that even if the texture is indefinite, the object irregular, the illumination flat, or the observer blind in one eye, there will be some stimulus basis for the immediate experience of the scene in depth.

The most obvious kind of depth in stereoscopic photographs, especially of the parlor variety, is the depth which appears at the contours of objects – the depth which makes objects stand out from the background and which gives one the impression that there is empty space between the object and whatever is behind it. The stimulus for this impression can be ascribed to an abrupt change in the disparity of fusing elements, or a change in the amount of disparity without any gradation. It is to be distinguished from the gradual change in disparity which is characteristic of a sloping surface, and from the abrupt (or gradual) change in the gradient of disparity which is characteristic of a bend (or curve) in a surface.

The Importance of Binocular Vision for Depth Perception. It has been commonly believed for many years that the only important basis for depth perception in the visual world is the stereoscopic effect of binocular vision. This is a widely accepted opinion in the medical and physiological study of vision, opthalmology. It is the belief of photographers, artists, motion picture researchers, and visual educators who assume that a scene can be presented in true depth only with the aid of stereoscopic techniques, and of writers and authorities on aviation who assume that the only kind of test for depth perception which a flier needs to pass is a test of his stereoscopic acuity. This belief is based on the theory of the intrinsic cues for depth, which is rooted in the assumption that there exists a class of experiences called innate sensations.

With the increasing tendency to question this assumption in modern psychology, the belief is left without much foundation. Depth, we have argued, is not built up out of sensations but is simply one of the dimensions of visual experience. The accepted belief is contradicted by the fact that persons having vision in only one eye, or temporarily limited to the use of one eye, see the visual world in depth just about as the rest of us do and get about in it without conspicuous loss in efficiency. There have even been one-eyed fliers, of whom Wiley Post was the most notable. It is also contradicted by the fact that many animals who do not possess overlapping binocular fields and who therefore lack disparate images – rabbits and rats, for instance – seem to discriminate depth and distance accurately since their behavior is nicely gauged with reference to it.[6] The implication is that the emphasis on binocular disparity as a cue for depth has been exaggerated.

The theory of retinal gradients as stimuli for visual depth implies that the gradient of disparity is only one of several, and that all of them are effective in everyday vision. The effect of any one may be experimentally isolated from that of the others, as we have tried to illustrate, but ordinarily they function as concurrent stimuli. Whether one or another of these gradients is more important for the resulting perception is something that can be determined only by experiment, and something that will probably vary with the conditions of stimulation. If the gradient theory is correct, binocular vision simply takes its place as a determinant, but only one determinant, of visual space.

Definition or Resolution in the Retinal Image

A disparity of the image in one eye relative to the other presupposes that both images are textured. Only with respect to a pattern of spots or lines could a disparity exist; there must be something to be incongruent. The homogeneous blue sky would not yield binocular disparity because it would not produce a patterned image. Binocular stereoscopic vision, therefore, depends on textured vision. The very mechanism by which the single eye gets a focused image, the reflex response of accommodation, must have something to do with texture since focusing can be thought of as a maximizing or sharpening of texture. What we call distinct or clear vision must be related to the texture of our impressions. Are all these facts interrelated? If so, what does optics contribute toward understanding them?

The Problem of Clear Vision and Visual Acuity. It is a fact, so far neglected in this discussion, that our vision of the

[6] Binocular stereoscopic vision has probably evolved as a sort of alternative to panoramic vision in which each eye has its own field of view with little or no overlapping. The primates, apes and men, have forward pointing eyes with coordinated eye movements, with convergence of both eyes on the same object, and with a total field of view approximating 180°; but many other mammals, specially those preyed on by carnivores, have eyes which point each to its own side, each moving independently, with little or no convergence and with a total field of view approximating 360°. The suggestion is that, in the course of evolution, the primates sacrificed the ability to see all the way around at once for an enhancement of the ability to discriminate depth (23).

world must be clear if it is to be useful. Anyone who suffers from nearsightedness or an alcoholically induced imbalance of the eye-muscles will agree. So also would anyone who has to drive a car in a dense fog. Clear vision is easier to describe than to account for, depending as it does on three separate sets of conditions, physical, optical, and neurological (see Chapter 4). Descriptively or psychologically, a scene is clear when it is sharp, or detailed, or definite. It is not clear when the impression is fuzzy, foggy, vague, dim, indeterminate, or blurred.[7]

The ability to see very distant objects or very small objects is called acute vision. Such ability is important in a number of human occupations, and is also theoretically interesting because there is a limiting size and a limiting distance, the two being interrelated, below or beyond which objects become invisible. A good many tests of visual acuity exist and a great many experiments have been performed on the conditions which affect it. The artificial scenes which have been set up to measure this kind of fine discrimination are interesting, and they may help to specify what has here been called visual texture.

These are some of the ways in which acuity can be measured:

1. Determine the smallest familiar form that can be recognized, starting with large sizes and proceeding to smaller ones. Block letters are usually employed. This is the familiar eye-chart devised by Snellen, and modified by others.

2. Determine the smallest form among a set of identical forms rotated to different positions which can be seen in correct orientation. For example, what is the smallest broken circle, or Landolt ring, for which the location of the gap can be reported? This test eliminates the factor of meaning or familiarity.

3. Determine the smallest noticeable interspace between two dark objects on a light background, just before they merge into one. Two points cannot be used unless they are big enough to be each visible, so the commonest practice is to present two rectangular bars side by side.

4. Determine the smallest black dot on a light background which can be reported as there.

5. Determine the thinnest black line on a light background which can be reported as there.

6. Determine the finest or most dense grid of parallel black lines on a white ground for which the lines can be distinguished as running horizontally or vertically. Such a device is an Ives grating.

7. Determine the finest or most dense black-and-white checkerboard which can be distinguished from an equivalent area of gray.

8. Determine the just noticeable misalignment of the two segments of a broken straight line. This is called vernier acuity. Our great sensitivity to such

[7] All the above adjectives are properly applied to perception only. They do not adequately describe the external flux of light, they certainly do not describe a process in the nervous system, and they probably should not even be applied to the retinal image when this is conceived strictly as a stimulus, that is, when the image is thought of as a complex of variations rather than as a picture to be looked at.

misalignment is an advantage in the reading of scales and pointers.

9. Determine the just noticeable curvature or bend in a straight line (or a straight contour). Our sensitivity to a deviation from the rectilinear is also great.

10. Determine the just noticeable disparity between the image in the right eye and the left eye in a setup viewed with both eyes. The effect perceived will be one of depth. This is called stereoscopic acuity.

11. Determine the just noticeable difference between the gray of a solid shape and the gray of its background, as evidenced by whether a contour (and hence the shape itself) is perceptible. This is usually called brightness-discrimination, although there is evidence that it is related to the other modes of acuity (103,6).

The Relation of Texture to Acuity. The sixth and seventh of the experiments or tests listed above demonstrate that fineness or density of visible texture is a measure of the acuteness of vision. Both a grid and a checkerboard are cyclical variations of dark and light. What do the remaining patterns suggest? Omitting the first test, the second and third involve the impression of a gap or interspace. However a gap may be defined geometrically, it can be considered a fundamental component of a texture. The fourth and fifth tests employ a spot and a dark line respectively. A line, of course, is a sort of elongated spot. Spots may be components of a texture and streaks or inlines (as distinguished from outlines) may also be components of a texture. The eighth and ninth tests, however, are of a different sort. They involve deviations from the straight quality of a contour. These deviations seem to be elementary impressions of shape rather than elementary components of texture. They have to do with the outline or margin of a physical surface rather than with its texture. The tenth test, of binocular disparity, is complex. The eleventh, however, points to an interesting conclusion, that a difference in brightness cannot be sensed without an accompanying impression of a margin (6). The stimulus for a margin seems to be a relatively abrupt gradient of intensity in the retinal image. Is it possible that the microgradient of intensity is the fundamental stimulus underlying not only the phenomenon of a margin or contour but also the phenomena of texture, visual acuity in its different modes, the focused image, and clear vision in general? Indirectly, if this were true, it would be the basic stimulus for the phenomena of edges, surfaces, shapes, and objects. Before answering this question we must consider a number of other problems.

What are the Conditions for Clear Vision? Sharp contours and definite textures in vision depend not only on accommodation of the lenses and normal eyes (the dioptric mechanism) but also on external or physical conditions and on internal or neurological conditions. Externally there must be reflecting substances and a clear medium, that is, one which does not scatter the hypothetical rays of light, as fog does. Internally there must be a fully matured nervous system and a dense mosaic of retinal receptor-cells. Disease or injury to the occipital brain, the optic tract, the nerves,

or the retina, will all reduce visual acuity. The system is intricate and apparently does not fully develop in children until 7 or 8 years of age inasmuch as visual acuity, poor in infancy, continues to improve until then. Between the external light and the internal nervous process there stands the mechanism of the eye itself. Here is where the most common defects are found which reduce acuity.

A primary function of the eye is to form a focused image, one in which there is experienced a minimum degree of so-called blur. A focused image depends on accommodation. Blur may result from a whole complex of anatomical and physiological defects, the simplest results of which are nearsightedness, farsightedness, and astigmatism. These particular effects can be compensated in part by supplementing the lenses of the eyes with artificial lenses.

Accommodation and Convergence as Cues for Distance. It will be recalled that as early as 1709 Bishop Berkeley believed that accommodation of the lens, together with the convergence of the eyes, furnished a sign for the distance of the object fixated and therefore gave it the appearance of existing in a third dimension. The precise cue employed would have to be the muscle-sensations involved in these reflex adjustments. The theory has persisted to the present day. After the invention of the stereoscope by Wheatstone in 1833 and the discovery of the binocular disparity between images, these muscle-sensations lost some of their importance as cues for depth, but the problem continued to attract interest (19). Present evidence, however, makes it doubtful that they furnish any data for depth-perception (121, p. 665-680).

Both the bulging or flattening of the lens and the pointing of the eyes are reflex adjustments. As reflexes they must have a stimulus. What has to be understood is that this stimulus cannot simply be light; it must be a condition of the retinal image with respect to order or pattern. Fundamentally this condition is geometrical. What is the sense, then, in seeking an explanation for depth in the adjustments of the eyes to a stimulus when the stimulus itself is something from which the explanation may be derived? Accommodation and convergence are responses of the eyes to a condition of their images (blur and disparity) which may concurrently produce that inner response we call "depth".

Blur as the Stimulus for the Reflex of Accommodation. The lens mechanism seems to operate on the principle of minimizing the condition in the retinal image which yields blur, not over its whole extent (for this would normally be impossible) but at the fovea. Since blur may result from either too thin or too thick a lens, the process is in all likelihood a sort of back-and-forth or trial-and-error one, not unlike that of focusing a lantern-slide projector on a screen. The resulting image has what is known in optics and the study of acuity as a maximum of definition or resolution. It seems possible that these terms can be analysed geometrically.

If we are forced to suppose that the eye is sensitive to stimulation of this sort in its reflex functioning, there is surely reason to believe that it can be similarly sensitive in its higher functions. The

elementary impressions of surface and edge, or texture and contour, may plausibly result from stimulation of this sort. If a reflex reaction can be a response to a geometrical condition of the image, so also may a quality of experience.

Is the Gradient of Blur a Cue for Distance? The accommodation of the lens which eliminates blur at the fovea necessarily produces blur in outlying regions of the image when the environment is a terrain or has a floor or walls. Physical parts of a surface at greater or lesser distances than the point of fixation will tend to be blurred. Does this constitute a gradient of increasing blur over the surface similar to the gradient of density of texture and the gradient of disparity? If so, might it be a supplementary stimulus for the impression of continuous distance?

The quality of blur does grade off from the center of the visual field, but it does not seem to change character in opposite directions. Lines or spots in the visual field have the same fuzzy character whether their objects are nearer or farther than the object in focus. The phenomenon may be observed in photographs. It is probably not, therefore, a univocal gradient like the gradient of crossed to uncrossed disparity with which it is always associated in looking at ordinary surfaces. Whether it could be experimentally isolated is doubtful; the writer's efforts to obtain a simple gradient of blur on a photograph were not successful. It can hardly, therefore, be an effective independent stimulus for distance.

The fact is, moreover, that when the eye is accommodated for any distance beyond a few feet the gradient of blur becomes a very minor matter. All of the field except that directly under the nose tends to be uniformly in focus; the gradient levels off. This results from the small aperture and small size of the eye considered as a camera; the depth of field (depth of focus in the photographer's terminology) is very great. The working limitations of the camera in this respect scarcely apply to the human eye. Hence it is, probably, that a gradient of texture over the field can be sensed in a single fixation without much interference from blur.[8]

The Definite, Resolved, or Focused Image. It is now possible to return to the problem of the relationship between acuity and texture. We are told by the students of the subject that visual acuity is defined as the smallest visual angle that can be resolved or the smallest shape which has definition. These terms are not perfectly satisfactory by themselves for, unless they are analysed geometrically, they come down to this: that acuity is what the tests of acuity test.

The geometrical analysis of an unblurred image provided by optics employs the concepts of ray and point, of the focal point, and of the circle of confusion. With a perfect lens and a perfect image, all light from a single point converges to a very minute point in the image. With other conditions a pencil of rays forms a more or less extended figure in the image instead of a point, the circle of confusion, and these overlapping circles account for blur.

[8] The writer is indebted to Professor Nora M. Mohler of Smith College for the computations on which these conclusions are based.

This kind of analysis is not the only one which is mathematically possible, however, and it seems likely that analysis in terms of area and intensity within the two dimensions of an image would be more profitable for the study of vision. From this standpoint, a focused image would be one in which the transitions from dark to light were as abrupt as the adjustments of the optical system permit. This analysis begins not with geometrical points but with gradients (in this case microgradients, page 73) and their slope. The formula would be that definition or resolution of an image is the degree to which the gradients of light intensity within it are as steep as the relevant conditions permit.

If this formula will account for the physically focused image – for the external and the optical set of conditions of clear and acute vision – it may help us to understand the dependence of acuity on the retina and the brain, the neurological conditions of clear vision. Most of the research on visual acuity by psychologists has been concentrated on this question (95), but it is outside the scope of our discussion.

According to this formula, the steepest gradient of intensity would be the condition corresponding to the sharpest visual contour. A series of alternating gradients as steep as possible would be the condition corresponding to a clearly perceptible texture or surface, both in the case of a surface with gross details and in the case of one with fine microstructure. The ability to see a very small spot, gap, or streak, as this is measured in tests of acuity, would also depend on the formation of an image with the maximum steepness of the pair of gradients involved, that is to say a focused image.

In all that has gone before, full illumination has been taken for granted. Acute vision, however, depends on a bright image as well as on a focused image; the ability to see detail falls off rapidly as the light reflected from objects grows dim. The above formula is also consistent with this dependence of acuity on illumination, inasmuch as the gradients of intensity in an image become less steep as the intensity of its brightest spots is lowered.[9]

The outcome seems to be that the texture, detail, and pattern of our visual perceptions, on which their spatial character depends, are themselves dependent on steps of luminous intensity in the retinal image. The large scale gradients of detail with which this chapter is concerned are based ultimately on small scale gradients of the intensity of light. Although such a formulation of the matter is tentative and incomplete, it has the virtue of bridging the gap between optics on the one hand and visual perception on the other.

Ernst Mach discovered long ago, by fusing a combination of black and white on a rotating disk, that a regular and uniform increase of light-intensity along a given dimension of the retina did not yield an impression of gradually increasing brightness as one might expect (77, p. 217). An impression of abruptly increasing brightness, a margin or border in other words, could be produced much more easily. The latter impression occurred

[9] The experimental evidence on acuity and on the relation of small areas to intensity is referred to in Bartley's survey (6) and in a recent review by Senders (95).

not only with an abrupt increase of light-intensity along a dimension of the retina, but also with an abrupt change in the *rate* of increase along a given dimension. This is the meaning of the Mach rings. The principle illustrated in Figures 40, 41, and 42 appears to be a fundamental one which applies to gradients of intensity on the retina as well as to gradients of the density of a texture.

The principle seems to be that a visual margin or border is given in experience by either a step in the level of luminous intensity (a steep microgradient) or by a step in the first differential of luminous intensity (a change in the rate of change). These two seem to be equivalent as types of ordinal stimulation. The retina responds to an abrupt change of change as readily as it responds to an abrupt change. Abruptness seems to be the critical condition. Perhaps the function of the retina has been misconceived in developing the theory that it responds to light rays and their directions. It seems to respond to gradients and their differentials instead. The retina is probably to be conceived as an organ of the body which is sensitive to grades of light, not points of light.

The Gradients of Aerial Perspective

In the course of the evolution of human vision, we might conjecture, all the existing variations within the retinal image have been utilized as stimuli for perception if they are consistently in correspondence with the actual lay of the land. Several of these projected gradations, differentials, and shadings have been described. One more which deserves mention is the gradient, or complex of gradients, known as aerial perspective. As a phenomenon it has been known to painters since Leonardo da Vinci first described and named it four centuries ago. The color of a landscape which stretches off toward the horizon becomes bluer and more filmy with increasing distance. The fact is that the retinal image of such an environment is constituted by pencils of light which have travelled through differing amounts of air. At the extreme top of the image the light may have passed through only a few feet of atmosphere, whereas at the lower margin (corresponding to the horizon) the light will have passed through many miles of atmosphere. Owing to differential scattering of wavelengths, the color of the light reaching the eye differs under these two conditions, even though the color of the reflecting surfaces may be identical. The simplified explanation sometimes given is that just as the sky itself is blue, so also the hue of the hills near the horizon is shifted from its proper quality toward blue in proportion to their distance. The gradient of this shift in color might be considered analogous to the gradients of size or density of texture.

We cannot be sure, however, that this increasing blueness in the visual field is a true stimulus gradient unless it can be shown independently to yield an impression of continuous distance. The same thing would be true of a gradient of haziness. No such experimental demonstration has been made; aerial perspective has apparently not been studied in isolation. The change of color is not simple and is comparatively slight as compared with the full range of color variations. It is unlikely that we can discriminate as

Bahnsen, from Monkmeyer

FIGURE 51. Aerial Perspective

many grades of this color change as we can grades of, say, texture density, and it is therefore doubtful that the color gradient could prove to be as effective a stimulus as the texture gradient. Aerial perspective depends on the amount of haze in the atmosphere, and may differ from day to day. It is certainly not discriminable in the immediately surrounding environment. In the absence of evidence, it would be safest to proceed on the assumption that a gradient of aerial perspective is not a stimulus in the proper sense of that term. Perhaps its function is more nearly what it has long been assumed to be – a cue, indicator, or sign of distance. If, as seems possible, the gradient is not always consistently in correspondence with the actual lay of the land, it is only reasonable to suppose that it should suggest the impression of distance, as red suggests warmth, without compelling the impression in the way a stimulus is supposed to do. The question, however, is a matter for experiment.

Aerial perspective is not intrinsically

related to the surfaces of the environment in a geometrical way as are linear perspec- and texture perspective. Neither are gradations of shading intrinsically connected with the geometrical shapes of surfaces and objects; they are only indirectly connected by way of their orientation to the source of light. The advancing and receding colors which painters employ to bring an object out from or back behind the picture-plane are also not connected with physical depth by any clearly understood principle. One might speculate that variations in hue or shading as such do not produce the same compelling impression of depth that gradients of texture, line, size, binocular disparity, and motion produce, just because they are not related to physical depth by geometrical laws as the latter are. Variations in hue and brightness can and do produce compelling experiences of outline, form, and pattern in the two dimensions of extensity, but their correspondence to experiences of solidity, depth, and distance is less precise.

Summary

The seeming poverty of the visual stimulus as compared with the richness of visual experience has apparently been exaggerated. Even the static momentary retinal images with which this chapter has been concerned appear, when analysed, to be adequate to account for the depth and distance of the visual world without the necessity of supposing a special mental process to supplement the images. The retinal image is an exact and intricate event, deserving of more respect than it has usually been given.[10] It is necessary only to give up the expectation of finding in it replicas of the experience we wish to explain and seek correlates instead. What right did we ever have to assume that the retinal image must copy the world, or that the form of an experience must be duplicated by the form of its image? We understand well enough that a visual stimulus is neither an object nor an experience of that object, but something which stands between them. What we have failed to understand is that this stimulus need not look like either its cause, the object, or its effect, the experience. It need only be a specific correlate of both.

Some of the previously unexplained features of the visual world are accounted for if we suppose the retinal image in each eye to be an array of steps, gradients, and changes in gradients of light-variation. Variations in texture and size, in shading, and in binocular disparity have now been described in these terms. They are in exact geometrical correspondence with the dimensions of the physical world, and there was evidence that they yield corresponding variations in perceptual experience.

---

[10] According to a leading textbook of ophthalmology by Duke-Elder (30, p. 764), the defects of the normal eye as an optical instrument are few. The most important kinds of distortion and even aberration found in ordinary lens systems are effectively corrected in the human eye. This fact is in contradiction to a fairly widespread impression that the eye is a poor optical instrument.

7

The Stimulus Variables for Visual Depth and Distance -- The Active Observer

The Gradients of Deformation of the Image During Movement of the Observer . . . . The Types of Retinal Motion . . . . Summary — The Sensory Analysis of Distance and Depth

Heretofore we have been talking about visual perception for the most part as if the observer stood motionless in the environment and kept his head fixed in one position. The normal human being, however, is active. His head never remains in a fixed position for any length of time except in artificial situations. If he is not walking or driving a car or looking from a train or airplane, his ordinary adjustments of posture will produce some change in the position of his eyes in space. Such changes will modify the retinal images in a quite specific way. Just as the image of the terrain in the right eye differs from that in the left eye by being horizontally skewed, so the image in either eye will be similarly skewed when the head is moved sideways for a distance of two and a half inches. Both effects are similar in principle to what astronomers know as parallax. The first is termed binocular parallax; the second is usually called motion parallax.

Every photographer is aware that even a slight movement of his camera during exposure will shift the image on the film, for it ruins his picture. The same kind of shifting of the image on the retina occurs all the time during vision, with the difference that vision is enriched rather than spoiled. The retinal image, of course, has a very different function than has the photographic image. It is not registered as an unchanging distribution of grains of metallic silver on a film, but as a flow of neural excitations and ultimately as visual experience which continues in time and changes as the image changes. Motion of the retinal image relative to the mosaic of sensitive rods and cones is therefore a normal stimulus-condition for vision and one which is almost continuously present during waking life. It must be remembered, however, that motion of the image produced by head movements is not the same as that produced by eye movements from one fixation to another.

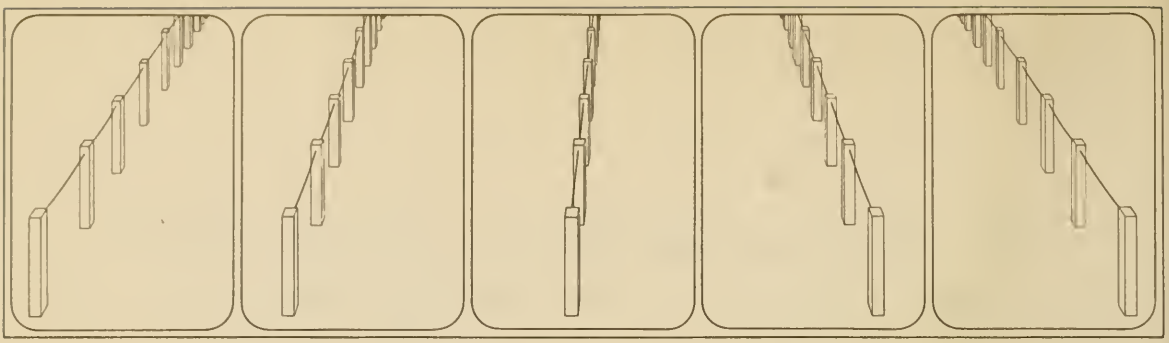


FIGURE 52. Successive Views of a Row of Fence Posts

The former is a stimulus for the perception of space, as we shall try to prove, and is also a precise sensory correlate of locomotor behavior; the latter has no such stimulus function. The former motion is one which deforms the image; the latter is one which simply transposes the image.

Figure 52 shows successive views of a row of fence posts. If you are moving at right angles to the line of posts the parallax (shift) of each post decreases as its distance increases. This decrease follows the same geometrical law as does the decrease in size. Consequently the image of the posts undergoes a horizontal skew which can be specified mathematically.

The Gradients of Deformation of the Image During Movement of the Observer

The visual field of an observer is alive with motion whenever his head moves. If the reader will fixate an object across the room and then move his head or change his posture, the shifting of contours can probably be noticed. Actually there is a deformation of the color patches in his field, and this corresponds to the deformation occurring in his retinal images. Like blur, double outlines, and the other characteristics of the visual field, this lively shifting of contours is visible only with special attention, and sometimes special practice, since what we ordinarily experience is the visual world, which does not manifest deformation. Only when we drive a car or ride on a train does it become so strong as to be unmistakable. A better idea can be obtained of what the retinal image must be like in this respect by holding a "view" camera with its shutter open in front of one's eyes and then moving it as the head would move. The inverted image on the ground glass screen presents a striking contrast to the stability of ordinary visual perception.

The fact of the relative motion of objects as a cue for their depth has been known for a long time and has always been included in the list of indicators by which the distance of an object may be judged. It has often been described as it appears from a moving train. Helmholtz wrote about it as it looks when walking:

> Suppose, for instance, that a person is standing in a thick woods, where it is impossible for him to distinguish, except vaguely and roughly, in the mass of foliage and branches all around him what belongs to one tree and what to another, or how far apart the separate trees are, etc. But the moment he begins to move forward, everything disentangles itself, and immediately he gets an apperception of the material contents of the woods and their relations to each other in space, just as if he were looking at a good stereoscopic view of it.

When we walk through open country with eyes fixed on the distance, Helmholtz wrote:

> ....objects that are at rest by the wayside.. appear to glide past us in our field of view in the opposite direction to that in which we are advancing. More distant objects do the same way, only more slowly, while very remote bodies like the stars maintain their permanent positions in the field of view.. Evidently, under these circumstances, the apparent angular velocities of objects in the field of view will be inversely proportional to their real distances away; and consequently safe conclusions can be drawn as to the real distance of the body.... (53, vol. 3, p. 295).

Helmholtz referred to the experience of depth as an "apperception" or as a "safe conclusion" about the objects in the field of view. He was thinking only of objects and not of a background or an array of elements. The relative motion he described is a variable of the retinal image itself when the latter is considered as a whole. When it is so considered, as a projection of the terrain or as the projection of an array of slanted surfaces, the retinal image is not a picture of objects but a complex of variations. If the relative motion is analysed out and isolated from the complex of other variations, it proves to be a lawful and regular phenomenon. Defined as a *gradient* of motion, it is potentially a stimulus correlate for an experience of continuous distance on a surface, as we shall see, and one no longer is required to postulate a process of unconscious inference about isolated objects.

We probably have a high degree of sensitivity to this kind of visual stimulation, that is, motion consisting of a change in shape as distinguished from motion consisting only of a change in location. The conclusion is indicated by the fact that, in a motion picture view, an unnoticeable optical distortion in the photographic image of a room becomes obvious at once when the camera is "panned" from one side to the other, being seen as a deformation of the walls and corners of the room. The distortion of the still picture of the room from normal perspective is not great enough to be noticed, but the slight change in the shape of the image produced by moving the camera lens is noticed easily, and it appears as an apparent stretching and contracting of the walls.

The Perspective of Visual Motion for a Projected Terrain. Consider an observer who is moving parallel to the ground, as he would be during normal locomotion. Let us assume that his eyes are fixated on the horizon, that is, let us disregard for the present any movement of his eyes either in pursuit of the gliding terrain or from one fixation to another. The way things appear to an observer riding in an open car on a clear moonlight night is a good example for anyone who can isolate this experience from distracting memories. When such a scene is looked at just for the sake of looking, the horizon, the stars, and all the field of view upward are motionless, but the world and the ground below flow past in a continuous stream. The flow vanishes at the horizon; but it increases downward and reaches its maximum on the road beneath. The flow is not like that of a river as it would appear from a bridge but rather as it would appear in perspective from the bank; it is a continuous deformation of the surface, not a movement in the ordinary sense of that term. In whatever direction one looks,

forward, to the side, or behind, the flow decreases upward in the visual field and vanishes at the horizon. There exists, in short, a *perspective* of this motion which is fundamentally similar to the perspective of the density of texture and the size of objects.

Considering the terrain as projected on a plane in front of the eye, the rate at which any element or object flows is inversely proportional to its physical distance from the observer, as Helmholtz noted. The motility of the world, like its density, decreases as it recedes. The geometry of this decrease is precisely like that illustrated in Figure 30 for ordinary perspective. Near elements of the ground change their angular direction from the observer (parallax) more rapidly than do distant elements. There is, in other words, a continuous gradient of the velocity of the ground "going by" from a maximum at the bottom of the visual field to zero at the horizon.

The Direction of Flow in Visual Motion Perspective. Motion, of course, always has *direction* as well as *speed.* The two may vary quite independently of one another, and it should therefore not be surprising to discover that the direction of flow in the projection of the ground during our moonlight ride does not vary in the same way the velocity does. We are referring to the visual field, of course, in which direction can vary through 360° like the hand of a clock. This visual direction of flow depends upon the physical direction of the spot in question, which varies like the pointer of a compass. The physical direction from the observer of any spot on the ground and the physical direction of the line of locomotion are objective directions which cannot be literally copied in a projected image.

To be specific, the flow of the terrain is visually downward in the field as one looks ahead from the driver's position; it is to the right as one looks to the right, or to the left as one looks to the left, and it is upward in the field as one looks behind. In other words, it is different in different visual fields of fixation as these vary from the forward direction of locomotion. As the observer gets successively different visual fields in turning his head around to the right the direction of flow changes successively in a counterclockwise rotation, and if he turns to the left the change is precisely the same except clockwise. Descriptively, the visual field ahead expands outward from a focus, the visual field behind contracts inward to a focus, and the visual field to one side or the other is being continuously skewed. The foci of expansion and contraction correspond objectively to the points toward which and away from which locomotion is aimed. Most of us have observed the expanding visual field ahead while driving a car, and the contracting visual field while riding on the rear end of a train.

These visual fields are represented in Figures 53 and 54 as a projected terrain viewed from an airplane during level flight. In Figure 53 the focus of expansion, or point of aim, is on the horizon, and the diagram would be the same in principle for a man on foot or for the driver of an automobile. The surface as projected is continuously deformed in the manner indicated, and each arrow is a vector representing the velocity and direction of flow of the

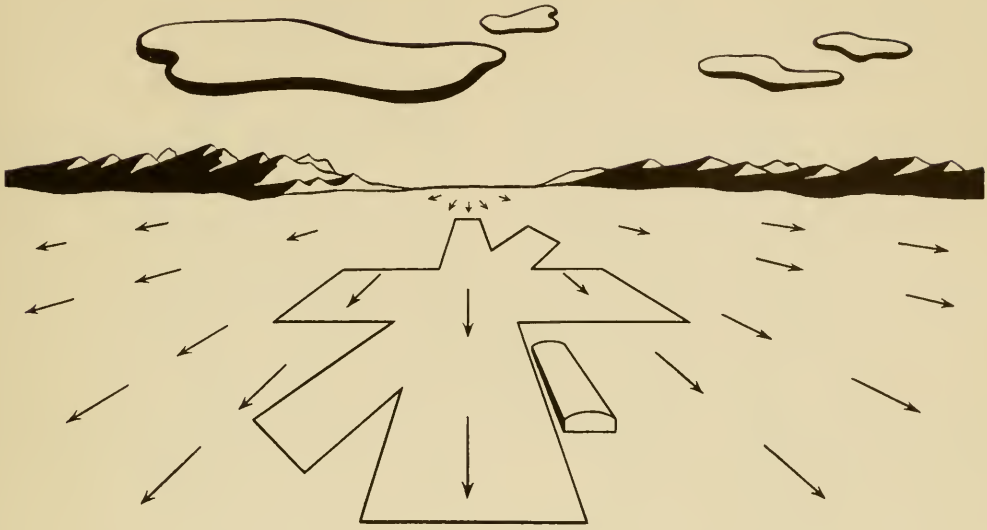


FIGURE 53. Motion Perspective in the Visual Field Ahead

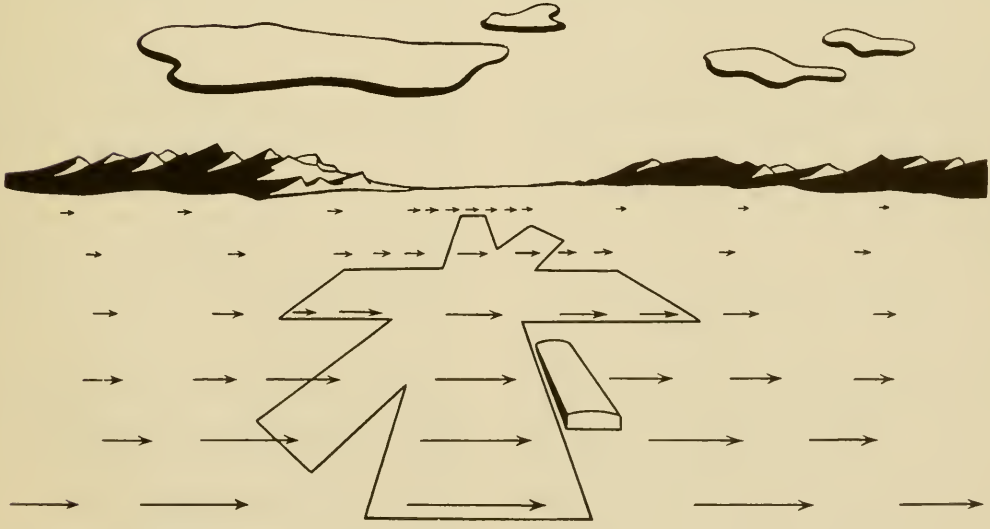


FIGURE 54. Motion Perspective in the Visual Field Looking to the Right

If the arrows are reversed, this becomes the visual field looking to the left.

surface-element at that point. On the picture-plane there is a gradient of decreasing velocity from the bottom up to the horizon, and also a gradient of changing direction from the midline to either side (which itself decreases upward). The mountains on the horizon and the clouds are at such great distances that the velocity of deformation approaches zero, and the sky is not deformed since it is not a determinate surface. If the direction of the arrows were reversed to make the field contract radially inward instead of expanding radially outward, it would then be the field of view looking backward.

Figure 54 represents a deformation familiar to passengers on trains and airplanes. On the picture the velocity gradient is similar to the gradient of linear perspective and, like that gradient, it vanishes on the horizon, so long as the observer is fixating the horizon. Actually, Figure 54 should be visualized as merging with Figure 53 on the left side and with the reversal of Figure 53 on the right to yield a combination of visual fields which cannot be projected on a plane picture. If the arrows on Figure 54 were reversed, the scene would represent the visual field looking 90° to the left of the line of locomotion.

It is interesting to note that if we could combine all these two-dimensional projections of a three dimensional visual world into a single scene, we would obtain a two dimensional space, in the geometrical sense, which is non-Euclidean.

It would have the properties of the theoretical space defined by the surface of a sphere considered as a two-dimensional surface, i.e. it would be boundless and yet finite, and it would return upon itself. The space composed of one's combined visual fields may be said to be a *curved* space in the sense that a point which traces a straight line will eventually come back to the position from which it started instead of travelling off endlessly in the same direction. In other words, if human beings had a visual field whose width included the entire horizon – if they could see all the way around at the same time like a rabbit – the field during locomotion would appear to open up ahead and close in behind in a rather astonishing manner. Such characteristics of the visual field created a great deal of difficulty for the early students of perspective and for painters who wished to represent a large sector of the visual world on a picture-plane.

Actually, of course, no rabbits and relatively few men have ever adopted the peculiar attitude of psychologists, artists, and geometers which enables them to see their visual field.[1] They are, with good reason, perfectly content with the visual world as it is normally perceived, conforming to the rules of Euclidean geometry. The world does not undergo any flowing deformation; it is seen to be stable and rigid and what moves is the observer himself. The ground does not move in a gradient of directions which change around the clock; the observer sees himself moving in a single direction in three dimensional space. Objects do not change position in relation to the observer; the observer sees

[1] The writer must admit that he has had no introspections from rabbits to justify this conclusion.

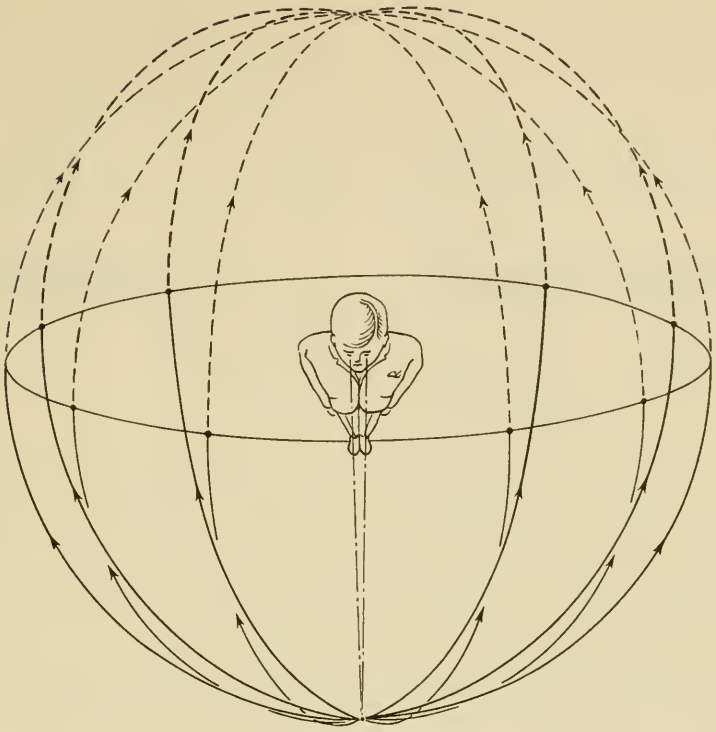


FIGURE 55. The Directions of Deformations in the Visual Field during Forward Locomotion, as Projected on a Spherical Surface around the Head

his own change of position relative to the objects. This flowing deformation then in the visual fields we have been describing is nothing but a visual symptom of the gradients of velocity and direction in the retinal image. We shall try to show how these gradients are stimulus correlates for perceived space and perceived locomotion.

The gradients of velocity and direction are invariable accompaniments of locomotion if the observer keeps his eyes open. The focus of expansion in the field ahead is an exact indicator of the point in the world toward which he is going; a shift of the focus goes with a change in the direction of locomotion and this provides him with a sense of a point of aim. The point of aim is, in fact, implicit everywhere in the visual field, and even when the observer does not look where he is going, he can in a sense see where he is going. This fact enables strollers, automobile drivers, and aircraft pilots to see a great deal more of the scenery than that which lies directly ahead of them, although admittedly it is a performance which should be practiced in moderation. So strict are the geometrical relationships between physical motion of the observer's body and retinal motion of the projected environment that the latter provides in fact the chief sensory

guide for locomotion in space. Retinal deformation is actually a kind of visual kinesthesis.[2]

The flow of the terrain if one imagined oneself to be looking vertically downward at it from a considerable altitude during level flight would be in general a flow from the top to the bottom of the visual field. The velocity would be a maximum at the point physically below the airplane, at the center of the visual field, and would decrease outward in all directions toward the horizon, 90° from the center of the field. The direction of flow would be divergent at the top of the field, away from the focus of expansion, and convergent at the bottom of the field, toward the focus of contraction. If the directions in the field were plotted in polar coordinates, the plot would look like a melon-shaped family of curves. It should be noted, however, that this downward looking visual field would have to be accompanied by eye movements, the effect of which we have not yet considered.

If one asks how the retina is stimulated when the observer looks vertically *upward* into a clear sky during locomotion, the answer is probably obvious. There is no ordinal stimulation of the retina, but only luminous stimulation. There is no texture, no surface, no motion, and accordingly no determinate world and no visual sense of one's own movement. Retinal motion presupposes the stimulus of retinal texture, as we pointed out at the beginning of this chapter. The air traveller who looks out into a cloudless sky gets the impression of floating in empty space even though he may be objectively moving at 200 miles per hour.

If there are cloud masses in the sky, however, these provide textured surfaces on which to anchor space. Some of the most interesting spatial experiences in flying are provided by the motion perspectives of broken cumulus clouds when passing through them. If, as frequently happens, the clouds take the form of a solid horizontal overcast or ceiling, the flow of this surface will be precisely the inverse of the flow on the ground, as represented in Figure 56. As in previous instances, an inverted gradient yields an inverted surface.

The Effect of Eye-Movements on the Retinal Gradient of Motion. The flowing visual field has so far been described with the limiting assumption that the observer's eyes are fixed on the horizon and are therefore for all practical purposes motionless in his head. During actual locomotion, however, the eyes are seldom on the horizon and, since all other parts of the world are flowing, they are seldom stationary. The eyes of an automobile driver, for instance, perform an endless series of downward drifts and upward jerks as he fixates points on the flowing road ahead of him, with occasional jerks

[2]The difference between this description of motion perspective and the established conception of motion parallax is that the former is allied to locomotion of the organism whereas the latter usually implies that the animal in question must move its head from side to side in order to obtain the cue for depth. Moreover, the retinal gradient of velocity on a ground surface is a stimulus-variable whereas relative displacement is conceived only as a cue. The general theory of retinal deformation as a type of visual kinesthesis covers all special cases of head-movement.

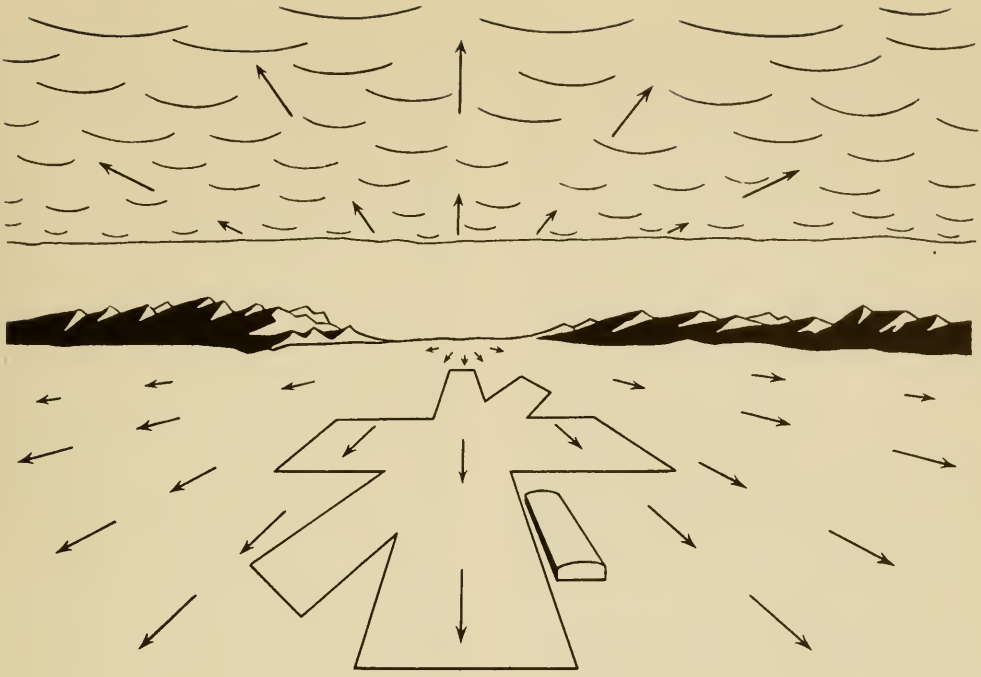


FIGURE 56. Motion Perspective with an Overcast or Ceiling

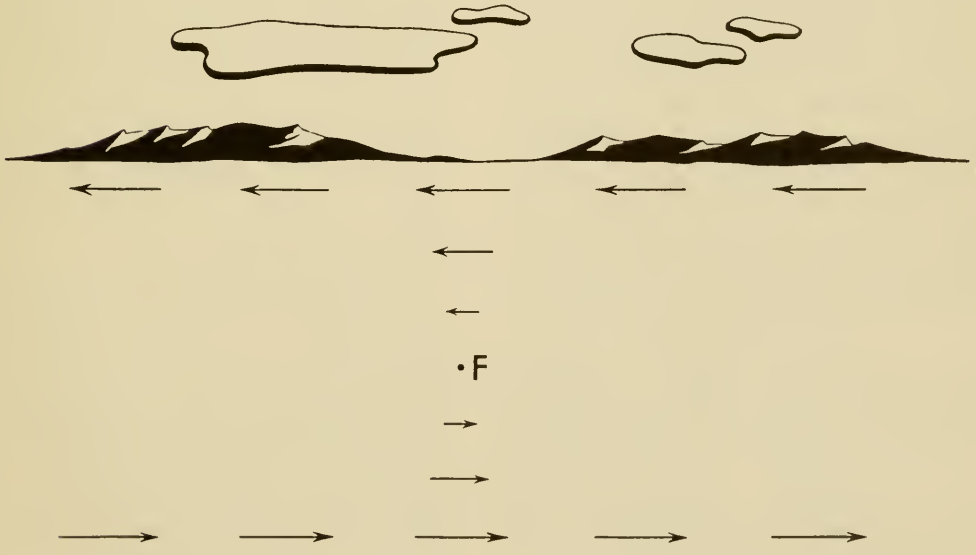


FIGURE 57. The Gradient of Flow Looking to the Right when the Observer Fixates a Spot on the Terrain

in other directions to pick up objects of interest at the side of the road. The eyes of a traveller seated on the right-hand side of a moving train make an endless series of drifts to the right and jerks to the left. The drifts are known as *pursuit movements*, and their function is to maintain the image of a selected moving spot or object at the center of the retina. The jerks are the type known as *saccadic movements*, the general function of which is to establish a new fixation, and they occur in the act of scanning the environment, in reading, and between the fixated pursuit movements just described. The question we now need to ask is this: What is the effect of the pursuit movements on the gradient of retinal flow which we have asserted to be a geometrical correlate of distance in the environment?

The problem is similar to the one encountered in connection with crossed and uncrossed retinal disparity as a correlate of distance, and the solution is of the same sort. Figure 57 represents the diagram of Figure 54 as it would be modified when the observer fixates a spot not on the horizon but on the ground halfway down. This is the scene frequently observed from a train window. The point of fixation is indicated by F. At this point the velocity of projected flow is zero, since the pursuit movement of the eye compensates for its change in direction and keeps its image stationary at the center of the retina with a considerable degree of precision. The flow is also zero on a line to the right and left of point F; all points above this line flow to the left and all points below it flow to the right. Although this deformation of the image might be supposed to be quite different from that when the point of fixation is on the horizon, the perceptions which result in these two cases are equivalent. The reader may check this observation for himself, with a piece of ruled paper on a table top substituting for the terrain. It is as if the eye's movement added a constant velocity toward the left, in the sense of vector addition, to the gradient of velocities to the right shown in the former diagram. The horizon and the clouds now move across the retina at the rate which the point of fixation, F, formerly possessed, *but in the opposite direction.* The gradient from the horizon downward of decreasing flow to the left is equivalent to the former gradient of increasing flow to the right. Positive and negative velocities may be added algebraically, and the *gradient* of motion remains constant on the retina whether the zero point of the gradient corresponds to the horizon, to a point halfway down the terrain, or even to a very near object on the ground. We know from physics that motion is relative to an arbitrary zero point, or frame of reference. The stimulus-motion of which we are speaking is a physical, not a phenomenal, motion.

The phenomenal motion of objects is not ordinarily perceived as relative to an arbitrary frame of reference. The stable visual world provides an absolute zero and hence there is an absolute sense of motion or rest with reference to the ground – both for oneself and objects. This fact is what makes the theory of the relativity of motion in physics difficult for the non-physicist to comprehend. Only in exceptional circumstances (such as looking

out of a train window at an adjacent train which fills the entire visual field) does it ever become equivocal whether the observer is moving and an adjacent object is at rest, or whether the opposite is true. In such circumstances there is no terrain surface on which to anchor the visual world. The same kind of equivocal motion can occur for an observer in a completely dark room who is presented with a single slowly moving point of light (28, summarized in 32). In fact, with no visual stimulation except a fixated light-point both motion and position of the point become indeterminate, and a physically stationary point may appear to make random excursions in any direction. This is the well-known autokinetic phenomenon. Its explanation, as Koffka understood (67), is that in the absence of ordered visual stimulation the point may as well seem to move as not and therefore sometimes does!

The effect of pursuit movements of the eyes on the deformation of the retinal image during locomotion, we conclude, is to add a constant to the motion of each point in the image but not to modify the gradients which are its essential characteristics. The variable which corresponds to physical distance in the environment must therefore be a gradient of motion-in-a-certain-direction, not a simple gradient of velocity as such. The horizon in the retinal image is a line which is determined by vanishing values of the stimuli of size and texture, and it has no intrinsic relationship to the gradient of motion. The latter gradient may have a kind of horizon of its own at the *anatomical* level of the fovea, that is, at the midline of the retina where motion vanishes during a pursuit movement of the eye. The two kinds of horizon coincide only when the eyes are fixed on the physical horizon – in the so-called primary position of optics. The horizon of motion is not a limiting value, like an asymptote, but a zero value on a scale of opposites. Whereas size in the image varies on an intensitive scale with an intrinsic zero, motion in the image varies on an *oppositive* scale with an arbitrary zero (38 p. 223).

Retinal Deformation with a Surface not Parallel to the Line of Locomotion. The foregoing descriptions and diagrams of the deformation of the image apply only to the case in which the movement of the head is in a line parallel to the material surface projected. Although this applies to normal locomotion with reference to the ground, it does not apply to all locomotion, nor to movement of the head with reference to the slanting surfaces of objects. A particularly important practical application is to the locomotion of a flier who is approaching the ground at a certain angle of glide preparatory to landing his plane. Distance perception, we are reminded again, is no mere visual luxury to be enjoyed in parlor stereoscopes but a biological necessity, one use of which is to enable us to get about without colliding with obstacles (41).

When an observer approaches a surface instead of moving parallel to it, a modification of its deformation is introduced in that the focus of expansion is no longer on the horizon of that surface but at a particular spot on it – the point of collision with the surface. The rule is that all deformation in a forward visual field radiates from this point. Crudely speak-

ing, the environmental scene expands as we move into it, and the focus of expansion provides us with a point of aim for our locomotion. An object in our line of travel, regarded as a patch of color, enlarges as we approach. It is not difficult to understand, therefore, why this expansion should be a stimulus for sensed locomotion as well as a stimulus for sensing the lay of the land. The behavior involved in steering an automobile, for instance, has usually been misunderstood. It is less a matter of *aligning the car with the road* than it is a matter of *keeping the focus of expansion in the direction one must go.*

When the focus of expansion is a spot on a vertical wall toward which a man walks, the flow is zero at that spot and increases symmetrically around it, if we disregard eye movements. When he approaches it at a slant, the flow is correspondingly asymmetrical, the velocity becoming greater on the near side. When a pilot comes into an airfield or on to an aircraft carrier, the direction and rate of flow of the ground are determined by the focus of expansion and the horizon in combination. The velocity increases outward from the focus but then decreases and approaches zero at the horizon. Figure 58 illustrates the field of a flier who intends to land on a runway. The gradients of motion are approximately represented by a set of vectors indicating direction and rate at various points. All velocities vanish at the horizon. The focus is a projection of the point on the ground at which the glide is momentarily aimed; if the glide is steepened the focus will move downward in the field and if it is made shallower the focus will move upward toward the horizon. It is therefore an indicator for the pilot as to where his wheels will touch the runway and hence it is a cue as to whether he is undershooting or overshooting the field. Inasmuch as either alternative is frequently fatal, the pilot has a vital interest in such cues. The working out of their interrela-

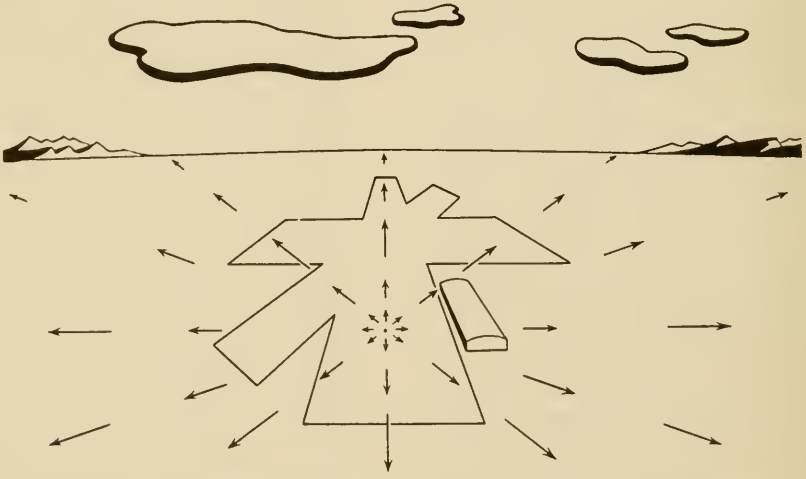


FIGURE 58. Gradients of Deformation during a Landing Glide

tions was a part of the research on aviation psychology during World War II (39, Ch. 9).[3]

Complex discriminations of the direction, altitude, and angle of flight, whether consciously perceived or unconsciously incorporated in the reactions of the pilot, are undoubtedly *learned*. But it is the *discriminations* that are learned, not the stimuli on which they are based. The retinal image of the novice flier is exactly like that of the experienced pilot; the difference is that the latter reacts differentially to variations of the image to which the former does not. The experienced pilot does not see *more* of the visual world than the novice but he sees a more *differentiated* visual world. The effect of his training is to enable him to make fine rather than gross discriminations of distance, altitude, angle of glide, angle of drift, speed of flight, the position and direction of everything, and therefore to see accurately a continuous visual world in which he himself moves with precision. In this sense, and only in this sense, is space perception a product of learning.

To a certain extent the pilot can be supposed to learn habits of *scanning* or *inspecting* the environment, habits having to do with eye-movements, and these no doubt improve the stability and comprehensiveness of the visual world and his orientation in it. Of this there will be more to say in the next chapter. Learning to *attend* to novel features of the world, to explore it, is something which psychologists do not understand at present. What the pilot cannot be supposed to learn is the impression of depth, distance, and altitude considered as an inference derived from unlearned sensations.

Evidence that Retinal Deformation is a Stimulus Variable for Space and Locomotion. Although relative motion in the visual field has always been accepted as a factor in distance perception, the theory of the retinal motion gradient as a stimulus for distance perception requires proof. The theory makes possible experiments which will either verify it or not. So far, only the theory has been constructed – actually no more than the outline of a theory – and the experiments are lacking. What is required is a proof that an artificially produced gradient of point-motions on the retina, in isolation from other gradients, will yield an observer the impression of continuous distance on a surface. Variations in the gradient must then be shown to produce variations in the slant of the surface.

This demonstration has never been set up and performed as an experiment, inasmuch as the gradient of velocity has not been isolated from gradients of ordinary perspective. An interesting method of attack on it would be with frame-by-frame photography of spots, lines and other texture-stimuli, that is, by animation of pictures. Such evidence as does exist comes from a program of wartime research which

[3]Since fliers in general are little inclined toward introspection, the analysing and describing of such appearances in the visual field is almost absent in the literature of aviation. A notable exception is Langewiesche's book, *Stick and Rudder*, which describes and illustrates the visual cues for landing in a way consistent with the theory here presented, and is full of acute observations on the space perceptions of the flier (70). Pfaffmann is another exception in that he has analysed the traditional binocular cues experimentally (87).

the writer directed in 1943-1946 (39) on the space perception of fliers and on the use of motion pictures to represent space. One goal of this research was to find out how effectively the perception of a three dimensional world can be aroused by the flat motion picture screen and by the two dimensions of a still photograph. It became evident that both still and moving pictures can yield a more adequate visual world than is generally recognized or understood. The theories of texture-perspective and motion-perspective are derived from this evidence.

If we take a motion picture shot, for instance, of the scene ahead of an airplane during an approach glide (Figure 58) and project it on a screen, the projection can be studied in two ways. The experimenter can plot by rough measurement the actual gradients of expansion on the screen. When this is done, the results are approximately those represented in Figure 58. Also the experimenter can present the screen picture to a group of trained fliers or of untrained persons and have them describe what they see, or identify the focus of expansion, or make other judgments. When this is carried out, the notable result is that no observer, trained or untrained, perceives an expanding screen picture and nothing more. Each onlooker gets a compelling experience of moving through space in a specific direction toward the ground. The angle of this movement and the point of its aim can be

FIGURE 59. A Landing Field during an Approach, at 1000 Feet, at 500 Feet, and at 250 Feet from the Edge of the Runway

judged by all – with varying accuracy, it is true, but nevertheless judged. The experiment was often repeated and the experience proved to be sufficiently reliable to justify making a series of such motion picture shots into an experimental test for measuring the aptitude of candidates for flying training, and even for measuring one aspect of the proficiency of trained pilots (39, Ch. 9).

The examination of many other motion-picture shots taken from moving aircraft, including views at right angles to the line of flight, yielded the conviction that the gradients of motion we have described are genuinely compelling for the experience of distance. Many of the phenomena of animated cartoons point to the same conclusion. When children observe Donald Duck diminishing rapidly in size on the screen, moving upward slightly, and retaining the same shape, it is likely that they actually see him whizzing off into the distance. The experimental *isolation* of these phenomena, however, remains to be carried out.

The Types of Retinal Motion

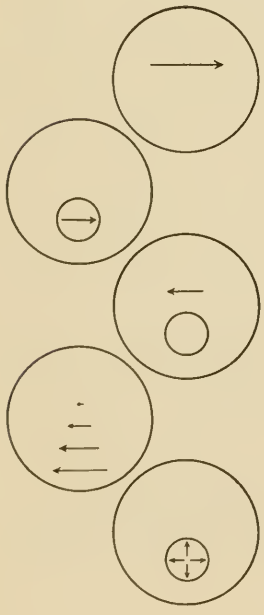
Ordinarily when we speak of perceiving motion we refer to the *movement of an object.* The problems connected with the motion of an object's image across the retina have been studied for many years. What we have been talking about, however, is the *movement of the observer* and the accompanying deformation of the environment's image on the retina. The latter has never been fully described and has been little investigated. Our concern has been with the perception of locomotion rather than the perception of ordinary motion, so far as locomotion is correlated with visual stimuli. It would clarify the matter if we could classify the different kinds of stimulus-motion that may occur in the retinal image and specify the physical situations which produce them.

We have already seen that the retinal image is a two dimensional projection of focused light on a sensitive anatomical surface. When we say that it undergoes motion, we must always mean motion with reference to that surface. One possibility is for the projection as a whole to move with reference to the surface. Another possibility is for a delimited part of the image to move with reference to the total stationary image (usually the surrounding image), or for the surrounding image to move with reference to a stationary delimited part. All three of these kinds of motion may be defined geometrically as rigid motions. A *rigid* motion is one of translation or one of rotation or a combination of these, but not one which involves any change of shape or distribution. A square which moves across the retinal surface, whether it moves with all its points tracing straight lines or tracing arcs of circles, remains a square. A total image which moves rigidly across the retina keeps the same distribution although it loses old and gains new parts at opposite margins of the retina. This rigid kind of motion relative to the retina has an analogy with the motion of physical objects – solid rather than liquid substances – and it is the kind analysed in classical physics and dealt with by the laws of motion. But it is not the only or even the most frequent variety of motion-

in-general which the retinal image may undergo. The image or any part of it may be expanded, contracted, or skewed in a number of ways. All of these ways are definable in terms of gradients of velocity and direction along one or another axis of the retina. They are also definable in terms of mathematical transformations of one image into another. This is the kind of motion which has been termed deformation. As with rigid motion, deformation may apply to the retinal image as a whole or to a delimited part of it.

These types of generalized retinal motion can be put into a very simple correspondence with the physical situations which produce them. Motion of the retinal image as a whole occurs with saccadic eye movements from one fixation to another, the head and body being stationary. Motion of a delimited part of the image occurs with fixated eyes when an object in the physical environment moves, with reference to the ground, at right angles to the line of sight.[4] Motion of the surrounding image occurs with the eyes in a pursuit movement when an object in the physical environment moves at right angles to the line of sight. (The perceptions resulting from these latter two modes of stimulation are surprisingly equivalent). Deformation of the total image occurs with eyes fixed on the horizon when the head of the observer moves in relation to the ground (locomotion). Deformation of a delimited part of the image occurs with stationary head and eyes when an object in the physical environment moves in depth – in any direction not at right angles to the line of sight. The relationships can be summarized in the form of a table.

[4] Actually, the retinal motion is non-deforming only when the object crosses the line of sight.



Mode of Retinal Motion	*Physical Situation*	*Perception of Objective Movement*
Rigid motion of the total image	Saccadic eye-movement in stationary environment	None, Perception of a stable world
Rigid motion of a delimited image	Stationary eyes, object moving frontally	Object moving frontally in stable world
Rigid motion of the image except for a delimited part	Pursuit movement of eyes with object moving frontally	Object moving frontally in stable world
Deformation of the total image	Movement of head in stationary environment	Movement of self in stable world
Deformation of a delimited image	Stationary eyes with object moving in depth	Object moving in depth in stable world

These modes of stimulation are to be understood as individual cases and not as alternatives. They may coexist in many combinations, and the resulting perceptions will also apparently coexist without interference or confusion. A man driving an automobile, for instance, can see his own movement and at the same time perceive the movement of a car approaching on a crossroad without obvious loss in the clarity of either perception. The deformation of the total image, considered as a set of gradients, appears to lose none of its stimulus function even though a number of other motions be summated with it.

Motion as an Independent Variable of Experience. The theory that all visual experience was made up of elementary sensations of flat color, each spot having an elementary quality of location, implied that motion was a perception which could always be analysed into a change of location of a color in time. Visual movement was not a simple but a complex experience, the sense of location being the primary fact on which motion depended. It seemed to be only reasonable that one could not see a thing as moving if one did not see it in successively different places.

The localization theory of space-perception, it has already been argued, was inconsistent with a good many facts of experience but nowhere was it more false than in this corollary. For, strange as it may seem to someone who has not observed it for himself, an impression of motion may be got without any impression of change in location. Although under ordinary circumstances the two go together, they need not do so. The clearest demonstration of this fact is the negative after-image of motion. One gets it after fixing one's gaze at a waterfall for ten or twenty seconds and then staring at another part of the scenery. It is sometimes observed out of a train window after the train has stopped. A better method is to watch the apparent expansion of a rotating disk on which a spiral line has been drawn so as to make the line appear to move outward from the center. The after-image of motion is confined to the stimulated area of the visual field (like an after-image of color) and it consists of a vivid impression of motion in the opposite direction, persisting for some seconds and only gradually dying out. The fact to be noted is that an unmistakable motion is perceived but that it does not involve any change of position. When we look away from the waterfall the foliage appears to move upward but it is not displaced. Out of the train window the ground seems to move forward but not to shift position relative to the window. After the rotating disk has been stopped it seems to contract, but it does not get smaller (although it may yield a queer impression of receding or retreating from the observer). One cannot avoid the conclusion that here is an experience of motion without an accompanying experience of displacement in the frontal plane.

The same type of experience may be obtained by having an observer fixate the center of a semicircular screen (a perimeter) and presenting motion in an area or window at the far periphery of his visual field. He may sense the motion correctly, as compared with no motion, but without

any impression of what the moving pattern might be, and even without any impression of what the direction of the movement is. An even more notable example of this kind of experience, because of its historical importance, is what has been called the phi-phenomenon (118). Successive exposures of two separated black spots will yield stroboscopic movement, one spot jumping into the position of the other. When the timing of the exposures is not optimal, however, the impression reported is one of movement without any moving spot, or "pure phi." Taking all these instances together, it seems certain that there must exist a visual quality of what might be called movement-as-such.

A combination of successive and adjacent order over the retinal mosaic would seem to be the fundamental stimulus condition for this impression of motion. The stimulus may be complex, in a mathematical sense, but the impression is simple. If motion is a variable of experience which can exist in its own right independent of the displacement of an object, there is probably also a variable of stimulation to which it will prove to correspond. The exact definition of this stimulus variable remains to be established. One fact about it, however, can be stated with some certainty: the stimulus for motion is not necessarily motion in the retinal image. To put it more precisely, it is not necessary to have the stimulation of a continuous series of adjacent points at a continuous series of successive instants. (Points and instants are fictions, in any event, as the paradoxes of Zeno ultimately demonstrated.) Instead, the stimulus condition for visual motion may be adjacent spots at successive moments with a considerable separation between the spots and the moments. There must be a combination of adjacent and successive order but neither the adjacent series nor the successive series has to be "dense," or "compact" as the number-theorists say. In short, the essential stimulus for motion is not physical motion.

The rule to apply is that the stimulus for a spatial impression need be only a correlate, not a copy, of the corresponding physical variable. Physical motion, like physical shape and physical depth, does not have to be duplicated in the retinal image in order to yield phenomenal motion. The paradox of the stroboscope and the motion picture – that we can see motion in a situation where nothing moves – is no longer a paradox if this rule is applied. One has only to relinquish the assumption that a stimulus must be a replica.

The Perception of Acceleration as Distinguished from the Perception of Motion. Normal Locomotion. It should be noted that the relationships of the types of retinal motion in the table refer to the perception of motion as such, which must be distinguished from the perception of acceleration or force. Perceived motion of one's own body seems to be mediated best by vision, whereas perceived acceleration and perceived force, which are actually the same thing, are mediated by stimulation from within the muscles (the muscle sense) and stimulation from within the inner ear (the labyrinthine sense). The retina is insensitive to forces acting on or within the body, and also insensitive to

an acceleration of the body, that is, a change of motion with time. On the other hand the inner ear is extremely sensitive to any force acting on the body, gravity for example, and to any acceleration of the body, but is wholly insensitive to uniform motion. The air traveller who floats in the sky at a uniform speed of two hundred miles an hour is witness to this latter fact. Stimulation of the inner ear and of the receptors in muscles do not provide a kinesthetic sense in the literal meaning of the term, but a force sense. Visual stimulation, with the possible addition of stimulation of the joints, is a better mediator of the *motion* sense.[5]

When a man walks, or, as we say, gets about under his own power, he experiences both motion and force. The perception of voluntary locomotion, as distinguished from passive locomotion, is jointly determined by two sources of stimulation, stimuli from the retina, on the one hand, and from the muscles plus the inner ear on the other. In the study of active locomotor behavior, an extremely important branch of applied psychology, neither of these sources of stimuli can be neglected. What has come to be called psychomotor behavior, for example operating a crane or a lathe, is only a special form of it. Walking, running, athletic achievements, automobile driving, flying, and many kinds of tool-using would all be included in the category of active locomotor behavior, for what we mean by it is behavior accompanied by the perception of moving from one place to another. As experienced, it includes sensory impressions of exertion or muscular action and of visual motion, the two being covariant. The product of these stimuli is something neither wholly motor nor wholly visual; it is locomotor action in a visual world. As we shall try to show in the next chapter, the perception of a stable, upright visual world also depends on covariation of the visual sense with the so-called body senses, and subsequently we may have a better understanding of why locomotion and the stability of space are intimately connected.

Not all locomotion, however, is active or voluntary. Driving a car or flying an airplane is a relatively passive action compared to walking, and for the passenger it is still more passive. Modern man has gone to great lengths to save himself effort during locomotion, and special problems arise in learning to use these locomotor machines. The muscular actions involved in steering, accelerating, or balancing an airplane with the stick are artificial and highly reduced actions as compared with walking or balancing one's own body, and the visual stimulation becomes proportionally much more important than the bodily stimulation in these relatively passive types of locomotion.

Visual Motion during Rotation. The distinction between motion and acceleration also makes possible a better understanding of a kind of movement of the observer which has not so far been men-

[5] This conception, which admittedly does violence to the accepted meaning of the term "kinesthetic" will be extended in Chapter 13. The above distinction applies only to bodily acceleration and motion, not to the motions and forces perceptible among visual objects. The latter have been studied by A. Michotte in *La Perception de la Causalite*.

tioned — rotation of his body and head. This is seldom, properly speaking, a form of locomotion. Prolonged passive rotation, as distinct from active exploratory rotation of the head, is an artificial situation not met with outside a laboratory. Motion of this sort has nevertheless received a great deal of experimental study, usually employing a rotating chair in which the observer is seated. The eye movements which go with it are called compensatory and they are collectively known as nystagmus. When the head is rotated the eyes move oppositely in such amount as to preserve an unchanged visual field, thus compensating for the rotation. As rotation continues, the eyes jump with a saccadic movement to a new fixation and then follow the environment once more, repeating these slow and fast movements as long as the rotation continues. Under the circumstances the observer should perceive a stable visual world with a sense of his head moving in it. Actually this is the result, but only if the rotation is not prolonged or rapid. When an observer is whirled in the usual rotating-chair experiment, the compensatory movements soon cease to yield precise fixations and the stability of the visual world breaks down. It then appears to rotate as a world, and the observer becomes disoriented and dizzy.

The visual appearances of the world during the rotating-chair experiment are complicated by the fact that the subject is stimulated by accelerations which affect the inner ear as well as by retinal motions. In this artificial situation the two kinds of stimulation necessarily come into conflict. The compensatory eye movements, for example, are aroused not only by the retinal stimulation of movement but by the inner-ear stimulus of acceleration. Either stimulus will produce eye movements in isolation from the other, as many experiments have shown. During any initial rotation of the head, the acceleration and the retinal motion are concordant stimuli producing the same response. But during the stopping of this rotation, after an interval of uniform rotation, the acceleration is reversed and becomes negative, although the retinal motion continues in the same direction during this period. The compensatory eye movements may then cease to have a compensatory function; they may be reversed in response to the negative acceleration, and this after-nystagmus may even persist after the head is physically at rest. If the eyes are open this kind of eye movement inevitably produces an illusory motion of the environment. The world then appears to rotate in the direction opposite the bodily rotation since the compensatory eye movements are compensating, as it were, for a nonexistent motion of the head. Other phenomena, however, may complicate this result.

The illusions which occur when a subject is rotated in partial darkness or in an environment showing only a few points of light should be particularly strong, and they are of considerable practical importance since this is the situation that is encountered in night flying. Graybiel and his associates have studied this situation and described what they call the oculogyral illusion which, like the after-nystagmus, is produced by compensatory eye-movements not having any compensatory

function (45). Together with other illusions due to acceleration in night flying maneuvers, it may help to explain the occasional disorientation to the objective world which a pilot suffers – sometimes with fatal results.

Summary. The Sensory Analysis of Distance and Depth

The traditional list of clues by which the mind is believed to infer a world of three dimensions usually includes the following factors: linear perspective and the decreasing size of similar objects with distance, the apparent size of objects whose real size is known, superposition of one contour on another or the covering of a far object by a near one, the distribution of light and shade over an object, the relative motion of objects or monocular parallax of motion, aerial perspective and the loss of detail with distance, binocular disparity, or stereoscopic vision, degrees of ocular convergence and of accommodation of the lens, and occasionally the factor of angular location of an object between the bottom of the visual field and the skyline or, in other words, the amount of ground between the observer and the object (19, p. 270).[6] Sometimes these factors are called signs or criteria of distance or, more frequently, cues. Whatever the term used, it is clearly implied that they are not precise geometrical correlates of physical distance but probable indicators, symptomatic rather than exact.

In the attempts to describe these factors it has often not been clear whether they referred to sensory experiences, or characteristics of the retinal image, or facts about the physical object. These are three quite separate classes of facts which should not be confused with one another. The covering of a far object by a near one, for instance, is a description which mixes physics and experience. It cannot explain depth perception, as phrased, since it presupposes the phenomenon which it seeks to explain – one object behind another. The joint contour in the retinal image is two-dimensional. How do we see depth at a contour so that one side of it appears near and the other far? That is the fundamental question.

[6] Brightness is sometimes listed as a cue to distance, the presumable assumption being that an object necessarily appears darker as its distance from the eye increases (19). Apart from the little-known effects of atmospheric conditions on visibility – aerial perspective – the assumption has no basis in physical fact. It is true that a point-source of light yields an intensity at the eye which decreases in proportion to the square of the distance. But an illuminated surface (an infinite number of theoretical points of light) yields the same intensity per unit solid angle at the eye when it is far as when it is near, and hence possesses the same brightness, within limits, under both conditions. Each theoretical point becomes theoretically dimmer, but the density of points becomes theoretically greater in exact proportion. In the ordinary environment of illuminated surfaces, therefore, brightness is not an indicator of distance.

The fact is, however, that in a dark room with no other sources of stimulation the more highly illuminated of two equidistant and otherwise equivalent surfaces tends to look the nearer. This is an empirical fact which has nothing to do with optics. It is not easy to account for. It does not imply that brightness is a clue, indicator, or sign of distance; it only poses a problem. So far as the writer knows, this empirical fact has never been observed except under darkroom conditions where distance is relatively indeterminate and where presumably the impression of distance may be affected by minor determinants which are inoperative in a normal environment of illuminated background surfaces. The experiment needs to be repeated.

The cues for depth have not been reduced to their components and precisely specified. They need to be analysed in terms of (a) the retinal image and (b) the corresponding impression in the visual field.

In this and the preceding chapter those variations of the retinal image which underlie the principal cues for depth have been described. The theory is that they are retinal gradients and steps of ordinal stimulation and that they are geometrically precise. As stimuli, they can be tested for exact correspondence with impressions of distance and depth in the perceptual experience we have called the visual world. If they are stimuli, it should also be possible to put them in correspondence with impressions in the relatively depthless visual field. In common terminology, they should correspond to sensations as well as to perceptions, although the meanings of these terms have been reversed by the argument that the perceptual impression is the primary one, immediate and independent, and the sensory impression the secondary one, obtainable only by analysing the perception. What are these secondary sensory impressions, then, which depth and distance look like when they are not seen as depth and distance? What are the features of pictorial vision which parallel the three-dimensional features of normal vision? A classification of them should serve as a useful substitute for the list of the cues for the third dimension.

The sensory impressions which go with the perceptions of distance or depth over a continuous surface might all be called varieties of perspective. Those which go with the perceptions of depth at a contour have to be given a more arbitrary name; they will be called sensory shifts. The first correspond with gradients of adjacent stimulation on the retina; the second with abrupt rises or falls in such stimulation. The varieties of perspective can be listed somewhat as follows:

1. *Texture-perspective.* This is a gradual increase in the density of the fine structure, the spots and gaps, or the extended pattern of either a part or the whole of the visual field. There is a great variety of textures for which no adequate names exist. The increase in density may run in any direction but very often it runs upward in the field. The impression of density turns into an impression of depth or distance during ordinary vision. This type of perspective has been illustrated in Figures 23 to 26 and 32 to 35. It merges into the next type.

2. *Size-Perspective.* This is a decrease in the size of the shapes or figures in the visual field when it is considered as an array of color-patches. It presupposes contours, or figures on a background, each of which may have its own texture-perspective.

3. *Linear Perspective.* This is size perspective when contours are rectilinear. It is a gradual decrease in the spacing (the size or dimension) between either outlines or inlines in the visual field. Since the edges of things in man-made environments are so often straight and since straight lines are easy to draw, this kind of perspective is the one we have learned to notice and pay attention to. It has been illustrated in Figures 36 to 38. All these perspectives can decrease to a zero limit of size or spacing (or to a maximum density

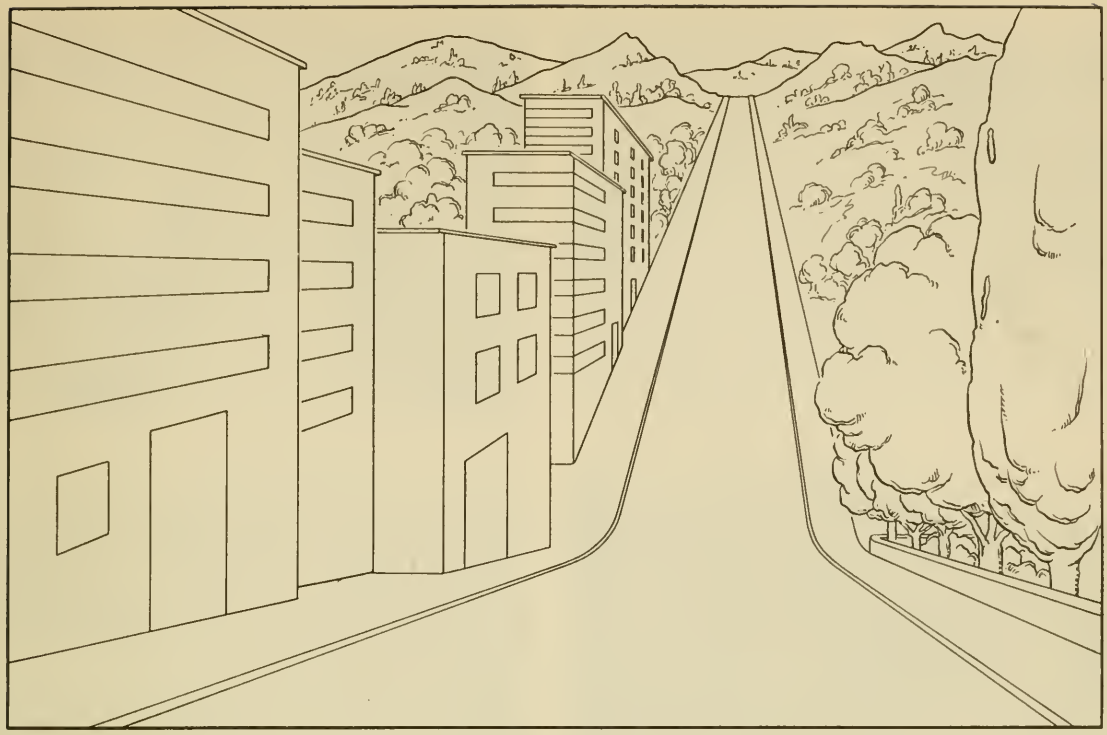


FIGURE 60. A Road Slanting Down and then Up

Can you find two gradients of converging lines in the drawing in addition to the gradient for horizontal converging lines? The picture was constructed with three different vanishing-points. *(Adapted from H. Buckley,* Perspective. *London, Sir Isaac Pitman & Sons, Ltd., 1947)*

of texture). A horizon in the visual field is, introspectively, a line at which these limiting qualities are reached.

When these three kinds of perspective, usually in combination, give way to the perception of a continuous surface such as the ground, the decrease in spacing or size (or the increase in density) gives way in turn to the perception of a *constant* spacing, size, and density. There results an experience which will be described in Chapter 9 as an *impression of constant scale.*

4. *Binocular perspective.* This is the perspective of double-images in the visual field and it is almost impossible to observe except at contours. The continuous change of double imagery for the elements of a texture, therefore, goes unnoticed. It is a change from a maximum of crossed double imagery toward uncrossed double imagery, as illustrated in Figures 48, 49, and 50. For any given plane surface there is a line of single imagery which is perpendicular to the direction of change and which always passes through the center of the visual field. The stimulus to which this kind of perspective corresponds is a gradient of the horizontal skew of one retinal image relative to the other, i.e. of the relative disparity at a given point. Both the gradient and the graded double imagery disappear when the observer closes one eye.

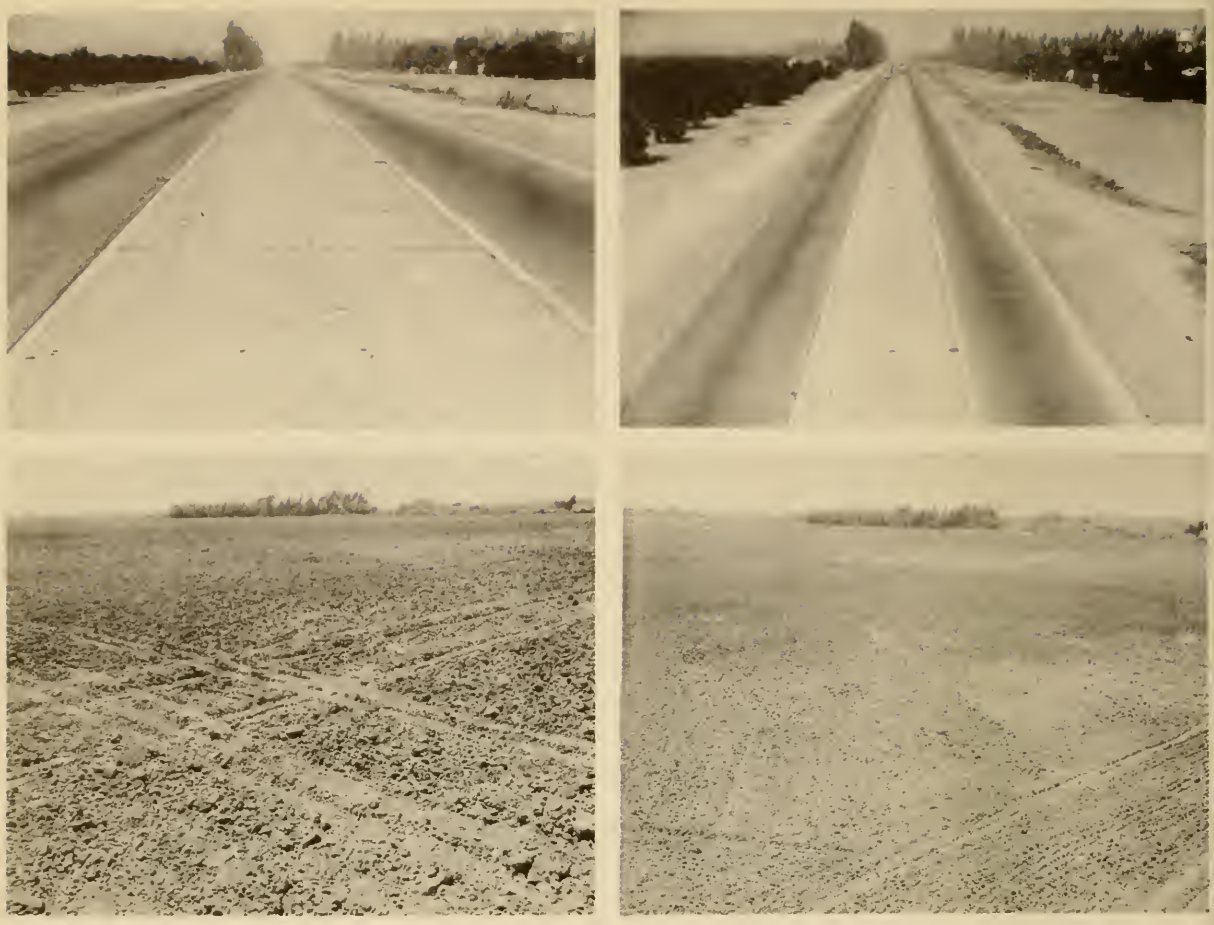


FIGURE 61. Two Scenes at Five Feet From the Ground and at Twenty-Five Feet From the Ground

The views on the left were taken at the height of a man standing, those on the right from a high ladder, but at the same spot on the ground. Note that there is a steep gradient of density in the left-hand photographs and a moderate gradient in the right-hand photographs. Note also that the more moderate gradient of density, in either case, produces an impression of a surface less inclined to the line of regard than the steeper gradient. Since the surface is interpreted as being level, the resulting impression is that the point of regard has moved upward. The right-hand photographs locate the observer high in the air, looking down. Another instance of the relation between the point of view and the slant of the surface is given later, in Figure 70.

5. *Motion-perspective.* This is a gradual change in the rate of displacement of texture-elements or contours in the visual field. The change is from motion in one direction through zero to motion in the opposite direction, and it also has a vanishing line, at right angles to the gradient, which passes through the center of clear vision. When the vanishing line of motion coincides with the horizon of texture, size, and spacing, the perspective of motion becomes easier to specify and describe. In that event the rate of deformation in the visual field simply decreases from the periphery toward the center of the field. The *directions* of the motion radiate from and

toward a pair of opposite poles (Figure 55) which are specific to the physical motion of the observer himself. The resulting perspective of expansion, skew, or contraction as the case may be is not difficult to observe, but it strongly tends to pass over into a perception of continuous distance or space. This perception is vivid and compelling.

Binocular and motion-perspective might be called perspectives of *parallax* whereas texture, size, and line-perspective are perspectives of *position*. Next to be listed are three which are independent of the observer's motion or position.

6. *Aerial perspective.* This is an increase in haziness, blueness, and desaturation of colors over the visual field. It has not been measured or precisely described. Unlike the other forms of perspective it is variable with the conditions of illumination and it does not rest on the geometry of optics. For this reason it seems improbable that it will ever prove to be a stimulus for the impression of distance, although it may be an indicator.

7. *The perspective of blur.* This is a decrease toward the center of clear vision of the quality of blur. Blur depends on texture elements and contours in the visual field. It is difficult to observe since the out-of-focus quality is never at the center of vision for a normal eye. When the lens is accommodated for any considerable distance, however, the gradient of blur tends to level off so that one may doubt whether it could serve as a univocal stimulus for the impression of distance. It is important only because it is more fundamental than the sensations of accommodation which are sometimes still listed as a cue for distance.

8. *Relative upward location in the visual field.* It has occasionally been suggested that the amount of background – the angular extent – between the lower margin of the field and a given object is a clue to its distance. The rule seems to hold for objects represented in pictures. Obviously, however, the clue is valid only when the background is taken to be the terrain rather than a wall or a ceiling. It serves mainly to illustrate the fact that the effective stimulus gradients in outdoor vision are usually those produced by the ground. Upward location in the visual field does not correspond to any gradient on the retina but only to a dimension of the retina on which a gradient of stimulation might occur. We shall recur to this phenomenon in Chapter 9.

The foregoing eight varieties of perspective all have reference to distance over a surface or an array of surfaces. There is another fundamental kind of distance perception, however, which was called depth-at-a-contour. The visual field contains only patches of color. What are the differences between these color-areas which yield depth?

9. *Shift of texture-density or linear spacing.* This is a change in the density of texture or the spacing of inlines which is sudden rather than gradual (Figure 41). It is usually coincident with a change in brightness or color in the visual field such as to produce a contour or a segregated form (p. 65). The parallel perception in the visual world is a recession in depth on the side of the increased density or decreased spacing. The effect is like

looking at a valley beyond the edge of a cliff on which the observer is standing.

10. *Shift in amount of double imagery.* This occurs when the texture-elements on one side of a contour are seen less doubled than those on the other side. The contour itself may be seen in single imagery (as *any* contour is when it is horizontal in the binocular visual field). When the observer takes a normal perceptual attitude he sees that side of the contour more distant which manifests a relative shift toward uncrossed double imagery.

11. *Shift in the rate of motion.* This occurs during a movement of the observer's head and consists of more rapid displacement of texture-elements on one side of a contour than on the other. If the contour is closed, the shape appears to move across the background. If part of the contour (the bottom usually) does not move across the background that part appears to be in contact with the background, or resting on the ground. The shift in rate of motion at a contour goes with a perception of one surface behind another, and the amount of retinal shift is theoretically an exact stimulus for the perceived distance between the surfaces.

The phenomenon of the superposition of objects is actually not a clue to the depth of objects but a perception which requires explanation. A man knows that a near object can partially obscure a far object but his retina does not, and the retinal explanation should be sought first. The preceding three factors supply such an explanation. The drawings of Figure 62, however, suggest that there are other factors, involving the *shape* of the contours, which help to determine superposition. There is no texture, double

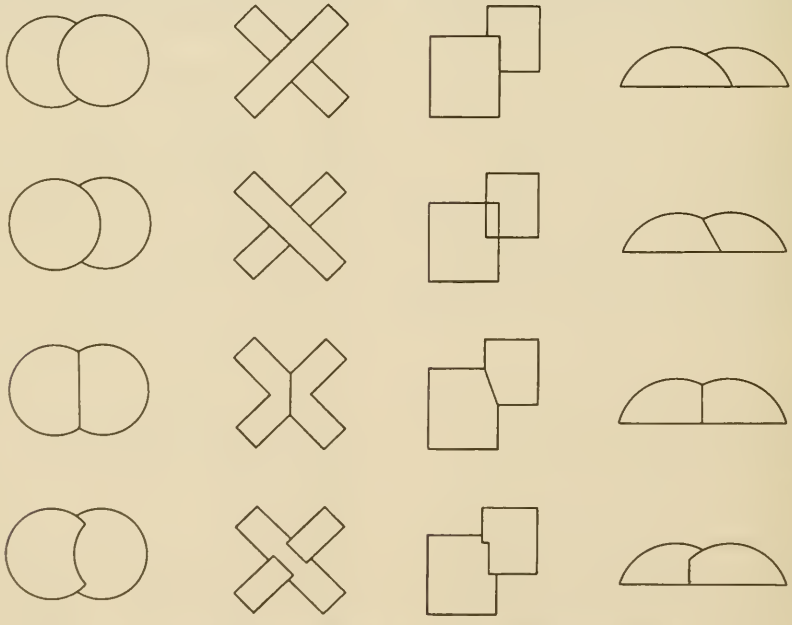


FIGURE 62. One Object in Front of Another

imagery, or relative motion in these drawings. They suggest the principle that the more complete, continuous, or regular outline tends to be the one which looks near. Is completeness, then, a sign or clue for distance?

12. *Completeness or continuity of outline.* We can reasonably assume that, if objects tend to have regular outlines, completeness, closure, or continuity[7] tends to be associated with the near side of a common contour and incompleteness to be associated with the far side. At the visual contour between an intercepting object and an intercepted object the side belonging to the completed outline usually has the coarser texture, the greater relative motion, and the greater crossed disparity. Consequently, a visual contour in isolation from these sensory shifts suggests greater depth for the incomplete outline and lesser depth for the complete one. The connection could be learned.

Besides continuous distance and depth-at-a-contour there is another major feature of three-dimensionality—the shape of an object in depth. The solid modelling of a surface, its protuberances, indentations, corners, curves, or flatness, is something which applies to backgrounds as well as to things. The principal explanation of this modelling of the world is probably the formula advanced in Chapter 6 for the slant of a bounded surface: that the *rate* of change of texture-density at any point in a projected image is proportional to the slant of the physical surface at the corresponding point.[8] An abrupt change of rate is the projection of a corner (Figure 40). A gradual change of rate is the projection of a curve. If this formula holds for the density of texture it should also hold for binocular disparity and rate of motion; the slope of these gradients is also geometrically linked to the slant of the surface projected (cf. Chapter 9). A supplementary explanation for the modelling of the visual world, however, is provided by the relation of *light and shade* to the convexities and concavities of the environment.

13. *Transitions between light and shade.* An abrupt shift in the brightness of adjacent regions within the visual field produces a contour, which is the necessary condition for a segregated shape or form. A shape in the visual field is coordinate with an object in the visual world. This rule is complicated, however, by the fact that there may occur shifts in brightness which produce not contours but the modelling of a surface. These are the brightness differences which we call differences in light and shade.

Light and shade, for reasons imperfectly understood, are not the same qualities in perception as white and black. The color of a surface and the illumination of it are perceived separately, although both must ultimately depend on the light-energy of the retinal image. Whatever may be the explanation of this paradox, the

[7] Ratoosh has given a mathematical formulation to the phenomenon of the continuity of an outline at its intersection with another outline (89).

[8] The slant of a surface, in this context, is its slant with respect to the line of sight not the line of gravity. It is the angle of confrontation of a surface.

transitions between light and shade seem to be capable of giving a surface the quality of shape in the third dimension as distinguished from its shape in two dimensions. This fact has been illustrated in Figures 43 and 44.

The relationship between these transitions of shading and the corresponding depth-shapes is not well enough known to be specified in psychophysical laws. A gradual transition, it is true, yields a curved surface and an abrupt transition an angled surface, or corner, but whether this depth-shape is concave or convex depends on complex factors. The geometry of light and shade changes as the direction of the illuminating light changes. Consequently it is impossible to decide whether this variable of the retinal image should be considered a stimulus-correlate for depth or not.

The traditional cues for depth, to summarize the last two chapters, can be restated as variables of the retinal image. They can also be described as they appear when one observes his own visual field, and can even be called sensations so long as one is careful to remember that they are only the visual symptoms of stimulation and not the causes or elements of perception.

Eight of the thirteen listed can be thought of as stimuli for perceptions of space. The remaining five are better conceived as probable signs, secondary to the others, or as having doubtful status.

8

The Problem of the Stable and Boundless Visual World

The Stable Visual World . . . . The Problem of the Unbounded Visual World

The distance, depth, and solidity which characterize our visual perception have been the subject of the preceding chapters. The visual world described in Chapter 2, however, has not yet been accounted for. Our experience of things is stable, upright, and unbounded although by rights it should not be so since neither the retinal images nor the visual field has any of those qualities.

The feature of the visual world which, as much as its depth, impels most of us to the conviction that it is there and is not an illusion or a picture is its stability. Samuel Johnson is said to have refuted Bishop Berkeley's doctrine that the world was all in our head by the simple argument of kicking a stone. He implied an ultimate trust in his muscular and tactual sense. An equally good reason, however, for confidence that things are sensed as they are is that they stay where they are. In view of this stability can the instability of the retinal image, shifting over the retina with every movement of the eye, be reconciled with the assertion that the retinal images are the basis of the visual world? The problem has been encountered before. The assumption that the visual stimulus was a static momentary image was only provisional and it should now be abandoned for good. We need to consider images as affected by exploratory eye-movements.

The Stable Visual World

As we asked in Chapter 2, why does the world not go shooting about as the observer shifts his fixation from one object to another? Helmholtz was aware of this problem, as he was of nearly all the other problems in vision, and formulated it as the question of how one recognizes stationary objects as such, in spite of the shifting of their images over the retina during eye movements. Helmholtz' answer was typical of his general theory: one sees objects as not changed because one learns to regard the retinal shift as merely the "sensory expression of the ocular movement" which actually corresponds to no change of the objects (53, III, p.63). This

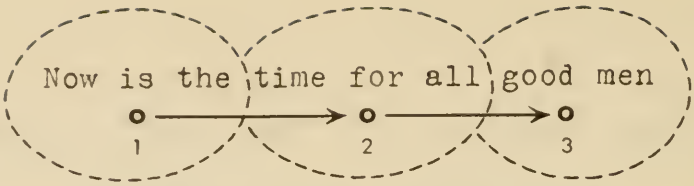


FIGURE 63. Successive Superpositions of Retinal Images

amounts to saying that we learn to see the change as not a change.

The problem can be formulated in different ways and it involves a set of related problems – all equally puzzling. It might seem that a quick and easy way of solving it would be to reply that the question asked is meaningless. To speak of the world, or of objects in general, as being displaced does not make sense, for what could it possibly be displaced with reference to? This reply, however, is no solution. The visual world is a response of the organism – amazingly complex but still a response – and it could be displaced with reference to the tactual world, that is, to the perceived surface felt by one's feet. Anyone who pushes on the outer side of his open eye (p. 31) can see the visual world being displaced, and anyone who has ever been dizzy knows that it can move. Why does this kind of perception not occur during normal exploratory eye-movements?

Consider the diagram of Figure 63. If the series of displaced retinal images which occur during the reading of the printed phrase were to be thought of as persisting after each fixation the result would appear as shown in the second line. The numbers in the first line indicate fixation-pauses. The dotted ellipses represent the region of clear vision immediately adjacent to the fovea. During a single second of reading, three fixations of the eye may occur, and the retina will have been stimulated by a complex of three overlapping images. The fact to be noted is that successive retinal excitations during eye-movements do not fuse with one another, as do successive exposures on a film or retinal after-images, but are integrated in a very different way.

The patches of color which make up the visual field continue to have a fixed direction-from-here when the corresponding patches of focused light are displaced over the retina. The absence of blur during saccadic eye-movements might be explained as due to a momentary central inhibition of visual experience during the interval (56; 121, Chapter 23), but this would not account for the absence of displacement. There must exist some compensatory effect produced by the eye-movement itself. Whatever neural process initiates the jerk of each eye from the old to the new fixation point must, at the same time, shift what

used to be called the "local sign" of every anatomical retinal unit. This innervation of the eye-muscles might be supposed, in geometrical terms, to add a constant displacement of opposite sign to the whole anatomical pattern of excitations or, what is the same thing, to transpose the axes of reference which give it a right-left (and up-down) quality in experience. What the physiological correlate of local signs or of reference axes might be we do not know. But we do know that the eyes are mobile relative to the head, the head relative to the body, and the body relative to the ground, and we know that the posture of all these organic systems with respect to the ground is maintained by reflex mechanisms. A chain of reference is thus provided by which an absolute standard of "straight ahead" might be established.

The requirement of some such theory as this has been recognized by psychologists for a long time (53, III, p. 570). As here formulated, it implies that a saccadic eye-movement involves both the neural activity producing the eye-rotation, which varies between opposites (right-left; up-down), and the neural activity produced by the shift of the image over the retina, which also varies between similar opposites. The movement of the eye and the movement of the image are by necessity reciprocal. A combination of two processes each reciprocal to another is obviously a constant, and such a constant product might then account for the non-displacement of the visual scene.

Some evidence for this theory may be obtained by trying it against a number of experimental observations. In terms of visual experience, the theory states that when the eye actively rotates to the right the scene should appear to be displaced to the left, but that it is not displaced because there is a compensatory shift to the right which cancels it. (Only a single eye is referred to, but the argument applies equally when both eyes are functioning.)

(1) If the eyeball is mechanically rotated toward the right the scene should appear to move to the left and will do so, according to the theory, because the compensating shift is absent. The reader may be able to verify this prediction for himself by pushing his eye with his finger.

(2) If the eye actively rotates to the right after a clear negative after-image has been established in the center of its field, the after image itself should not appear to move at all because it is not displaced on the retina; according to the theory, however, it will appear to be displaced to the right because the compensatory shift is present with no retinal motion to be cancelled. This result is a matter of common observation.

(3) If the eye is innervated to rotate to the right but actually does not move because of a paralysed eye muscle, the scene should not appear to move on any grounds of retinal stimulation but, according to the theory, it should move to the right because the compensatory shift to the right is present, as in the last experiment. That it does just this is apparently well known to ophthalmologists and the phenomenon was described by Helmholtz (53, III, p. 245). If the central neural mechanism is normal, a patient who tries to move his eye to the right in these

circumstances has an illusion that things move to the right.

(4) One final observation may be given. If the eyeball is mechanically rotated to the right after a negative after-image has been produced in the field, the after-image should not appear to move on account of any compensatory effect. Only if one assumes some such nonsense as that visual sensations are projected outward from the eye like the beam of a magic lantern should it move at all. As nearly as I can observe, the after-image in this experiment is stationary. The scene on which it is superposed, however, is displaced to the left, as in the second observation. Helmholtz, as one might guess, observed this phenomenon too, although he did not interpret it, and his report is that the after-image does not move although the screen on which it is projected does (53, III, p. 244).

The implication of all this is that the directional stability of the visual world might be a product of activities which are inverse to one another — the activity of the motor centers for eye-movement and the sensory centers for retinal motion. The visual world then could possess a standard straight ahead direction-from-here and a constant direction at all other points. Koffka has discussed the problem in the light of its history (67, p. 384 ff.) and suggested an explanation essentially similar to the one given above.

The Upright Visual World. The visual world is stable in another respect in addition to direction. The world always appears upright and, little as it may seem so, this is actually a very curious fact. The world has the useful but unappreciated property of remaining horizontal and vertical in visual appearance even when we lie on our side or tilt our head and thereby turn our retinal images around their centers like a couple of radio dials. Although the contours and edges considered as color-patches within the visual field usually appear to tilt under these circumstances the visual world remains linked to the vertical and horizontal axes of gravity. The eyes, it is true, do roll in their sockets a few degrees to compensate for slight tilts of the head to right or left but they do not counter-roll enough to put a twist in the optic nerve. Why then does the world not tilt when the image does?

The answer seems to be that when a man lies on his side or puts his head in an abnormal posture he gets two forms of stimulation which vary concomitantly, first, the rotation of the retinal image around its center, and second, the off-balanced pull of gravity both inside the inner ear and on the bilateral musculature (42, p. 303). The gravitational and the visual stimuli vary in a reciprocal relation to one another; the product of these two variations would be a correlate for the upright visual world.

Evidence for such a theory as this is no more than indirect since the existing experiments on the problem have not been based on it. Most of them have sought to prove either that visual stimulation was a primary and posture a secondary determinant of upright vision or vice versa. Although asking whether posture is prior to vision or vision prior to posture is like the juvenile puzzle of the hen and the

egg, a good many students of the problem have been guilty of it, including the writer himself (42).

Koffka, (67, p. 213 ff.) believed that the upright character of the phenomenal world despite a tilted image was explained by a tendency for the main lines and contours of the image (floor, walls, horizon) to become the reference-axes on which an impression of tilt would have to depend. He based the theory partly on a report by Wertheimer (118) that if an observer keeping his head upright looked through a tube at the surface of a large tilted mirror, the tilted scene gradually righted itself in perception and, after a period of time, looked normal. This observation has been questioned, however, by the writer who repeated but could not verify it. He produced an optical rotation of the visual world by Wertheimer's mirror-and-tube arrangement and on another occasion by a pair of reversing prisms placed before the eyes, but the tilt did not disappear (42). The observation, therefore, remains in doubt. Koffka's theory, in any event, went somewhat beyond this phenomenon to assume that the principal lines of the visual field when the head and eyes were tilted but the physical environment was upright immediately determined the vertical and horizontal axes of the phenomenal world. The main lines of organization are seen as the main directions of space. The implication is that vision, and vision alone, accounts for the vertical and horizontal appearance of things, and that only the laws of sensory organization need be invoked. It should be noted that Wertheimer's phenomenon was a case of discrepant or conflicting cues, whereas Koffka's application was to the ordinary case of covariant or reciprocal cues.

Insofar as Koffka was demonstrating with this theory that anatomically fixed axes on the retina – the nervous connections of a hypothetical pair of lines engraved on the retina – are no explanation of why we see things as vertical and horizontal, he was surely correct. The theory of Cartesian coordinates on the retina is a myth. But our visual sense of the vertical and horizontal, unlike our visual sense of texture or contour, cannot be ascribed to the prevailing main lines of retinal stimulation alone. Pure order in the image might explain a contour but not the orientation or direction of a contour. The visual vertical and horizontal (when apprehended correctly) have reference to the direction of gravity. The direction of gravity is reliably indicated by postural stimulation, in ordinary life. Neither the uprightness of the visual world nor its stability can be understood if we confine our attention to a shifting, swivelling retinal image. These features of space are inseparable from the feeling of the ground under our feet and the feeling of standing up, of moving about, and of looking. The tactual and kinesthetic stimuli which arouse these feelings ordinarily co-vary with the visual stimuli and the product is something which is neither visual nor postural. It is not that we stand with our eyes or that we see with our muscles, but rather that we both stand in and see a steady world.

The writer's discussion of the problem, written in collaboration with Mowrer, has already been referred to. It made the error of assuming and trying to prove that

postural cues were logically and genetically prior to visual cues in determining how we maintain postural equilibrium and perceive an upright visual world. A motor theory in opposition to Koffka's visual theory of the perceived vertical and horizontal was proposed. The evidence considered was taken from experiments on conflicting visual and postural cues, the unrecognized assumption being that when two kinds of determinants seem to underlie the same kind of experience one of them must be the true basis for the experience and the other must be secondary.

The props for either a motor theory or a visual theory have recently been knocked out by the studies of Asch and Witkin (2, 3, 4) who repeated the observations made by Wertheimer and the writer with more subjects and under more varied conditions. Several of their experiments made use of a boxlike room open on one side, into which the observer looked, and which could be physically tilted while the observer himself stood upon a level floor. He had to adjust a movable rod within the room to what he thought was the true vertical. Their experiments show that when the direction of the postural vertical and the direction of the main lines in the visual field do not coincide, the usual result is not a complete domination of either but a somewhat unstable compromise between the two directions. Whether one or the other set of discrepant cues tends to dictate the observer's perception of the vertical depends on the circumstances and on the observer. The implication is that postural and visual stimulation are both determinants of the upright character of the visual world. When they are concomitant, as they are in voluntary change of posture in a physically normal environment, perception of the vertical remains correct and equilibrium is maintained. When, however, they are set in conflict with one another by an ingenious experimenter, or in the "haunted-swing" of an amusement-park, or during the act of banking an airplane, the perception of the vertical tends to be unstable and is likely to be objectively incorrect. In this situation, the organism is forced to search for reliable cues to the direction of gravity, and the perception is objective only to the extent that reliable cues are discovered and used. The flier, for instance, usually has to learn that the objectively reliable cue in banking an airplane is the visual horizon. He always has to learn that the objectively reliable cues in night flying are the instruments.

The question of the reliability of the cues for space perception, including the cues for depth and distance, arises when there are discrepancies among them. In the preceding chapters there has been no mention whatever of this question. It is a valid question, of course, and it might just as legitimately be asked about *stimuli* for space perception, as it is asked about *cues* or *clues* for space perception. Brunswik gives it great emphasis in his treatment of the objectivity of visual perception (15). The neglect of it in the present book stems from an intention to concentrate on the theory of those spatial perceptions for which the determinants are supplementary to one another, not discrepant, and for which the stimulus conditions are optimal rather than im-

poverished or inadequate.

The Transposition of the Retinal Pattern. The theory of a reciprocal effect upon the excitations set up by a moving image on the retina still leaves another paradox unsolved. A moving image as an event of physics and optics can be understood by analogy with the displacement of a camera film relative to the image focused on it. But a moving image as something which initiates vision in the retinal receptors can be understood only by such analogies as our moving electric sign (p. 56). The pattern of excitations in the retina, the optic nerve, and the brain implies a set of localized excitations which are anatomically displaced by an eye-movement. The corresponding shift of the patchwork of colors in the visual field relative to its margins also implies a set of localized excitations. But then why are the steps and gradients of excitation unaffected by the displacement? The difficulty was first formulated with reference to geometrical forms, and the question was why a square could still be seen as a square after a wholly different set of retinal elements had been excited. This is the problem of the transposability of forms. It led first to the doctrine of form-qualities or special sensations of form and, when this was recognized as naming but not solving the problem, it provided a point of departure for the whole of Gestalt theory (Chapter 2). The solution offered by Gestalt psychology was that a process of configuring or organizing the excitations could be assumed in the central nervous system.

Evidently, however, the problem is more fundamental than the problem of form as something possessed by objects. Actually, what is transposable on the retina is the whole panorama of the retinal image: the steps of texture and shading, the gradients of variation, and the inflections of these gradients. As contours, surfaces, depths, and objects they yield a stable visual world. As stimuli, therefore, they must be equivalent when one set of retinal receptors gives way to another. The paradox goes deeper than the equivalence of a square; it is the identity of the retinal image itself which comes into question. What equivalence can there be between two sets of nerve-cells which bear a purely anatomical relation to one another?

The answer suggested in Chapter 3 was simply that there exists an ordinal pattern of excitations. Steps, gradients, and inflections of a gradient are mathematical facts which are the same no matter where they are located. In an array of adjacent nerve-cells an increase in excitation from one cell to the next is a kind of serial order which can be transposed without causing a permutation or an inversion of the order. The set of numbers ..345.. are similarly equivalent in order to the set of numbers ..789.. but not to 987 or 879.

Can one suppose that the organism is capable of reacting specifically to an order of excitations? Can a perceptual impression correspond to what seems to be a mathematical abstraction? We do not know, but the suggestion may be more reasonable than might at first appear. What, exactly, is a stimulus? Physiologists and psychologists often assume that

it is a kind of physical energy acting on a single receptor. We shall name this the one-stimulus-to-one-receptor theory and we note that it gets into trouble when a sensory surface composed of adjacent receptors is considered. Moreover the physiology of single receptors, or what is known of it, suggests that the effective stimulus is a change of energy acting on the receptor. Some psychologists are led to the assumption that the organism responds differentially to variables of stimulation but not, strictly speaking, to isolated stimuli. In psychophysical experiments, what the observer reacts to is the relation between two things or events and almost the whole of our knowledge is founded on judgments of "greater", "less", or "equal" (102). Change, variation, and relation are no less abstractions than is adjacent order. Perhaps the concept of ordinal stimulation only makes explicit what has often been assumed.

We must not forget, of course, that there is a fixed anatomy of the retinal receptors and that the appearance of the visual field bears witness to it. A transposable order is not, then, an abstract, timeless form of the sort that Plato believed to underlie the world. A transposable order would be meaningless unless it were embodied in an order of adjacent anatomical elements.

The Transformation of the Retinal Pattern. The retinal image, as we have seen in the first part of this chapter, may undergo not only a transposition but also a de-

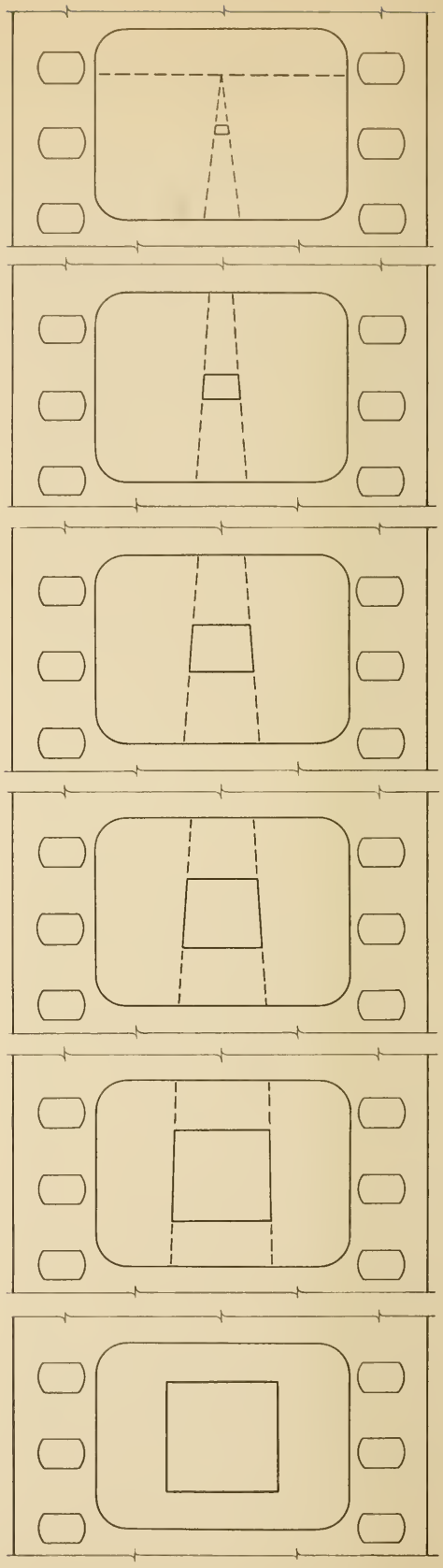


FIGURE 64. Successive Transformations of the Image of a Square of Pavement while a Person is Walking along a Street

formation. As a whole and in every part, it undergoes deformation whenever the head moves, but the stability of the visual world remains unchanged. The only effect on perception is an impression of subjective locomotion. Mathematically, this change of the image is analogous to a projective transformation. This is a kind of transformation which, as it were, stretches but does not tear the form. Point-to-point and line-to-line correspondence is preserved, but the shape is altered. Transformations are usually represented on a plane, however, whereas the retinal image is a projection on a curved surface. As a matter of fact, the actual retinal image on a curved surface is related to the hypothetical image on a picture-plane only by such a non-rigid transformation (see Figure 28, Chapter 6). The geometry of transformations is therefore of considerable importance for vision, and it is conceivable that the clue to the whole problem of pattern-perception might be found here. If we add to the classical problem of the transposability of the retinal image the requirement that it must be transformable as well, the problem may emerge in a new light.[1]

The transformation of a given pattern, mathematically defined, does not simply destroy the pattern as one might at first suppose. A transformation is a regular and lawful event which leaves certain properties of the pattern invariant. The study of the invariant properties of geometrical forms which have undergone transformations is known as topology and the study of projective transformations is known as projective geometry. The writer cannot claim to be an expert in either of these branches of mathematics. The general application of such principles to the deformation of the retinal image during locomotion by the observer has not been attempted, so far as I can discover, by any psychologist or mathematician. At the crudest level only, then, it seems clear that not every pattern can be regularly transformed into any other. A doughnut cannot be transformed into a cube by any continuous process. But a transformation can be applied to a given pattern without affecting certain of its general properties. The skull of a chimpanzee can be transformed into the skull of a man, as D'Arcy Thompson first suggested, by a fairly simple geometrical operation (105, p. 1085).

Moreover a series of transformations can be endlessly and gradually applied to a pattern without affecting its invariant properties. The retinal image of a moving

---

[1] According to Courant and Robbins (25), any mapping of one figure on another by either a central projection or a parallel projection (or a succession of such projections) is a projective transformation. In such transformations straight lines are invariant, intersections are invariant, and the order of points is invariant. Lengths and ratios of lengths are altered, but the ratio of two ratios of length (the cross-ratio) is invariant.

Cassirer has suggested that the geometry of transformations provides a solution for the puzzle of perceptual constancy. But he assumes that the process of perception must necessarily involve a search for constancy, an objectification of sense data, or a discovery of the invariant properties of shapes which have undergone transformation (21). What he suggests is that the mind transforms the retinal images, whereas the suggestion above is that transformations of retinal images are equivalent as stimuli.

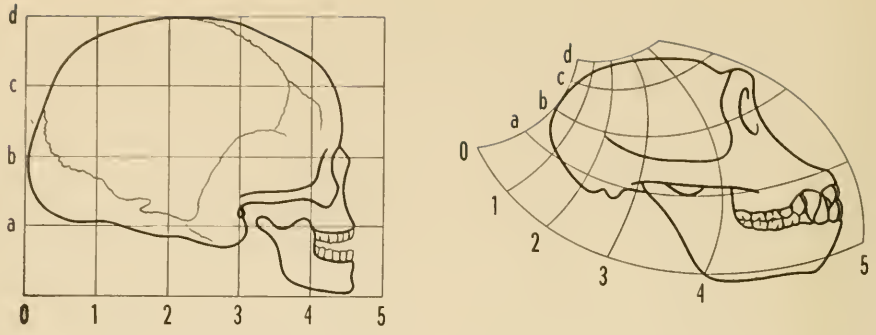


FIGURE 65. Human and Chimpanzee Skulls

(Redrawn from D'Arcy Thompson, Growth and Form *(Macmillan, 1942)*

observer would be an example of this principle. Perhaps the clue we are seeking lies in the invariant properties of such a continually changing retinal image. Only these properties would be capable of providing the stimulus basis for a stable and unchanging visual world. A man who walks about in his back yard never has the same retinal image of it twice unless he comes back to the same spot, puts his head at exactly the same point, and fixates in the identical direction. Nevertheless he perceives the same environment throughout his wanderings. He also, of course, perceives his wanderings. If the retinal images have constant properties and also undergo a continuous transformation perhaps we can account for both his visual perception of the environment and his visual perception of locomotion in it.

The suggestion is that we need to understand the geometry of the transformations of the retinal image in order to explain why its successive changes can all be equivalent for perception. The changes co-vary with muscular action during active locomotion, as we have noted, but during passive locomotion little or no bodily stimulation exists, and hence we probably cannot fall back upon it for the whole of the explanation. Certain features of the retinal image are preserved during a locomotor transformation – order, continuity, points to points, and straight lines to straight lines – while certain other features are not preserved – angles, the congruence of shapes, and the metric properties of lines. The features that are preserved may be the mediators of a stable visual world and the features not preserved the mediators of the visual impression of motion.

Evidently the retinal image needs to be considered as a changing event in time, and the analogy with a static picture is thoroughly misleading. The order into which the image can be analysed ought to include not only the *adjacent* order which has mainly concerned us but also the *successive* order which it equally presupposes.

The Problem of the Unbounded Visual World

The problem of why the visual world is stable is actually related to another problem—why it is unbounded. Ordinary visual perception is not delimited by an oval-shaped boundary, nor does it have a clear center and a vague fringe. These are the characteristics of that unusual kind of visual experience, the visual field, which we get when we fixate a point and take note of the experience, concentrating on how it feels to see. As we noted in Chapter 4, the center, the fringe, and the boundary are reflections of the optic anatomy: the fact of the fovea, the thinning out of sensory cells, and the margins of the retina.

Prolonged fixation interrupts the normal course of vision by inhibiting the exploratory movements of the eyes. As has been emphasized, the eyes are extremely mobile organs. In the activities of everyday life the center of clear vision will shift as often as a hundred times a minute, and during reading or while driving a car the rate of fixations will exceed this figure (121, Chapter 28). Can we find an explanation in the facts of ocular movement for the absence of the above characteristics in the visual world — its lack of boundaries, its more nearly uniform clarity and its possession of what we might call a panoramic character? The depth and distance of the visual world can be accounted for. Its stability or unchanging direction-from-here one can at least struggle to explain. But why is it unbounded when the stimuli consist of fragmentary images?

The Function of Saccadic Eye-Movements. The movements of the eyes have been studied for many years. In the half century since Dodge invented a photographic method for recording them objectively and described their fundamental types (26), a great deal has been learned. The shifts of fixation when we look at pictures, for example, have recently been carefully studied by Buswell (17). A typical record of how the eyes behave is shown in Figure 66. Note that the fixation point moves all over the picture, but not in an orderly fashion. There are frequent zigzags and irregular jumps. No simple or clear relationship appears between the order of fixations and the order of elements which compose the picture.

Some theorists have believed not only that eye-movements could account for the perception of locations and distances in space but also that they might explain the perception of patterns and two-dimensional forms. The eyes might be supposed to trace the outlines of things and thereby provide cues to their shapes. But the actual records of eye-movements have never supported the theory. They tend to be more or less like the record illustrated. A related but more plausible theory is that the composition of a painting is something which enforces a particular sequence of fixations in the onlooker, the eyes being drawn from one point to another by the spatial order of lines and masses in the painting. It is widely believed by artists and art critics. But the theory cannot be literally true in the face of the illustration given, and if there are obscure relations between pictorial com-

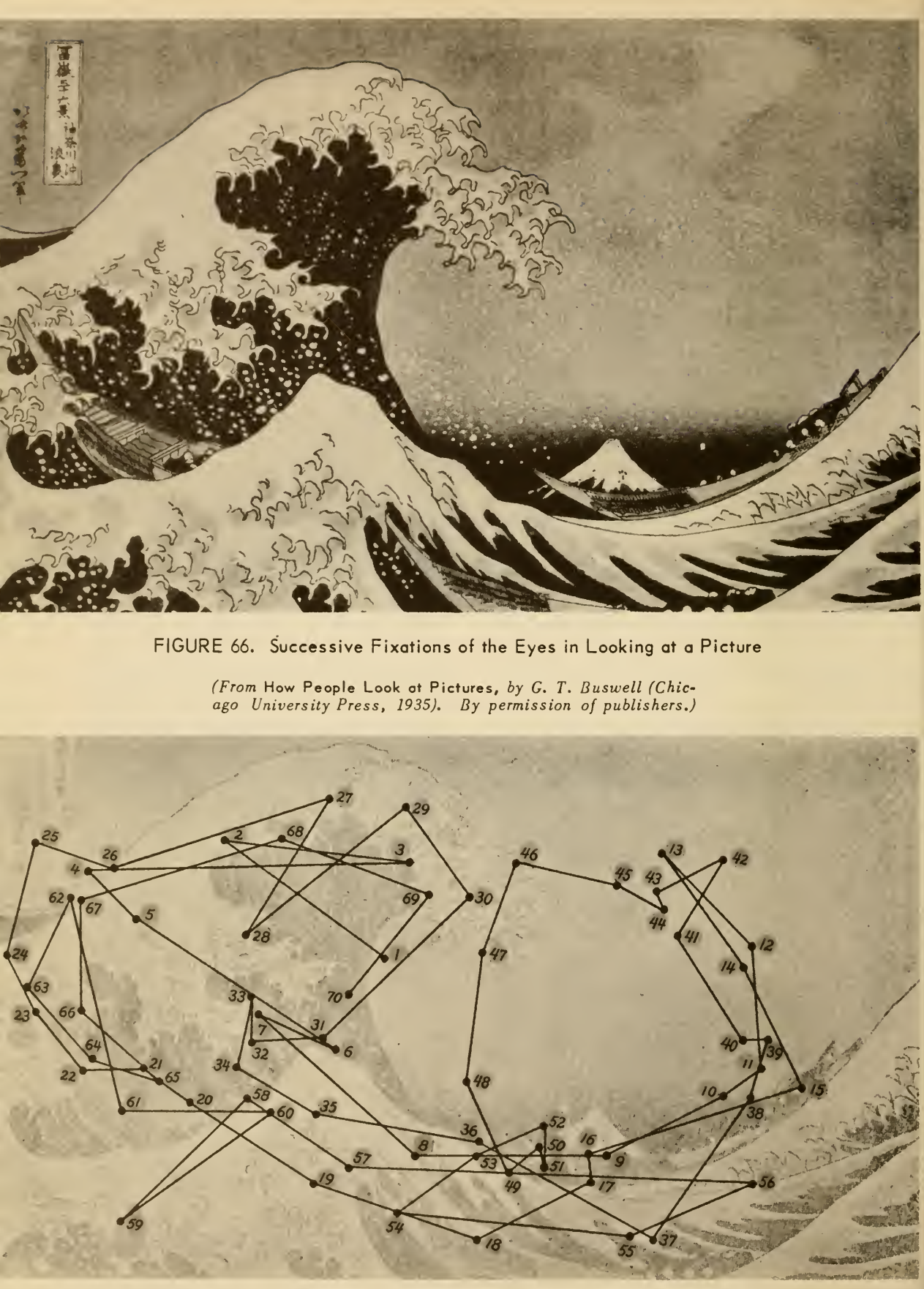


FIGURE 66. Successive Fixations of the Eyes in Looking at a Picture

(From How People Look at Pictures, *by G. T. Buswell (Chicago University Press, 1935). By permission of publishers.)*

position and fixation sequence they have not yet been discovered. The evidence suggests, for vision in general, that the sequence of fixations is more or less random, and that each single fixation tends to fall on an object of attention (rather than the background) and on a point of interest within an object.

The function of saccadic eye-movements is probably not that of perceiving or appreciating any kind of order in the retinal image. It is more probably to enable us to register the environment in two ways which, in the use of a camera, are mutually exclusive. Eye movements enable us to see the environment over a very wide angle; fixation gives us narrow angle vision for fine detail. The eyes function so as to obtain the advantages of both a very wide-angle camera lens and a telescopic or narrow-angle lens. With the same eyes, we can either look all around or look at a single object.

Helmholtz understood some such function as this. The intent of vision, he said, is to see as distinctly as possible, with both eyes, various objects or parts of an object in succession (53, III, p. 56). The sole purpose of the mobile eyes is to permit a kind of light-absorptive pointing. Any kind of ocular movement which does not produce the distinct imaging and centering on the retina of successive parts of the environment in each eye cannot be performed. We cannot move our eyes evenly and slowly across the environment. We cannot move one eye upward and the other downward. We cannot converge the eyes without accommodating the lenses. We cannot hold the eyes still with reference to the head when the head rotates. All we can do is to let our astonishing ocular reflexes take care of the intricate task of looking.

One slight but important exception to Helmholtz must be taken. We must not forget the total retinal image. What gets centered on the fovea are successive small regions of an extended image covering 180° which, as an image of optics, is equally distinct everywhere. These regions correspond to the so-called objects of attention or points of interest. It is therefore not true to say that what happens on the retina is a succession of fragmentary images as we did in the beginning. There occurs a succession of *overlapping* images, 180° wide, only the centers of which are registered by the nervous system in fine detail.

The Problem. We are forced to conclude that the visual world cannot be perceived all at once. The process of perception cannot rest upon the image of a single fixation such as yields a momentary visual field. To see more than this takes time, and requires a succession of images. The product of these successive impressions, however, is such that, paradoxically, all awareness of the succession has been lost. Unquestionably the panoramic visual world depends on a temporal series of excitations and just as unquestionably the succession of the excitations is not represented in the final experience.

Evidently, the abstractions which we call space and time are not as distinct as they have been assumed to be, for space cannot be apprehended except in time. These abstract ideas, however, arise out of concrete experiences, and it is these experiences that concern us. At the

stimulus level we have a succession of overlapping images. At the level of experience we have a panoramic world, all parts of which are concurrent. In ordinary visual perception there is no sense of either the *sequence* of fixations or the *time-lapse* between them. This is to be contrasted with auditory perception, where the sequence of experiences is identical with the sequence of stimuli. A heard sound begins, proceeds, and ends, but a visual object does not, although its patch of color enters, remains, and leaves the foveal region. 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. The stream of your consciousness is more like the multiple stream of events than it is like the stream of fixations.

The successive retinal images, it must be remembered, are not distinct entities like the pictures of a comic strip. They overlap and, to that extent, are transpositions of the just-preceding images. Instead of saying that the visual world is based on a succession of images, therefore, it is possible to say that it is based on a continuous but changing image. The retinal image may be defined as a pattern or as a process. The problem, accordingly, may be formulated in two ways. First, how are successive patterns on the retina integrated to form an unbounded visual world? Second, can a compound of adjacent and successive order be defined which would provide a stimulus basis for the visual world? We shall take them up in order.

The Integration of Successive Images. Successive excitations of the retina must be integrated by memory. How else could our present perception of the extended world arise except through memory of just-past glimpses? The kind of memory required to explain perception is not, of course, that commonly understood by the term. It is often called primary or immediate memory. It is the kind of memory which makes possible the apprehending of a melody. Its similarity to perception has been emphasized by calling a melody a temporal Gestalt, and there are some striking parallels between such experiences as visual patterns and auditory rhythms (67, ch. 10). Primary memory is the kind which enables us to hear a series of seven or eight spoken digits and repeat them immediately as if we were reading them from a page. Beyond seven or eight, however, we begin to have trouble, and the limit for a given individual is known as the memory span. Similarly, a certain number of strokes of a clock can be apprehended without counting them; five o'clock does not have to be counted whereas twelve o'clock does (10, ch. 5). All the experimental evidence implies that successive stimuli somehow endure and are integrated to a single experience. Introspectively observed, the conscious present does not seem to be a point where the future meets the past but a considerable range of events. Although the concept of primary memory has been derived from the study of auditory perception, it must apply with even greater force to visual perception, where successive integration

is so complete that the observer can be wholly unconscious of his fixations.

We could assume that each fixation is followed by a primary memory image which fades very slowly, each segment of each image being fitted to every other by the compensatory displacement mechanism already described. The result would be a panoramic world which, at any given instant, will have faded least in memory at the point of regard and most at the point directly behind the head (depending on how recently one has looked around). The fading would not be a decrease of clarity like that observed from the center to the periphery of the visual field but only a decrease in reportability or that quality of experience which goes with recency, and hence the visual world can be said to be everywhere equally clear.

What the physiological basis of memory might be, we do not know. Each momentary excitation might be assumed to leave some after-effect or trace in the brain, but the working out of a complete theory of successive traces is faced with many difficulties (67, ch. 10-13). Even a theory comprehending the primary memory-image, the eidetic image, and the image of imagination is lacking, and the relation of these to the after-images and other after-effects of prolonged fixation is unknown.

Sequence and Scene. It is interesting to compare the integration of the visual world with the way in which a series of motion-picture shots builds up a scene. The first kind of integration is an unconscious process. The second kind is one whose rules have puzzled even the greatest of film artists. A shot, which may last only a few seconds, is the unit of the motion picture camera, and a spliced series of shots comprises a scene, which is the unit for telling a story. It is no accident that a series of shots is called both a scene and a sequence in the terminology of films. A good film editor can make us see a scene by means of a sequence. A concurrent situation combining, for instance, persons and places can be created by successive camera views if their sequence is consistent with that of natural apprehension rather than that of some arbitrary logic. Just as the visual world must be scanned in order to be seen, so the situation must be represented in some little understood natural order. Each camera view persists in memory, but that is not sufficient for a clear understanding of the situation; the fixations of the camera on the points of interest must conform to certain rules.

A traditional film sequence consisting of a long shot, a medium shot, and a close-up reflects merely the tendency of an onlooker to look around in a new situation before he narrows his attention. The instantaneous cut from one shot to another is analogous to a saccadic eye-movement. Scenes, on the other hand, are separated by fades, dissolves, or other visual devices for representing a time-lapse. Between the heroine bound to the railway tracks and the hero galloping across country to the rescue occurs a cut which makes the time of the two events simultaneous. The alternating order of shots in a chase sequence is easy to understand; where the difficulty comes is in arranging the sequence for a scene of complex human interaction.

Our hypothesis implies that one guide to motion-picture cutting would be to determine the sequence of fixations which an onlooker would unconsciously perform if he were taking in the scene as it would really occur. What would he look at first, what next, what would he then examine? The formula would not be literal, since the camera cannot look in the way that the eyes do, and the limitations of the camera must be allowed for. Such a formula might serve as well, however, as the obscure intuitions and traditional practices which govern most film-making, and it has the virtue of being subject to experimental study.[2]

The problem of how to produce a scene from a sequence is not confined to the motion pictures. The novelist, and in fact all artists but graphic ones are confronted with it. Abstractly, the question is how to produce a series of impressions which, for the reader, the hearer, or the observer, creates an objective world flowing in objective time. A successive order of impressions must be made to yield an adjacent order of things.

The Compound Order of the Changing Image. The foregoing explanation for the panoramic character of the visual world is satisfactory for a stationary observer but the case of a moving observer puts a bad strain upon it. Although the stationary observer may be considered to have a succession of retinal images, the image of a moving observer undergoes continuous transformation. Must we analyse the latter into a succession of distinct patterns and assign a specific memory-image to each? How many memory images shall we assume — one for each ten seconds of locomotion? One for each second? One for each thousandth of a second? It would be simpler, if this were possible, to define the retinal image as a continuous process in time and leave the memory images entirely out of account.

Consider the perception of simple motion — the movement of an object across a visual field, for instance. If we were to be consistent, the integration-by-memory theory would have to be applied to this experience. Only if each retinal element successively excited leaves a trace in immediate memory could the past excitations be held together to yield the impression of motion. We would have to conclude that motion is never perceived, strictly speaking, but only remembered. Contour, form, color, surface, depth, and distance, being instantaneous, are perceived, but not motion. This conclusion comes near to reducing the memory theory to an absurdity.

A simpler assumption would be that motion corresponds to a variable of retinal stimulation. The variable is a correlation of the adjacent and the successive order of excited elements. This assumption raises no questions as to whether the impression of motion is based on memory, inference, or the interpretation of eye-muscle sensations. We need not assume any short-circuiting in the brain such as Wertheimer proposed to explain strobos-

[2] Discussions of film-montage, such as that of Eisenstein (31), have been concerned with the dramatic effects of successive shots rather than the effect of a coherent world in time. The former, however, surely depends on the latter, and the simpler problem should be studied first.

copic movement and the phi-phenomenon (118), and we need not decide, in an outmoded terminology, whether motion is a sensation or a perception. We only suppose that it has a stimulus-correlate.

This assumption is consistent with the evidence that motion is a simple quality of experience and it has been implicit throughout the last two chapters. If it is valid for the movement of a patch of color across the retina, however, why should it not also be valid for the continuous deformation of the retinal image? If it is, we are committed to the position that succession is just as much a property of visual stimulation as distribution. Moreover, if successive order participates in and determines visual experience over a short interval of time, why should it not do so for a long interval of time?

Considerations such as these point to a general theory of space perception which has interesting implications. Let us first define the physical environment in which a man lives as the one in which he gets about. Depending on the criterion of mobility selected, this might be his house, his neighborhood, or his city. At any given physical point in such an environment there is one and only one ocular image which a standard human eye will produce when it is pointed in a given direction. This image is unique. (Since the principle holds for either eye we need not refer to the second image.) If the eye rotates at that point in its peculiar saccadic fashion the images are individually unique, and the 360° panorama of images is a unique collection. Note that this fact is not psychological but physical; the flux of light is unique at that point when it is focused as an image. An eye can explore the flux of light at a given position like a blind man feeling an object on different sides in succession, and the panoramic image is just as immovable as a fixed object. The complex order of steps, contours, and gradients of this potential 360° image is unchanging, and the momentary images merely sample it.

If now we suppose that the eye moves from one point to another in the environment (as its possessor goes about his business) the image becomes a continuous serial transformation which, as a series, is unique to the path travelled. The series has no real beginning and no end during waking hours; any momentary cross section is specific to the momentary position and, over a long period of time, the serial image will have sampled the light flux at a great many positions. Assume next that the panoramic series and the locomotor series are combined, as they must be if an observer both scans his environment and moves about. The combination yields a range of images in two dimensions which corresponds to the whole of a three-dimensional environment, independent of any given point of view. It will not only be stable, panoramic, and unbounded, but it will approximate the range of images possessed by another man who lives in that environment in the degree to which they have been in the same places. It will be something very much like objective space.

An image of this sort is extended in two dimensions with respect to distribution and is extended in a third dimension with respect to sequence or time. It might be called a train of momentary images, but we

must remember that the panoramic series (related by overlap) and the locomotor series (related by continuous transformation) must be combined in the same train. This image-in-time is a conception so abstract and so strange that to call it a visual stimulus may strike the reader as absurd. A boundless, concurrent, and public world is not an absurdity, however, although it is abstract, and if it is not a pure illusion there must be a basis for apprehending it. To the abstraction on the side of experience and behavior there must correspond an equivalent abstraction on the side of physical stimulation. Tracing the latter may be difficult but the attempt, at least, does not involve us in an absurdity.

9

The Constant Sizes and Shapes of Things

Why Do Things Look as they Do? . . . . The Constancy of Perceived Objects with Respect to Color . . . . The Constancy of Perceived Objects with Respect to Shape . . . . The Perception of Foreshortened Surfaces . . . . The Constancy of Perceived Objects with Respect to Size . . . . How is the Distance of an Object Seen? . . . . The Perception of Distance and of Scale . . . . The Rigidity of Visual Dimensions . . . . Does Size Constancy Break Down at Great Distances? . . . . Conclusion: The Objectivity of Experience

For a long time psychologists have been puzzled about what has been called the approximate constancy of visual things. Objects look much the same size whatever their distance from the observer, and the same shape at different angles of regard, or from different points of view. It is true, of course, that when you attend strictly to the appearance of objects the rule does not hold, but then, we have argued, you are having a different kind of visual experience. Will a psychophysical theory of perception throw any new light on this problem? The traditional explanation for this constancy of objects in ordinary uncritical perception has been that we correct our sensations of the size and shape of things by remembering their true size and shape. Even apart from the objection to original sensations there are indications that this explanation is not the true one. Birds, for example, can discriminate correctly between a large and a small object when the larger object is so much farther away that its image is smaller. Does this imply the memorizing by the bird of the true sizes of all the objects it sees? Animals in general do not behave as if what they saw was at first a flat visual field and subsequently a three dimensional visual world. The behavior of animals and birds suggests that they react differently from the outset to objects at different distances. We may suspect that the so-called problem of constancy is actually only one aspect of our larger problem — the problem of the perception of the visual world with all of its objective characteristics. The aim of this chapter is ultimately to show that the question of why things retain their

sizes and shapes under different circumstances is a false question. Only if it is believed that perceptions begin as patches of color in a visual field does the question arise.

Why do Things Look as they Do?

To what extent can we now account for the perception of a phenomenal world which is adequate for behavior? This is the question which we set out to answer in Chapter 2. The visual world, it will be remembered, differs from the visual field in a number of ways. First, it has depth or distance, and it includes the experience of solid objects which lie behind one another. Second, it is Euclidean in the sense that neither the objects nor the spaces between them appear to change their dimensions in perception when the observer moves about. This is a general way of saying that they tend to remain constant. Third, it is stable and upright; things as seen have constant directions-from-here when the observer moves his eyes and the perceived ground remains horizontal when the observer tilts his head. Fourth, it is unbounded; our experience of the world does not have any visible margins or limits such as the visual field or a picture has. Finally, it has a characteristic to which we have scarcely referred but which, in a way, is the most important of all: it is composed of phenomenal things which have meaning. Even if the visual world is primarily an array of spaces, surfaces, and contours, it is secondarily an array of familiar objects, persons, and symbols. We have been trying to account for the perception of the material world, but objects have significance as well as solidity, and any theory of space perception must at least recognize this fact.

Of these five general properties, which can we now account for with a stimulus-theory of perception? We may begin at the end of the list and work backward. The meaning of things, surely, cannot be explained solely by their optical stimuli. Nearly all, if not all, meanings are learned by experience and therefore depend upon memory. So much do they vary from one individual to the next and from one culture to another that we are accustomed to say that different people do not even perceive the same world. We usually realize, however, that such a statement is an exaggeration. The world we refer to is the world of values and meanings rather than the world of space and shapes. The purest examples of such meanings are incorporated in those visual objects which are used as symbols — printed letters and words. Although a Chinese character will be seen as the same shape, at least approximately, by all men, it has meaning only to someone who has learned it. We shall return to this problem in Chapter 11. Before we can investigate why things have meaning, however, we must know why they are seen as material things. We assume, for the present, that the solidity and separateness of things must develop in the vision of the infant before meanings can begin to be attached to them. It seems probable that only as the child can identify things by shape, size, and color does he learn their significance for his needs or their use as conventional signs.

The unbounded or panoramic appearance of the material world, according to the suggestion made in the last chapter, is

one which might have a basis in retinal stimulation. This possibility, however, depends on the assertion that time is a variable of the stimulus or, more accurately, that sequence and arrangement are coordinate dimensions of stimulus variation. This is a plausible but abstruse assertion, stimulating to the imagination but not immediately productive of clear-cut experiments to prove it. The alternative hypothesis is to rely on the more familiar concept[1] of memory. The explanation would then be that successive fixations or glimpses of the environment are integrated in primary memory to yield a panoramic visual world. Stimulation, conceived as momentary, fails to account for the panoramic world.

We have seen that the stable and upright character of the world can be successfully explained by stimulation if we take note of the body as well as the eyes. One may assume concomitant stimuli which are reciprocal to one another. The shift of the retinal image and the movement of the eye to its new fixation are then in opposite correspondence. Likewise the tilting of the retinal image and the deviation of the body or head from its normal posture are correlated but opposite. In both cases, the visual and postural stimuli would yield a joint variable of complex stimulation which would normally be in correspondence with an unchanged scene.

The distance, depth, and solidity of the visual world have been considered in detail. They can be explained by stimulation if the world is conceived as an array of surfaces. Continuous distance is in correspondence with certain gradients in the retinal image of a surface or the combined images of two eyes. Depth at the contour of an object which stands out from its background is in correspondence with an abrupt step in these gradients. Solidity, or depth-shape, is in correspondence with the slopes of these gradients and their variations in slope. If these types of psychophysical correspondence are upheld in future experiments, the gradient theory can be said to explain the tridimensional properties of visual perception.

There remains one more item on our list, the second. The visual world might be described as *rigid.* A given object tends to remain constant in size as the observer moves toward or away from it, and constant in shape as the observer moves around it. Perceptual objects, instead of being deformed as the color-patches of the visual field are when the observer changes his viewing position, remain approximately the same, dimension for dimension. This constancy of size and shape also appears to hold true for the ground or the floor, and for any segment or part of the background. We shall find some evidence that it also holds for the distances between objects — the shapes of the intervening spaces — even though the corresponding color-patches in

[1] The fact that physicists have found it useful for certain problems in astronomy and in the structure of matter to assume that time is a dimension of space — the fourth dimension — is suggestive for psychologists but is apt to lead into high-sounding guesses. There may be a significant parallel between problems of relativity physics and problems of visual stimulation, but until the latter have received a mathematical formulation the tracing of the parallel will be speculative.

the visual field be shrunken, expanded, or transformed. This tendency for visual dimensions to remain constant is a fact of perception. It is also true, of course, that measured dimensions remain constant, but this is a fact of geometry and engineering, not of perception. Perceptual estimates are less accurate than physical measurements, but both are facts. The axioms of Euclidean geometry, we shall propose, are abstractions from both of these facts. When it is asserted that the world we tend to perceive is a three-dimensional Euclidean space, the assertion really means that the dimensions of things and of their interspaces tend to remain constant in perception.

The constancy of visual objects with respect to size and shape needs to be considered in detail, and it will be the main subject of this chapter. By implication, however, the problem is not limited to objects in the literal sense but applies to the whole of a visual scene, to the ground on which objects normally rest, and even to the dimensions of abstract space. If our approach is correct, the problem is a corrolary of a much more general question: How do we perceive a world which is consistent with our actual behavior – the visual world of ordinary experience?

The Constancy of Perceived Objects with Respect to Color

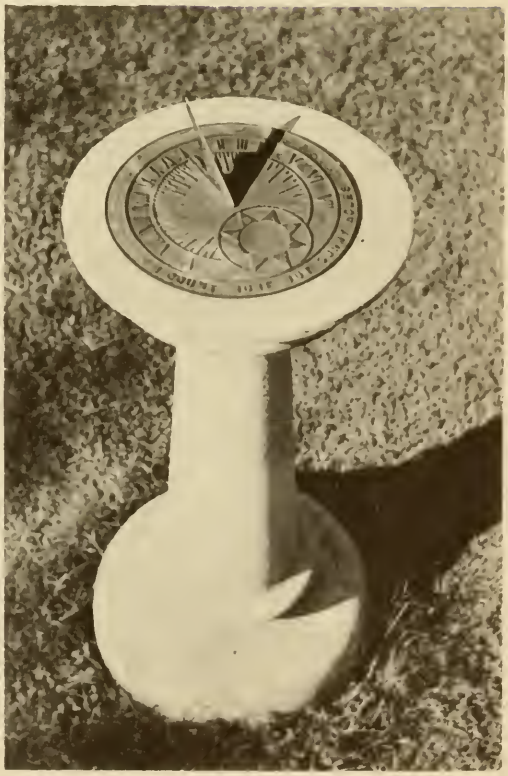
As a preliminary to the main problem, we might consider briefly the nature of color perception. Although colors are traditionally supposed to be sensations, and although the correspondence of hue, brightness and saturation to the variables of light energy is perhaps the most basic fact of the science of vision, the sensory colors can be seen only under the artificially controlled conditions of the laboratory or by the practiced eye of the painter. Such are the hues seen in a spectroscope or the color transmitted through a pane of ground glass in a dark room: they are disembodied colors floating in a visual field rather than the colors of objective surfaces in a visual world. They look filmy and insubstantial and appear at an indefinite distance, in contrast with the colors of objects in daylight illumination which appear to be localized on and to be part of the surface of the object in question. The latter are known as surface-colors, whereas the former have been variously called film-colors, expanse-colors, reduced colors, and the like (61). Only when color is thus disembodied or separated from a localized surface does its brightness correspond with the intensity of the light on the retina.

The colors of objects in the ordinary sunlit visual world are not the same as the colors of the patchwork in the corresponding visual field. The untrained observer cannot see these differences, but a landscape painter can, for he has had to learn that the disembodied color of an object is the color which must be reproduced on his canvas. Just as a table top cannot be represented as a square, so a white surface in shadow cannot be represented as white. In pictorial vision the former is a trapezoid and the latter is gray. The surface-colors of everyday perception remain fairly constant despite changes in the illumination of different

FIGURE 67. Brightness Constancy

Fritz Henle, from Monkmeyer

Camera Guild



parts of the world and despite changes in the illumination of the whole world between sunrise and sunset. It is as if the colors of objects in perception corresponded to the chemical and physical properties of their physical surfaces independently of how well these are lighted, and hence in seeming independence of the light-intensity of the retinal image. Perception seems to deviate from its stimulus and, in the words of Thouless (106), to show "regression toward the real object." This is, of course, a valuable achievement for the perceiving individual insofar as it helps to identify objects at dusk or in deep shadow.

A variety of theories (79, 67) has been proposed to explain this tendency in perception, based on the assumption that the disembodied color is first aroused in the retina and that the object color is then perceived by some added process. This is not, however, the only possible assumption. One might suppose that a disembodied color corresponds to light stimulation *without* ordinal stimulation and that a surface color corresponds to light stimulation *with* ordinal stimuli for a determinate visual world. Both types of correspondence would be strict, but the latter would involve more variables and would be expressed in a more complex function. As an example take a protuberance which, as we argued in Chapter 6, may be considered an elementary depth-shape. Although both sides are physically white, the left side is lighted and the right side is shadowed. Why does it appear to have the same color on both surfaces? Possibly because the high-to-low step in the brightness of the retinal image yields an impression of depth and therefore cannot at the same time yield an impression of a difference in color between the adjacent surfaces. The bright-dark stimulation on the retina could yield the disembodied colors white-gray, but in that event the perception would be depthless. (So it would actually be if the protuberance were looked at through a tube which hid its outer margins.) The alternative is for the bright-dark stimulation on the retina (in combination with texture and binocular disparity) to yield a protuberance, but in that event the perception is of both sides with the same color.

The conception of two kinds of seeing, the natural kind and the introspective kind, is confirmed by some of the experimental results on brightness-constancy. Both MacLeod (79) and Henneman (54) have demonstrated that two different attitudes are possible for an observer in such an experiment, the objective attitude and the subjective attitude. The former is naïve and is directed toward the real object. The latter is critical, analytic, or photographic, and is directed toward the stimulus. The two attitudes are not mutually exclusive, for there may exist intermediate stages. MacLeod found that "a shift in the observer from the objective to the subjective attitude is sufficient to reduce considerably or even to destroy completely the phenomenon of color-constancy" (79, p. 45). Henneman demonstrated that the objective attitude yielded a high degree of brightness constancy (62 per cent) while the subjective attitude yielded a great reduction in constancy (25 per cent). The writer's interpretation of these facts would be that in the first

case the observer tended to see a visual world whereas in the second case he tended to see a visual field.

The problem of color-constancy is much broader than the problem of perceiving the whiteness of the shaded side of an object. It includes the perception of the whole field of colored surfaces seen at different levels of illumination or in artificial or colored illumination, and the perception of surface hues as well as black-white qualities. The experimental evidence is voluminous and there are difficult questions such as how the general level of illumination over the visual world is sensed by the observer. An introduction to the subject is provided by MacLeod (79), Katz (61), and Koffka (67).

The implication of this digression into color-constancy is intended to be this: that even colors – the supposedly pure qualities out of which vision is built – tend to be intimately connected with surfaces, slants, and edges. Color as it is embodied in space is affected by spatial stimulation even though it is true that spots and grades of color provide the basis for spatial stimulation. The innate attribute of extensity which color has been supposed to possess turns out to be not the simplest kind of space but merely indeterminate space.

The Constancy of Perceived Objects with Respect to Shape

Is the reason for our seeing an object as possessed with constant dimensions from whatever position we view it simply that we know the object? Assume that we have handled it and perhaps even measured it or, if not, that we know the laws of physics about all material objects, including unfamiliar ones. Assume that what we know modifies what we see. These familiar assumptions will explain a great many of our perceptions, including an object with constant dimensions, but whether it is necessary to call upon so intellectual a process to explain this particular kind of perception is questionable. Conceivably the rigid object has a correlate in retinal stimulation even though it certainly does not have a copy.

The retinal image as a whole and in every part undergoes a continuous transformation as the observer moves about. The images of objects, moreover, are deformed when the objects move with reference to the observer (page 34). As you walk up to a mailbox to post a letter, for instance, the projected shape of the object goes through the series of transformations shown at the top of Figure 68, and when you face a door that is being opened, its image is transformed as in the drawings just below. Nevertheless, the mailbox and the door retain the same shape so long as you adopt a natural attitude toward them or, in our terminology, so long as you see them as part of the visual world. On separate occasions when you see them at novel angles of regard, they still possess their proper shape in three dimensions. Is this entirely a matter of knowing the mailbox and knowing the door?

An answer can be given only by experiment and only by simplifying the situation. As described above, the conditions are complex. Constancy experiments in the psychological laboratory, therefore, ab-

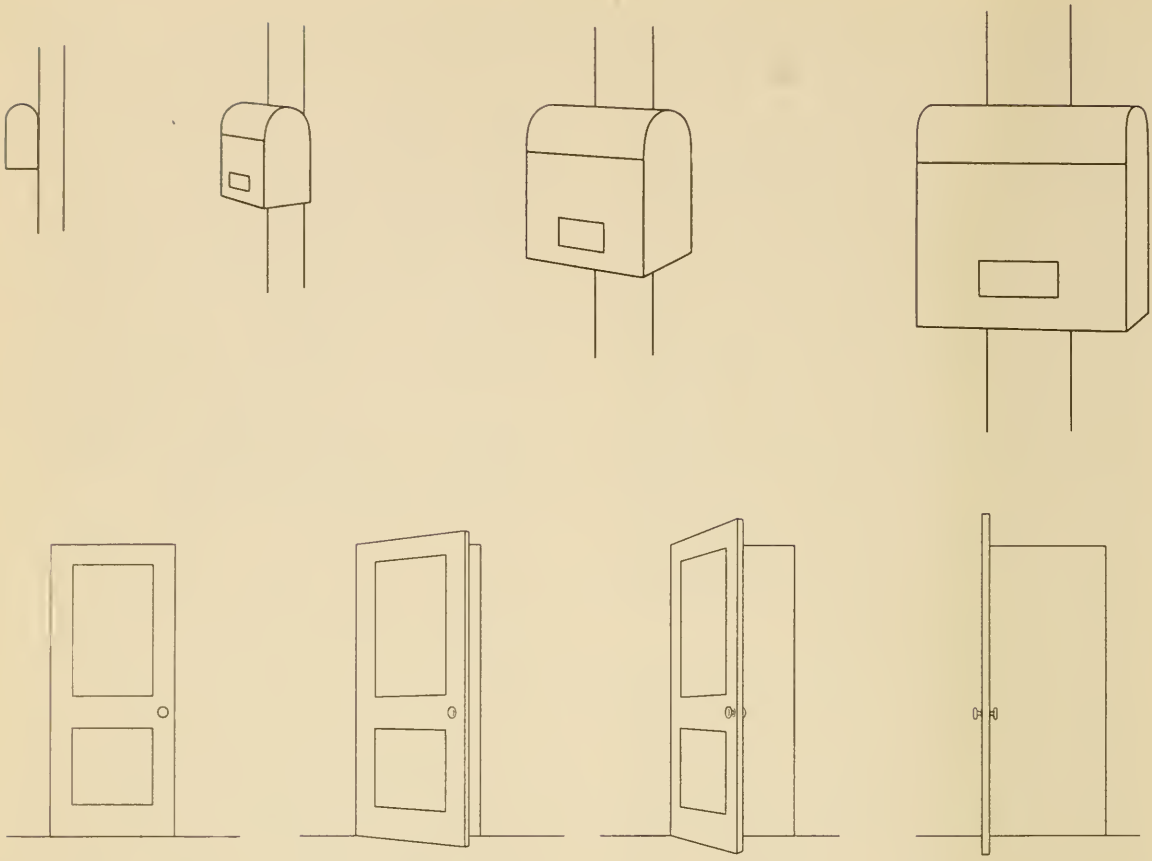


FIGURE 68. Constancy of Shape

stract from the perception of mailboxes and doors in several ways. First, the object whose shape is to be judged is isolated from the background of surfaces in an ordinary environment and is placed on an artificial background. The surfaces surrounding the object are often flat cardboard screens, or something equivalent, and frequently the environment is darkened. Neither the object nor the observer is allowed to move. Second, the object to be judged is itself artificial. The chosen shape is arbitrary and therefore unfamiliar to the observer. A typical object would be a rectangle or an ellipse cut out of cardboard. Third, the shape to be judged is reduced to the flat outline of a single surface which is then slanted with respect to the line of sight. What the observer sees is not a true solid object but an isolated face of an object. The observer has to judge the particular dimension of this outline which, being slanted, is optically compressed – for example the altitude of the rectangle or ellipse. A scale of graded rectangles or ellipses is provided for the judgment. Fourth, the outline to be judged is changed by the

experimenter on successive trials so that the observer cannot become familiar with the object, and the slant of its surface is altered in a random manner from one trial to the next.[2] The conditions of perception are simplified and reduced as compared with ordinary vision. Nevertheless, the significant result of all these experiments is that the observer can fairly successfully match the slanted object with an objectively equal but unslanted object, so long as he takes a naive attitude toward what he sees and so long as cues for the depth of the object are visible.

The most obvious implication of this result is that we do not have to be familiar with an object in order to see the same shape at different angles of view. Knowledge and past experience of the object in question are not essential for constancy. The constancy of its dimensions must depend, instead, on our ability to see it in three dimensions. For the particular object used in the experiments – the flat face of a solid object – this means our ability to see the slant or tilt of the surface.

There is another result of these experiments which has seemed even more significant and which has therefore received more emphasis – the fact that constancy is often incomplete. In general, if the observer takes a critical attitude toward what he sees, or if he is asked to judge the apparent shape of the object, his judgment becomes a compromise between the objective shape and the shape as it would be projected on a picture-plane. If he is asked to pay attention to the projected shape his judgment is still more compressed, but it seldom or never reaches the true projected shape. If the cues for depth perception are absent the judgment may approach the perspective shape. These compromise judgments are usually taken to be the typical results of the experiments, and they are expressed by computing an index of the amount of constancy, or the degree of phenomenal regression toward the real object (106). The index is zero for a perspective judgment and 100 for an objective judgment. Is it not true, however, that this index of constancy begs the question? Using such an index suggests that one is measuring a perceptual process and that this is superimposed on a primary sensation of shape. The process is assumed to be one which, figuratively speaking, corrects the sensation. It implies that the visual field is the sensory basis for experience, while the visual world is a perceptual accomplishment of the organism.

If one questions this traditional explanation of the experimental results, what theory is better? The basis for an adequate theory has been laid by Koffka (67, p. 228 ff.). Any perception of the stimulus object involves two components, the shape and the orientation. These two aspects of the percept are, as he says, coupled together. The shape is not experienced in isolation; it is always a shape-in-a-given-orientation. We can suppose that the perceived orientation combined with the apparent shape yields a constant shape. If the orientation is

[2] These conditions are fairly typical for experiments on constancy of shape. The most recent is that of Stavrianos (100), whose report summarizes the earlier work.

seen correctly, the constancy will be complete; if the slant is not visible, there will be no constancy.

An effort to verify this theory has recently been made by Stavrianos (100), using tilted cardboard rectangles. The idea was simply to determine whether observers who failed to perceive the tilt correctly also failed to perceive the tilted dimension correctly by a geometrically equivalent amount. It could be predicted that if an observer judged the degree of tilt as less than it was, he should then judge the rectangle as just so much shorter than it was. This determination proved to be not as easy as it might seem, and the prediction was not perfectly confirmed. The trend of the results, however, especially in one experiment, was consistent with it.

The Perception of Foreshortened Surfaces

In formulating the problem of shape constancy the tendency has been to think of shape as a kind of disembodied geometrical form and to think of depth as a kind of disembodied third dimension. The kind of shape which manifests constancy, however, is an outline attached to a surface, and the kind of depth which is relevant to constancy is the slant of a surface. The perception of a surface is the central problem, then, if we want to understand the seeing of a shape in depth. Experimenters have simplified the general phenomenon of shape-constancy so as to deal with the outline of one flat face of an object, but they have not thereby eliminated the surface in which both the shape and the impression of slant are embodied.

Figure 69 illustrates the optical compression which the square face of an object would undergo if it were slanted backward from a frontal position.[3] The outline is said to be foreshortened, that is, it is shortened along the fore-and-aft dimension but not along the dimension which serves as the axis of rotation. The texture of this surface is represented as a checkerboard, the size of the units of texture being much exaggerated, in order to illustrate that the same optical compression which affects the outline also affects the texture. How is the slant of such an outlined surface sensed by an observer, and what is the relation of the slant to the foreshortening of the outline? One basis for slant is provided by the principle stated in Chapter 6, that the retinal gradient of density in the image of a physical surface bears a constant relation to the slant of that surface. The steepness of the gradient is proportional to the degree of slant, and the direction in which the density increases is related to the direction toward which the surface faces. The principle was intended to apply to any variety of visual texture – from the finest microstructure to a gross structure of inlines or repeated patterns. The steepness of a texture gradient in a certain direction, then, is a potential stimulus correlate for perceived slant. The steeper a gradient of density along a projection, the greater will be the compression of the texture, as one can note in Figure 69.

---

[3] The slant of a surface may be defined with reference to a line of regard and a surface perpendicular to it. Slant is any departure from this plane by rotation on a horizontal or a vertical axis.

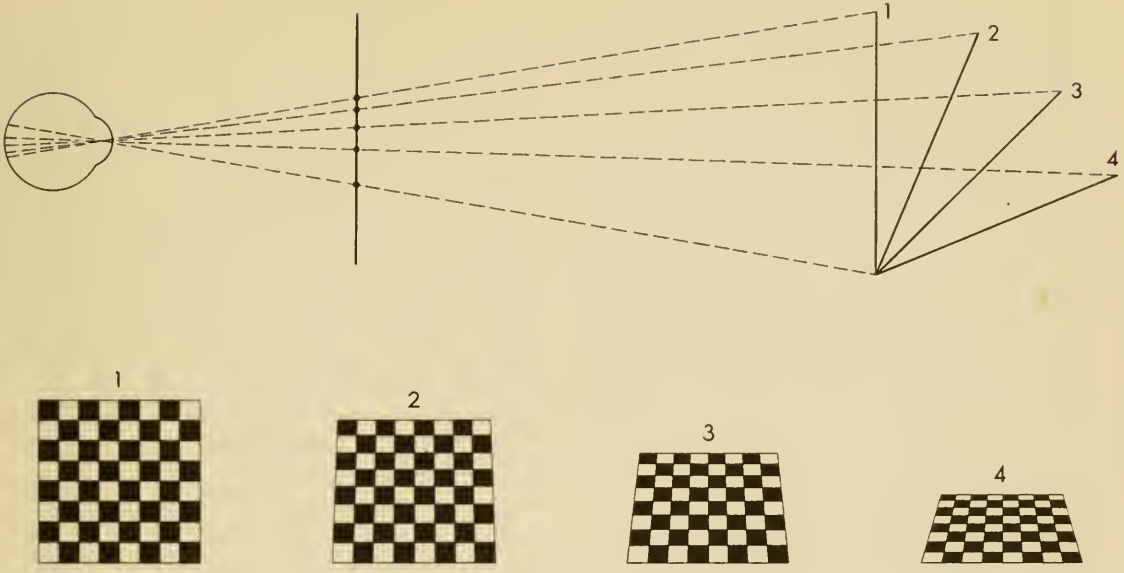


FIGURE 69. One-way Compression of Texture and the Foreshortening of a Contour

A slanted surface manifests a one-way compression of texture — an increase in mean density along one dimension of the projected image relative to the other. The texture — the actual surface quality — of the face of an object is foreshortened when the outline is foreshortened. Although the compression of outline is much easier to observe than the compression of texture, especially in the case of surfaces with a fine microstructure, the latter is also important for the perception of slant. If the outline of the tilted object in Figure 69 were an ellipse of unknown dimensions instead of a square, the one-way compression of texture would still yield a perception of slant although the foreshortening of outline would be indeterminate. The objective dimensions of the ellipse could then be perceived, and this result is actually what occurs when the experiment is tried.

Other gradients than that of texture-density are, of course, supplementary stimuli for the perception of slant on a delimited surface, particularly the gradients of binocular disparity and of motion. The cues of disparity and motion are supposed to yield depth rather than slant, but these two concepts have been shown to be fundamentally inseparable, and the rule relating the gradients to the slant of a surface can be stated. The steepness of a gradient toward crossed disparity running from the far to the near margin of a physical surface is proportional to the slant of that surface (Chapter 6, Fig. 49). Likewise the steepness of the retinal gradient of motion over a physical surface bears a constant relation to the slant

of that surface. In general, the greater the slant of a physical surface the steeper will be all these retinal gradients, and the steeper they are, we may postulate, the greater will be the impression of slant.

These two gradients of disparity and of motion, however, are not primary ordinal stimuli. The first is dependent on different simultaneous projections registered in two eyes. The second is dependent on different successive projections registered in a single eye. One is a gradient of the relative skewness of the two binocular images and the other a gradient of the deformation of the image. Both presuppose the existence of a retinal image possessing an outline and usually possessing texture. Both gradients can be eliminated in a shape constancy experiment if the observer looks with one eye and holds his head motionless since this leaves only the bare outline and the texture as the stimuli for slant. When they are thus eliminated the tendency toward shape constancy should still be evident if the density of texture is visible.

The tendency to perceive the dimensions of things as constant, therefore, can be ultimately reduced in the shape constancy experiments to this question: how can we judge a height relative to a width when one of these dimensions is foreshortened on the retina? Figure 69 suggests the following kind of answer: a dimension is judged as the amount of surface perceived, and a surface is composed of texture-elements. The gross number of elements, or amount of surface perceived, along the compressed dimension is just the same as the number or amount along the uncompressed dimension when the two are equal. This is true because the elements are compressed along with the dimension. The amount of surface perceived, therefore, will be in proportion to the dimension of the surface and not to the dimension of the image.

If you ask, "Why is the retinal compression not seen?" the answer is that it *can* be seen with special attention and a special effort to disregard the slant. (Some impression of slant, however, is normally compelled and the observer therefore cannot see all the compression.) Without the special effort, however, the retinal compression yields an experience of slant, not of compression. This quality of slant is a correlate, not a copy of its stimulus. When the conditions of the experiment are such that the texture of the surface becomes indeterminate (in a darkened room, for instance) and when the gradients of disparity and motion are also eliminated, a dim outline on the background may come to be the only remaining variable in the retinal image, and in that event the slant should become zero. The observer no longer sees the dimensions of a surface, but the proportions of a pure depthless shape — an abstract shape presumed to lie in the frontal plane.

The Constancy of Perceived Objects with Respect to Size

The size of the retinal image of an object is a very poor indicator of the size of the object, just as the shape of the retinal image of an object is a poor indicator of the objective shape (15). So far as the light which composes the image is

concerned, the object might be either something small and near or something large and far off. The cone of light rays which enters the eye and is focused as an image does not differ substantially whether it comes from five yards or five hundred yards. Figure 70 illustrates this principle. The photographs show two charts for measuring visual acuity, one chart being enlarged to four times the size of the other and placed at four times the distance of the other. They subtend exactly the same angle on the retina just as they do on the photographic film, and this fact may be noted when the two charts are projected side by side. The letters on one chart, consisting of E's in various positions, are just as distinguishable as the letters on the other chart, for the two charts are optically equivalent.

In the right-hand scene, when the observer puts his head in exactly the position to align their two sides, the charts come together and appear to be of identical size and to be at some paradoxical but identical distance. The two surfaces actually seem to be continuous with one another. In the left-hand scene, however, the chart on the right appears to be much larger than the one on the left and at a much greater distance, although the legibility of its letters remains the same. When it goes back into the distance it looks automatically and inescapably larger. The photographs illustrate, therefore, both the law of the visual angle for the retinal image and the tendency toward size constancy for the corresponding perception.

By way of digression, Figure 70 also illustrates two degrees of steepness of the gradient of linear perspective. The more obtuse the angle at which the lines of a retinal image converge to their vanishing point, the steeper is the gradient of convergence. The steeper the gradient of convergence over a surface, according to our hypothesis, the greater will be the impression of slant, that is, the more it will look inclined to the line of sight. Observe the floor of the corridor in the two photographs and compare the angles made by the edges. In the second scene one appears to be looking down at the floor, relatively, whereas in the first scene one appears to be looking more nearly along the floor. These gradients, in combination with others on other surfaces, establish a point of view in each scene, one being approximately from a standing posture and the other from a kneeling posture.

The non-variation in the perceived size of things in spite of variations in their distance has frequently been studied under laboratory conditions. The experiments on the problem generally involve the setting up of an unfamiliar object such as a stick or a cardboard square at a considerable distance from the observer and a series of varying comparison objects at a conveniently close distance. The observer's task is to judge the size of the far object in terms of the near objects. With a naive attitude, and under favorable conditions for seeing the distance the estimates are fairly accurate. Constancy is then said to be complete. This accuracy, however, is only possible if there are adequate cues to the distance of the far object. When the cues are reduced or eli-

PICTURE 1

FIGURE 70. Two Objects with the Same Visual Angle

minated from the scene, correct judgments of size no longer occur (57).

These results suggest that there is no such thing as an impression of size apart from an impression of distance. In the case of an unfamiliar object, its size is necessarily a size-at-a-given-distance. One ought really to speak of size-distance perception, for the two are "linked together," in Koffka's words, both optically and perceptually. The question is, how is the impression of distance obtained? Since it is not produced by the cone of light rays which constitutes the image of the object, it must be produced by the light rays which constitute the image of the background of the object.

How is the Distance of an Object Seen?

Figure 71 illustrates the two definitions of the problem implied by the theory of cues and the theory of gradients. In the upper drawing the near object projects a retinal image twice the size of the far object. How can they be seen equal? The different cues for the distance of each object must correct or compensate for the different retinal sizes. It is easy to under-

PICTURE 2

stand why cues like accommodation, convergence of the eyes, binocular disparity, and superposition received all the emphasis in this theory, for the objects are represented as if suspended in empty space. How the empty space is seen gets no explanation unless one assumes that distance is computed in the brain by a mechanism similar to that of the optical range-finder – every act of perception being an exercise in trigonometry. In the lower drawing the same objects are represented as if resting on or attached to a surface. The spots of the surface project as a background for the objects; the gradients of density, deformation, and relative skew provide a continuum of distance in which the distance of each object is then fixed. The gradient theory undertakes first to explain how continuous distance is visible and second to account for the distance of objects.

Granting that a longitudinal background is necessary, how is the distance of an object fixed on the background? Let us assume that an object is seen where its contour interrupts the background – at that distance and no other – except when depth-at-a-contour brings it forward in distance. This latter effect is produced mainly by a step in the rate of deformation or disparity at the contour. We are assuming that in

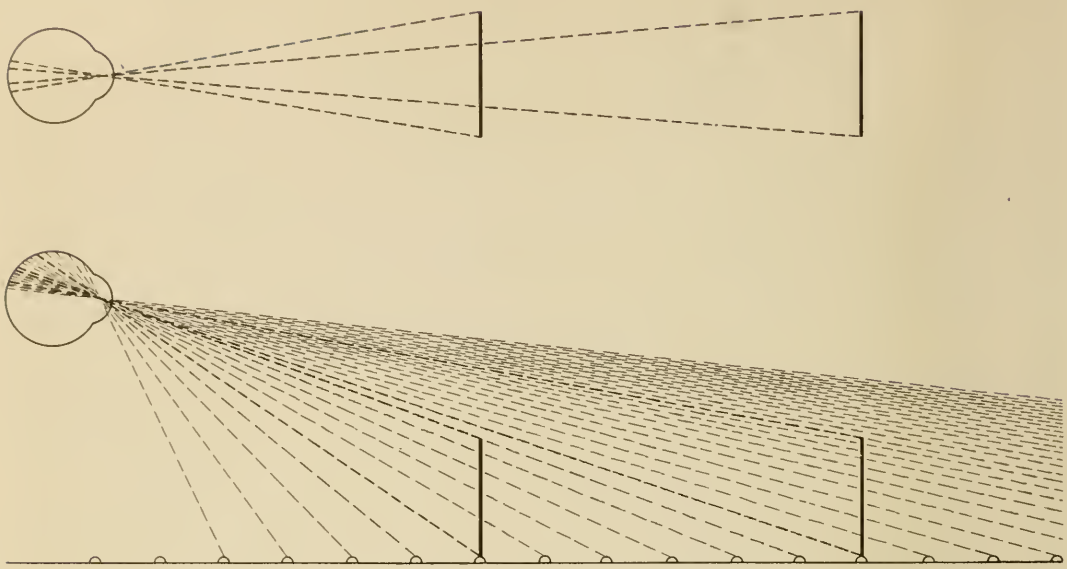


FIGURE 71. Two Objects of the Same Size but Different Visual Angles

the absence of what is called relative motion or stereoscopic depth a contour is seen on the background. This would explain why, in the monocular vision shown in Figure 71, the object appears where it physically is instead of nearer and smaller within its cone of rays.

One is free, of course, to assume instead that an object is seen resting on or attached to its background because we know from experience that objects are either at rest or physically attached to a surface. Although the effect of past experience no doubt contributes to the explanation, it is not a necessary assumption.

The former hypothesis can be tested experimentally. Figure 72 represents a set-up for doing so. The upper photograph is approximately what the observer saw in the experiment: a long table with a patterned cover on which two white oblong objects appeared to rest, one large and far on the left, and one smaller and nearer on the right. This impression is clear even in the photograph. The observer looked through a small hole in a large screen at the near end of the table, using one eye and keeping his head motionless. His angle of view was approximately as represented.

Actually, the object on the left is the smaller and nearer, as can be seen in the second photograph. It is simply a cardboard rectangle raised about three inches from the table-top by a rod which is invisible to the observer. Its distance is given by the optical contact of its base with the background. The dimensions of the rectangles are such as to project identical images at their respective dis-

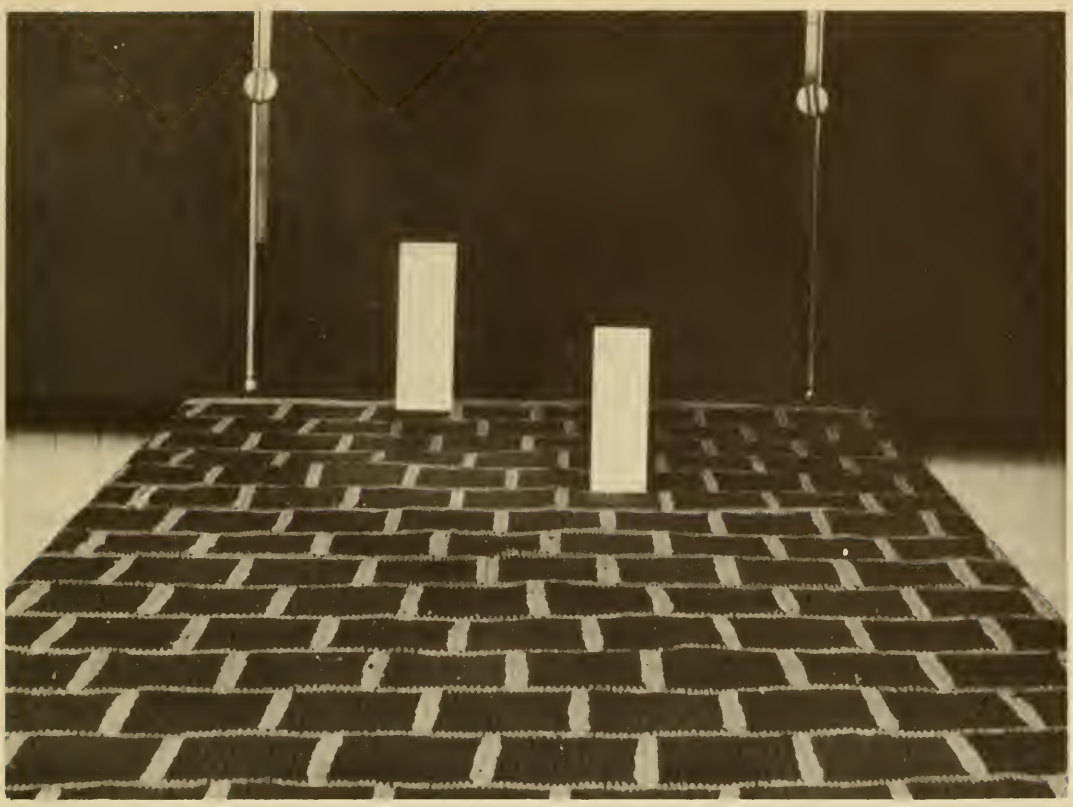
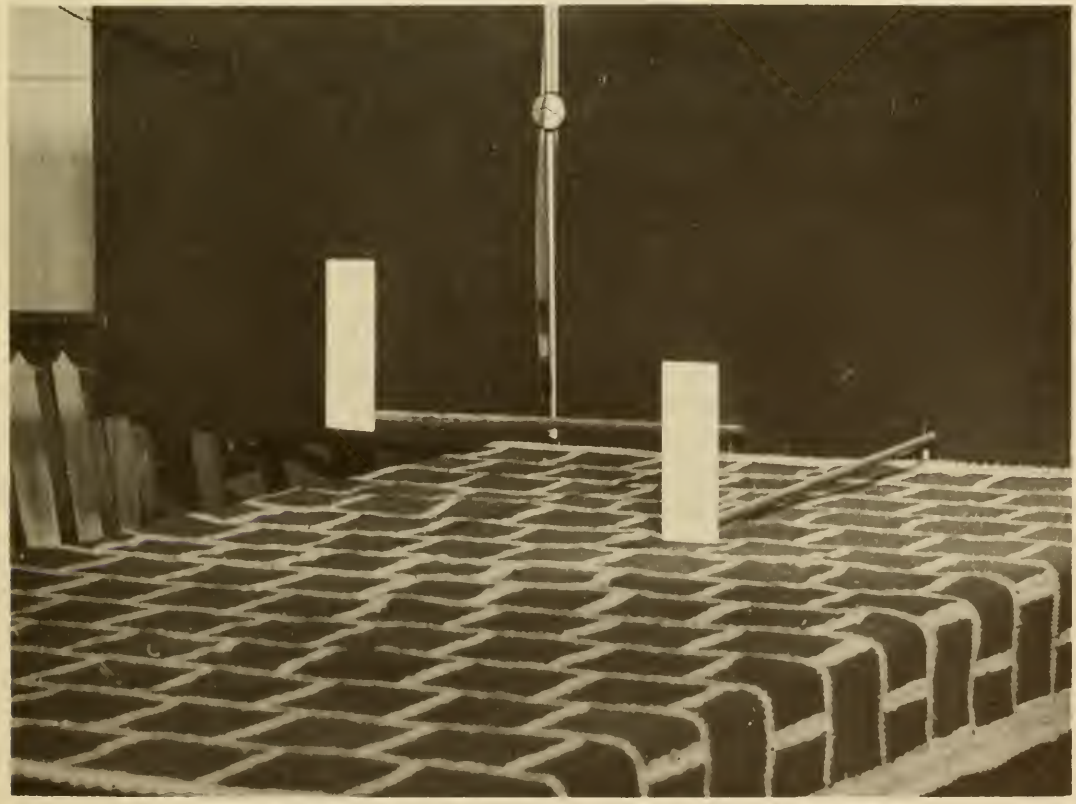


FIGURE 72. Distance as Dependent on Contact with the Background



tances, so that the two images in the upper picture are physically the same size.

To an observer unaware of the physical set-up the illusion is astonishing, but when he is permitted to look at the objects with both eyes and to move his head, he suddenly sees the left-hand object as much smaller and suspended in the air. Even with full knowledge, however, the original appearance returns, or strongly tends to do so, when monocular motionless vision is reemployed. The conclusion is that a perceived object under these conditions recedes within its cone of light rays until its surface is continuous with the background surface. Other things being equal, two optically adjacent surfaces tend to be adjacent in depth – a phenomenon which was also observed in the acuity charts of Figure 70.

Painters have always known that the higher up an outline is placed within the frame of a picture, the farther away it will appear, even when no background whatever is represented to give perspective. This is the recognized cue of relative upward location in the visual field. Of two objects in a perfectly blank frame, the upper will appear to be farther away. The explanation is probably that a blank background suggests a terrain or floor – the fundamental visual scene of Figure 19 – more strongly than it suggests a wall or a ceiling. A terrain or floor is always present to vision whereas a wall or ceiling may not be. "Upness" is therefore a fairly reliable cue to the distance of an object in the visual field, given the principle of contact with the background and granting the great frequency with which the ground is the background. When this cue is studied in isolation, as it would be in an empty frame, it should be regarded as an associated variable, an inference, or a probable indicator of distance, not as a true stimulus, since although the stimulus gradients on the ground produce an increase of distance upward in the visual field, the stimulus gradients on walls and ceilings do not (Figure 26, page 71).

The Perception of Distance and of Scale

The evidence of the size-constancy experiments is not entirely summed up in the conclusion that the size of an object tends to remain the same at different distances. Some of the experiments point to a more general statement; they imply that the dimensions of things, large or small, are *comparable* at different distances. This general formula comes closer to describing the ordinary experience of the world in which we live since, under favorable conditions, we can estimate a range of sizes at a range of distances. As an extreme case we might take a flier looking for an emergency landing field. He will estimate the sizes of all the cow-pastures for ten or fifteen miles in any direction including the one directly below, all being unfamiliar and no one being the same as another. These facts suggest that we perceive a quality in the visual world which might be called *scale*.

For objects at any given distance we possess a subjective scale of sizes from very large to very small such that we can judge immediately and rather accurately not only when two trees are equal in width but also whether one tree is half, or twice, or three or four times the width of another. Size perception, to describe it accurately,

involves a scale of sizes in relation to which any given size or dimension is unique. Such a scale is psychological; it is something we carry around with us and it is implicit in the very process of perception; it is therefore not to be confused with the conventional scale of meters or feet which depends on a set of operations using carefully constructed pieces of wood or metal. The conventional scale is a psychologically complex affair, depending as it does on the learning of techniques and concepts, whereas the implicit scale of visible size is a primitive feature of perception. [4]

The size constancy of objects, in the light of this conception, is a by-product of the constant scale of the visual world at different distances. Scale, not size, is actually what remains constant in perception. The gradient theory can account for this kind of constancy, for one can assume that the perceived scale of the background is a function of the same stimulus variables which yield the continuous distance of the background – a different function, it is true, but equally dependent on stimulation. The size of any particular object is given by the scale of the background *at the point to which it is attached*, and that is why its apparent size is linked to its apparent distance.

The impressions of scale and distance are so related to one another that with increasing distance there goes an unvarying scale. This is the rule for ordinary perception of the visual world. If the observer tries to see the world in perspective like a flat picture, however, the relation between them is the same but the impressions are different: with unvarying distance there goes a decreasing scale. The reason for the landscape-painter's ability to see perspective in his visual field is that he has retinal stimulation which ordinarily yields joint impressions of both distance and scale. When he makes an effort not to see the distance, the scale is correspondingly diminished, and this altered impression is what we call seeing the perspective.

Figure 73 illustrates what is sometimes called the perspective-illusion. All three cylinders are the same size on the page. It is not an illusion at all but a demonstration that apparent size depends on apparent distance. An illusion may be defined as a perceptual judgment or estimate which is consistently not in agreement with measurements of the object giving rise to the perception. This definition is clear enough for objects; it is when pictures come into consideration that our thinking about illusions is apt to become confused. Insofar as this picture is a substitute for objects (and our chronic habit is to see pictures thus) the increasing size of the cylinders is not illusory. Insofar as this picture is itself an object (consisting of black lines on white paper)

[4] The study of psychological scales, or subjective scales, is still in its infancy. Enough is known, however, to suggest that they permeate all forms of perception and that the traditional emphasis on perceived objects has too long diverted psychologists from concerning themselves with perceptual scales. A program of experiments on various perceptual scales (area, length, angular size, numerousness, and so on) is currently being conducted under the direction of Dr. John Volkmann at Mt. Holyoke College, but this research has not yet reached publication.

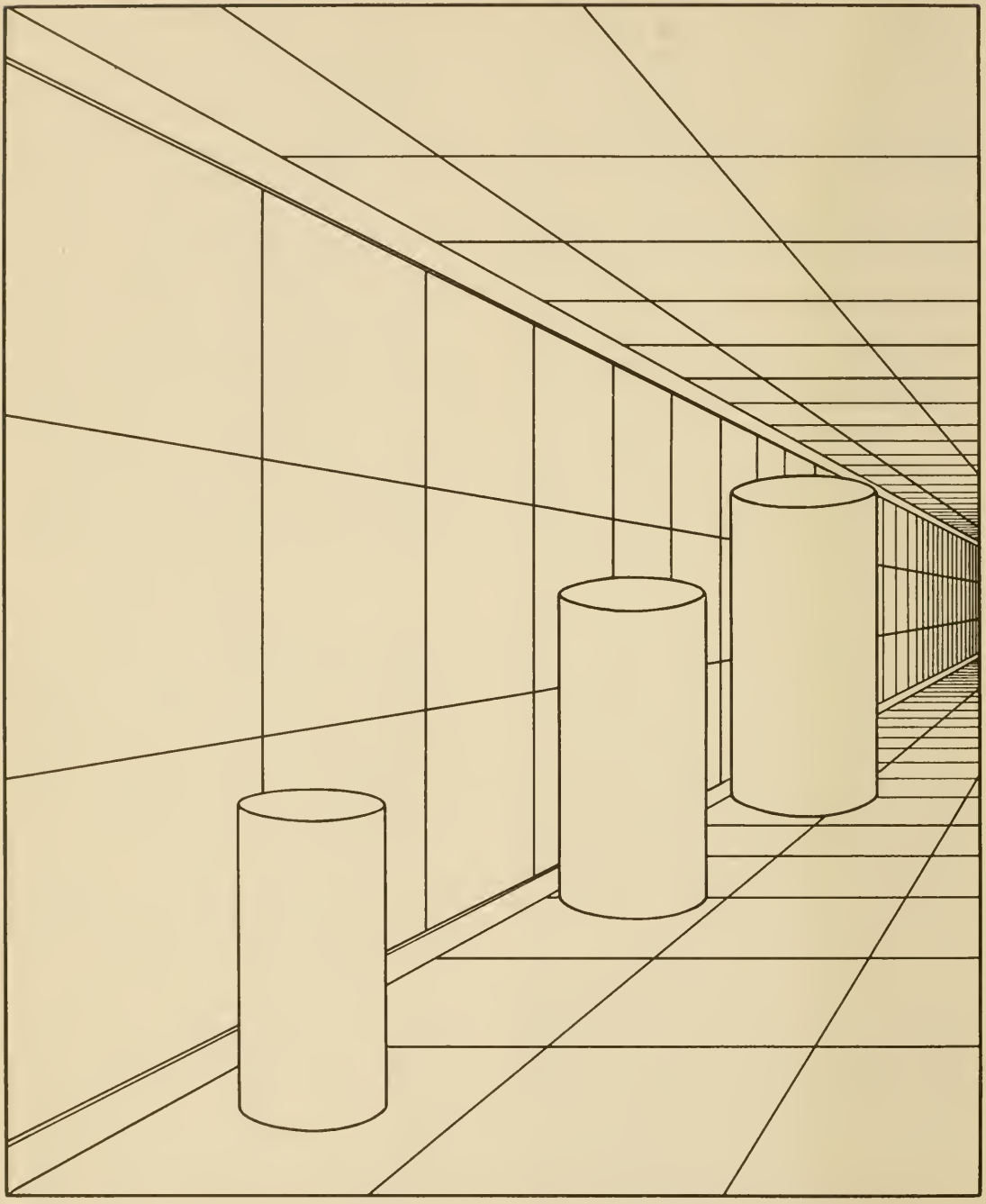


FIGURE 73. Size as Determined by Distance

the increasing size of the cylinders is illusory.

The Rigidity of Visual Dimensions

It was proposed at the beginning of this chapter that constancy tends to hold for the distances between objects as well as for objects. If this is true it would help to explain why we find it so easy to conceive of abstract geometrical space as we do. Although the tendency toward a constant size and shape of interspaces with varying distance and angle of regard is suggested by ordinary observation, the phenomenon has been neglected by experimenters.

During the war R.H. Henneman and the writer set up an exploratory experiment which, incomplete though it was, may be worth describing here. The observer was seated at the end of a thoroughly cluttered room containing tables, cabinets, boxes, shelves, and furniture. Among these objects he had to estimate 20 specified dimensions, some being the dimensions of solid things and some being dimensions in the open air between them. The distinction was not as clear as it sounds, for there was always a background surface behind any dimension. They varied between extremes of two and forty inches. The subject made his estimates by pulling out a steel measuring tape to match the specified dimension. It should be noted that the length, width or height was always optically diminished by distance, and might also be foreshortened, relative to the tape. The estimates of the fourteen subjects varied among themselves and from one dimension to another, as would be expected, but the over-all mean error was a slight overestimation. Size constancy was the rule. The significant result, however, was that the interspaces could be estimated without notable uncertainty. Dimensions in the air were judged with less accuracy and with a tendency toward underestimation as compared with dimensions of a solid object, but nevertheless with an approximate constancy of size. The implication is that the distances between things tend to be visually rigid as well as the things themselves.

Does Size Constancy Break Down at Great Distances?

One more wartime experiment is relevant to the kind of theory being developed. Although the perceived size of an object which recedes in the distance has been recognized not to diminish at the same rate as its retinal image, no one has ventured to suppose that it does not diminish at all. At some eventual distance the object ceases to be visible, and what is easier to suppose than that it does so by way of becoming smaller? It has therefore been assumed that size constancy necessarily breaks down at large distances, and perceived size then tends to become perspective size. The implication is that in outdoor space or aerial space, as contrasted with the room-sized spaces of the psychological laboratory, the appearance of things necessarily tapers off toward the horizon and the features of the terrain perforce look smaller than they are.

The experiment to be described provided a test of this assumption, for size-estimates were obtained out to a distance approaching bare visibility of the object. The situation is shown in Figure 74, which represents several trials of the experiment

FIGURE 74. Judgment of Size-at-a-Distance

FIGURE 74 (Contd.) The bottom photograph is the same scene as the one opposite it except that the camera is higher, "looking down".

(36, p. 201 ff.). From the observer to the low hills in the background the distance is a half-mile or more. The field is the sort of perfectly level cultivated land found in the coastal plains of southern California, and it was selected after much exploration. Its texture was fairly even, but without any furrows to give linear perspective. The procedure was as follows. A wooden stake, from fifteen to ninety-nine inches in height, was planted at a previously measured distance by one member of the experimental team who then hid in a conveniently invisible irrigation ditch. The observer, in complete ignorance of its size, faced about and estimated its height, either in scale-numbers, or by saying "smaller than one" or "greater than fifteen." These judgments were repeated for different sizes of the stake and at different distances, 150 judgments being obtained from each of fifteen observers. The averages of the estimates were then computed for the different sizes, distances, and observers.

Let us examine the results for the seventy-one inch stake at different distances, since they are typical. The correct match was with scale-number 12, which is also seventy-one inches high. The mean estimate when it was planted at the same distance as the scale (14 yards) was 71.9 inches (S.D. = 1.8) showing great accuracy. The mean estimate when it was planted at 224 yards, a three-minute walk down the field, was 75.8 inches (S.D. = 7.3). The mean estimate when it was planted at 784 yards, a ten-minute walk and nearly to the end of the field, was 74.9 inches (S.D. = 9.8). The four intermediate distances were comparable. This outcome is surprising, for at 784 yards – nearly a half-mile – a man-sized object is beginning to be difficult to make out.

Our question as to whether sizes necessarily become smaller in perception before they reach a vanishing point seems to be answered in the negative. Under favorable conditions for seeing distance, as we have here, an object can apparently be seen with approximately its true size *as long as it can be seen at all.* Its size does not become smaller but only more indeterminate. In this experiment there was no question of an index or amount of size constancy, for that was complete, but only of the increasing variability of the judgments with increasing distance.[5]

The critical reader may notice in Figure 74 that there are one or two ways of consciously inferring an approximate size for the distant stake. Inferences of this or a similar sort are always possible to some extent in judging objects on the ground. With the exception of one observer who was a professional psychologist, no one became aware of these cues during the experiment. What the average subject perceived was merely a size-at-a-distance.

Conclusion: The Objectivity of Experience

The general implication of this chapter, taken with the preceding ones, is this: the objectivity of our experience is not a paradox of philosophy but a fact of stimulation. We do not have to learn that

[5] See (39) p. 210, for other details of the experiment.

things are external, solid, stable, rigid, and spaced about the environment, for these qualities may be traced to retinal images or to reciprocal visual-postural processes.

Such a conclusion upsets completely the traditional interpretation of perception. The experiments of Gestalt psychology undermined it, but they did not overthrow the conviction that somehow we construct our world of things and events out of impressions which are themselves not thing-like. The conception of sensory organization implies a putting-together of non-objective elements in perception. On this theory, the data of sense still have to be translated into an awareness of objects and events. According to the present argument, however, the objective world does not require for its explanation a process of construction, translation, or even organization. The visual world can be analyzed into impressions which are object-like, and these impressions are traceable to stimulation. The fundamental impressions obtained by introspection are not colored bits of extensity but variables like contour, surface, slant, corner, motion, distance, and depth, in addition to color, all of which correspond to the variables of a distribution of focused light. These impressions do not require any putting together since the togetherness exists on the retina. The suggestion is that, philosophers and estheticians to the contrary, order exists in stimulation as well as in experience. Order is just as much physical as mental.

10

Geometrical Space and Form

The Space of Geometry . . . . The Problem of Visual Form . . . . A Psychophysical Approach to Form Perception

In Chapter 5 it was argued that the abstract space of points, lines, and planes was a poor conception with which to begin the analysis of how we see, for no one has ever seen it. The investigation of space-perception got off to a bad start by taking the space of geometry as the phenomenon to be explained. We chose to study the visual world instead. Now that we have a theory of the concrete phenomenal world, however, do we also have any insight into the conceiving of abstract space? Why do we find the postulates of geometry satisfying and why is empty space so convenient and easy to imagine?

Similarly, we chose to study objects with surfaces and edges instead of geometrical forms. Triangles, squares, and circles, however, can be drawn on paper and can be perceived. How they are perceived is a puzzle. What is the status of these abstract geometrical figures in psychology?

The Space of Geometry

If the tendency toward invariant sizes and shapes of things in spite of variations in their distance and in the observer's point of view holds true for interspaces as well as objects, it is possible to imagine that something like matter or an ether is also constant in these respects. If the ground tends in perception to have the quality of scale, so might the boundless firmament above and the earth beneath. If this scalar quality seems to extend as far as we can see, perhaps it never ends. All these features of visual perception can be abstracted, and this abstracting was what the Greek geometers were the first men to do. If the world is emptied of objects and only their ghosts in the form of points, lines, and planes are imagined, one thereby simplifies visual thinking. Size constancy then finds expression in the postulate that two straight lines in the same plane can be drawn which, however far they are extended, remain equidistant. This is a form of Euclid's famous parallel postulate, which is one of the characterizing features of "self-evident" geometry. Empty space as thus conceived has been described as analogous to a box without sides. The unvarying sizes and shapes of things also find expression in the abstract idea that a geometrical form can be shifted

about and made to coincide with another identical form. This leads to the axioms of congruence and the identity of dimensions and angles in Euclidean geometry. Empty space is rigid, as are solid objects.

The self-evidence of these postulates, however, was somewhat shaken when, centuries after Euclid, geometers began to study optics and perspective and to examine their visual experience . In the visual field, parallel lines meet at a vanishing point, and this suggests that a location called infinity is thinkable. What are the consequences if parallel lines are assumed to intersect in a specific way? Objects change shape in the visual field as they move. How then can one describe these transformations? Things expand in the composite visual field from one pole and then contract to another pole 180 degrees away. What kind of geometry does this imply? These were not, of course, the literal questions which the non-Euclidean geometers asked themselves, but the connection between the study of vision and the new geometries is too suggestive to be accidental. In recent times these geometries have flowered and, in their applications to physics and astronomy, have been popularized. Everyone has heard of curved space and nearly everyone is puzzled by it. Why is such a conception at once reasonable and unreasonable?

If geometry is an abstraction from experience, then Euclidean geometry is an abstraction from our experience of the visual world whereas non-Euclidean geometry is an abstraction from, or is at least suggested by, our experience of the visual field. The latter is a more sophisticated experience and is harder to attain. The non-Euclidean geometries, correspondingly, developed later in the history of mathematics and are harder to understand.

The space of Euclid, the Cartesian coordinates, and the Newtonian universe were based on the visual world of human behavior. We are, after all, terrestrial animals and our actions presuppose the ground, upright posture, and forward locomotion. These abstract to three dimensions, and to a rigid space with absolute location and absolute motion. This rigid space is perfectly adequate for terrestrial measurements and the analysis of terrestrial events, just as the visual world is adequate for locomotion and our dealings with ordinary objects. If it is not adequate for astronomical events and measurements involving the speed of light, as the theory of relativity suggests, it is at least the primitive conception of space from which the more erudite notions are derived.

Whether the space of the physical universe as a whole is or is not Euclidean is a problem for physicists and astronomers. The suggestion has sometimes been made, however, that perceived space is non-Euclidean, and such a conclusion has recently been reached by Rudolf Luneberg on the basis of a mathematical analysis of binocular disparate images (76). Most of us are bewildered by such a conclusion. The first difficulty with it is that it fails to distinguish between an imagined space and a perceived world. To say that a space is not Euclid's space may be intelligible but to say that the visual world does not follow Euclid's postulates violates common sense. A more important difficulty, however, is that it rests on a confusion between pictorial seeing and

everyday seeing. The visual field, to be sure, is non-Euclidean in the sense that its geometry is based on perspective. The visual world, however, is an experience of quite a different sort. Its geometry is presumably based on the practical biological necessity of estimating the dimensions and judging the shapes of the environment we live in, for only if things keep a certain rigidity of appearance can we identify them for what they are.[1]

The Problem of Visual Form

The so-called constancy of the shapes and sizes of things seems to be a corollary of the perception of a visual world whose surfaces have the quality of slant and the quality of visual scale. Underlying this visual world is always the primary surface of the ground. We can now better understand why the world does not seem to expand as we move forward in the environment, although the retinal image does. The expansion is a stimulus correlate for the sense of moving forward. An object does not appear to contract as it recedes, for the contraction is a stimulus for its recession. The face of an object does not look compressed when seen at a slant since the compression is a correlate of its looking slanted.

The changes of the retinal image which, it always seemed obvious, ought to produce a change in the size and shape of the object are actually stimulus variables which yield changes in the distance and the orientation of the object. The size and shape of physical objects are not represented in the retinal image although they are specified by it. The natural assumption has been that an outline on the retina yields an outline in perception and hence, that a deformed outline should yield a deformed percept. This assumption is reinforced by our interest in the lines and forms which we can draw on paper, for which it holds true that a modified drawing yields a modified percept. When the retinal deformations of outline do not have the expected effect in perception, we are faced with a paradox. How can the shapes and sizes of objects be constant in experience? Perhaps the fundamental error lies in making the original assumption. Conceivably, a deformation need not yield a change in the perceived form. But if that is true, what enables us to have the experience of a visual form in two dimensions?

Form, as we refer to it here, means projected form – a silhouetted shape as con-

[1] When Luneberg suggests that perceptual space is the hyperbolic type of non-Euclidean geometrical space it only confuses me. The obstruse and theoretically unclear set of facts which it seems to account for (the alley experiments) are not obtained in a situation with full illumination and optimal conditions for depth perception. The argument is based entirely on an analysis of binocular disparity of images, leaving out of consideration the geometry of perspective as it applies to size, texture, motion, and other types of stimulation. Perceptual space as we get it under optimal conditions – with constancy of size and shape – is so plainly and simply the space from which Euclid abstracted his geometry, and this conception is so illuminating for all the constancy experiments which yield 100 per cent constancy, that to deny it for the sake of the alley experiments seems unjustified. The application of mathematics to space perception is a fertile field, but the conclusions will be no better than the assumptions with which the mathematician starts.

trasted with a shape in depth. This abstract geometrical form is not, we have argued, a primitive spatial impression at all. The primitive impression is a form-in-depth, and its two aspects of projected form and depth as such are abstractions never experienced in isolation. What we are concerned with is a conceptual experience represented by outlines on paper.

The main stumbling-block in the whole history of our efforts to understand perception, one might venture, is the tendency to think of form as two-dimensional only. The form-on-a-plane, the geometrical form, the drawn form — such forms are literally nothing but the shadows of things, that is to say, their projections. Although Gestalt theorists have recognized that depth is as fundamental in perception as mere extensity, their central problem and guiding concept was the shadowy form, not the substantial one. Might it not be that the dynamics of projected forms, so diligently studied, consists not in the laws of form as such but in the laws that relate them to solid objects — the laws of projection and transformation?

This way of considering size and shape points toward a conception of visual projected form which is very different from the traditional one. Perhaps a closed outline is not an independent entity as we have tended to think but some kind of variable. Perhaps we should conceive form not as a thing but as merely one of the variables of things. The projected shape of a perceived object would then be only one of its visual qualities among others such as the slant of its surfaces, its size, its color, its texture, and its distance, all of which can vary continuously along a scale or dimension. If we cease to think about forms as a set of geometrical entities and concentrate instead on the transitions between them, as we have learned to do for colors, our thinking about the visual process may be clarified. This endeavor will be made in the remainder of the present chapter, but first, we may ask, what stands in its way?

Against this conception there stands the tendency to think of a form as unique. How can there be more or less triangularity in a triangle, we ask? A form, according to the emphasis of Gestalt psychology, is something more than the sum of its parts. "A shape," to quote Koffka, "is itself and nothing else" (67, p. 175). As an assertion that configurations are not reducible to elementary point-sensations this emphasis is surely correct, for the triangular quality of a triangle cannot be derived from the points as such. It does not, however, tell us how to analyse the perception of form or how to understand the ways in which we see and discriminate forms. It suggests that form is unanalysable. If every form carries its own law there can be no laws common to all forms. As a scientific hypothesis this emphasis is of no value whatever, once the principle is accepted that we need to study forms and not retinal points. Although unique and unanalysable forms play a leading role in Gestalt theory, the actual accomplishments of this theory consist of attempts to find the laws of form.

The only laws of sensory organization discovered, however, have been principles such as proximity, similarity, symmetry,

and good continuation (119). The criticism can be made that they are intended to explain only why the perceived form is different from the retinal form and that they are not relevant to the main question: Why is the perceived form specifically related to the retinal form? Only if there is a specific relation is the perception good for anything, since only thus can it be related to the outside world. This relation cannot, it is true, be one of simple pictorial correspondence but it does not need to be pictorial as long as it is specific. The Gestalt theorists, however, being aware that the perception is not a copy of the retinal image, assumed that it could not be wholly specific to the retinal image (like other theorists before them) and went on to account for this discrepancy by a theory of dynamical processes in the brain. The laws of sensory organization were the expressions of such processes. These principles of physiological self-arrangement gave a simple explanation for errors in form perception. In some respects self-arrangement in the brain seemed a better explanation of the errors than a distorting effect of well-remembered or frequently experienced forms, but in other respects it did not, and a controversy which was never settled arose over the role of sensory organization versus past experience in form perception (121, Chapter 4). Neither a dynamical brain process nor an interpretive brain process, however, is relevant to the primary question of the psychophysical relation between retinal and perceived shape.

The writer has concluded, after puzzling about visual form perception for a good many years (35), that all the experiments in which a subject is required to draw (or recognize) a visual pattern from memory or from just previous observation yield no information whatever about why the subject's response is like the original pattern. That is a problem of psychophysical correspondence, analogous to why a certain wave length looks red. What these experiments do yield information about is how a subject learns to discriminate between similar patterns and how we learn to conceptualize objects. Such learned discriminations are of the greatest importance but they must be founded on unlearned discriminations, and it is the latter with which we are now concerned.

A Psychophysical Approach to Form-Perception

How do we judge the shapes and estimate the dimensions of the environment we live in? This is the question too often forgotten in discussing the perception of abstract projected form. For a mobile animal like man, the very essence of environmental shapes and dimensions is that they are successively transformed on the retina. If these locomotor transformations all yield perceptions of the same objects why could not other transformations yield the perception of all possible different visual forms? Certain regular transformations (expansion, contraction, a one-way compression, a certain kind of skew, and a simple transposition over the retina) go with a perception of the same shape. They are related to the geometry of perspective and parallax. Certain other transformations, not experienced as continuous during locomotion, go with the perception of different shapes. They are not confined to

the geometry of perspective and parallax.

Conceived thus, visual outline forms are not unique. They could be arranged in a systematic way such that each form would differ only gradually and continuously from all others. Such dimensions of similarity among contour forms have never been explored (and there must exist an enormous number of such dimensions) but the fact that modern geometry can specify general modes of transformation suggests that the exploration ought to be possible. "The concepts of modern geometry," according to Cassirer (21, page 8), "derive their precision and true universality only from the fact that the intuited particular figures are not considered as pre-given and rigid, but rather as a kind of plastic material capable of being moulded into the most varied forms."

The seemingly infinite variety of visual forms is the crux of the difficulty. This manifold has usually been reduced to order only by classification, beginning with triangles, squares, circles, and the like. The result is a set of groups or mutually exclusive categories analogous to the classes of individual things and persons implied by Aristotelian logic. The Greeks, and especially Plato, thought of geometrical forms in this way and the tendency has persisted. Classification is not, however, the only or the best way of ordering a manifold. Serializing is more apt to bring out the fundamental relations between things (75). If we are ever to understand exactly what yields a perception of shape we must study the dimensions of variation of visual shapes.

Geometrically considered, visual forms do not fall into any mutually exclusive classes or categories except the so-called transformation-groups, and the closed visual outlines now being considered are all in the same transformation-group (25, Chapter 4 and 5). Any closed form can be transformed into any other closed form by a perfectly gradual or continuous change. The psychological quality of shape which changes with such a continuous transformation is what needs investigation. The normal or standard shapes familiar to everyone are no more than special points of anchorage on a continuous dimension of variation. They are norms or standards of reference for shape, not entities of shape.

Figure 75 is a chart of a few very simple dimensions of variation. It illustrates the modes of transformation which comprise the manifold of all rectangular figures. The square lies at the center of the chart, but there is no unique square, for the equal-sided shape varies in size along the diagonal. The discriminable qualities of shape are very evident but they are not so easy to name as one might suppose. Thinness to thickness and tallness to shortness are perhaps the most obvious terms to apply. The dimension of variation running from upper left to lower right is a quality of shape which, however it be named, appears to some observers as simple as the supposedly elementary quality of warmth to coldness.

The chart may be conceived to include either transformations of the same object, which we get when we move about in the environment, or transitions between different objects, which we perceive as qualities of outline-shape. The dimension from lower left to upper right is a change in projected size such as accompanies

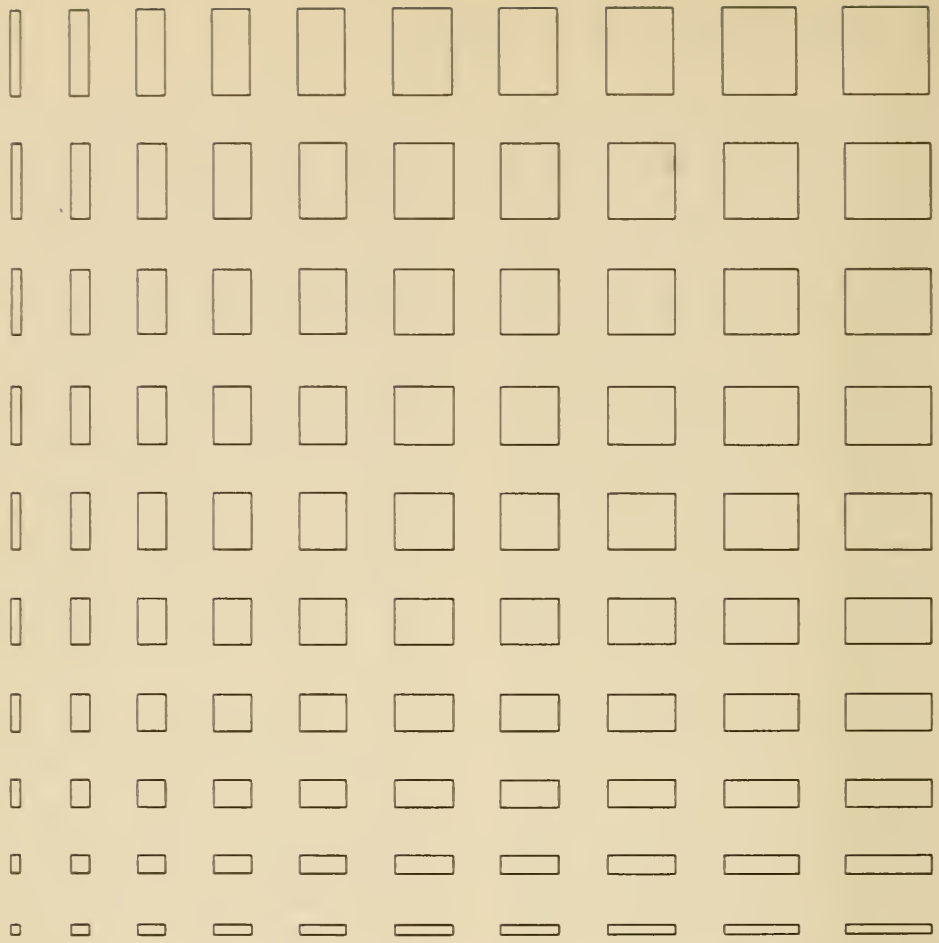


FIGURE 75. The Transformations of Rectangular Shapes

These transformations are based on parallel projection rather than on polar (or perspective) projection.

getting progressively closer to a given object. Whether it is seen as a change in distance (with size constant) or a change in the size of objects (with distance constant) is indeterminate; only with a normal environment does one or the other alternative become realized.

The main significance of such a chart is that it makes possible psychophysical experiments on the variables of outline shape in the frontal plane. Having a method of systematically varying the stimulus, one can determine the corresponding variations in the perception, as psychologists have done for the variations of color, sound, and the whole gamut of the qualities and intensities of experience. In so doing, one can hope to uncover the nature of the specific relation between the stimulus and the impression of shape.[2]

[2]This suggestion, like much of the chapter, is vague and speculative. As this book goes to press the writer has begun a set of experiments based on the hypothesis that a flat outline shape on paper *never* yields a determinate perception; it only yields a presumption of a real shape. Perhaps shape cannot be studied intelligibly apart from the slant of its surface.

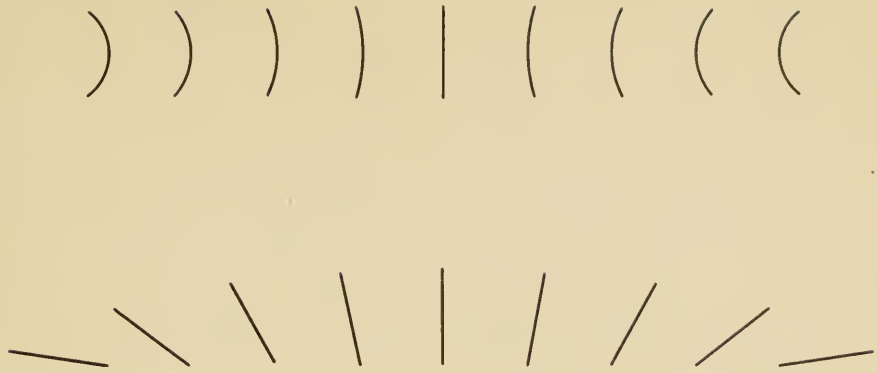


FIGURE 76. The Qualities of a Simple Line

The writer suggested some years ago (38) that a visual line or border has two variable qualities, besides length. One is left slant...zero slant...right slant, and the other is convex...straight...concave. A line looks as if it had those phenomenal properties and behaves in perception as if it had them. The two dimensions of variation are as much sensory as are the hue and brightness of color. They could be termed the quality of direction (linear slope) and curvature (linear shape). The names are inadequate, but the geometrical variables to which they correspond are exact; the stimulus variables are the first differential and the second differential of a curve in analytical geometry. The first and second differential are sometimes explained as the slope of a curve and the sense in which the slope is changing. Mathematically, these two variables determine a curve at all its points. Phenomenally, the two corresponding qualities determine a visual line or border in all its (conveniently chosen) segments. If one specifies the direction and curvature of a short visual line has one not specified the entire experience?

The evidence for this unorthodox analysis of border sensations came from two sources. First, it was possible to vary line segments in these two systematic ways and observe the corresponding variations in the qualities of direction and tilt. A psychophysical correspondence could be established (38). The variables could be isolated and the qualities could be discriminated. Second, it was a fact that tilt and curvature changed in the same way as do sensory qualities like white-gray-black or warm-neutral-cold during and after prolonged stimulation. When a line was stared at, the tilt (or curvature) tended toward neutral, and thereafter an upright (or a straight) line looked tilted (or curved) the other way. In short, the familiar after-images of color were paralleled by after-images of tilt and curvature, slight, but definite in amount (36, 37, 38, 43). The after-image of color could be understood as

a shift in the psychophysical correspondence, and so could the after-image of tilt or curvature.

If curvature and direction are the variable qualities of a border, there is a possibility that a closed border — a form — may be reducible to variable qualities. The variables of a rectangular figure have already been illustrated. The changes of curvature and direction around a contour which determine its shape may become enormously complex. They seem to be integrated or organized to yield qualities of a higher order. The quality of closure itself, one must admit, does not appear to be a variable quality. It goes with "thingness" and suggests a theory of dynamic organization which is difficult to analyze. When a line becomes a contour, the variations of shape appear to jump to a higher level. But there is no reason to assume that shape cannot be reduced to its variations and that a form is unanalyzable.

The theory of visual form which this approach suggests is rather different from that of contemporary Gestalt theory. The latter is based on the celebrated laws of visual organization discovered by Wertheimer (119), expanded by Koffka (67), and often generalized as the laws of all perception. These principles were based on drawings, chiefly of points, lines, and curves, in different arrangements or regularly varied compositions. Elements of this sort can vary in their spacing, curvature, orientation, and modes of intersection. The fact is that the points and lines look like groups or suggest complete outlines in these artificial arrangements. Moreover, the groups or outlines which tend to occur depend on (1) the proximity and (2) the similarity of the elements to one another, on the tendency of lines (3) to make outlines (closure), and on the tendency of outlines (4) to be smooth (good continuation) and (5) to make simple forms (good shape). These factors were interpreted as laws governing the appearance in perception of a figure on a ground. The implicit assumption was that points and lines have to be unified for a figure to appear in perception. The factors could therefore be taken to indicate forces of neural organization operating among the elements.

The laws of organization have already been criticised on the grounds that they do not account for accurate perception, Koffka acknowledged the determining role of the retinal image only in the postulate of external forces of organization which set limits on the internal forces of organization (67, p. 138). In addition to this criticism it is also possible to doubt whether the particular arrangements of points and lines which Wertheimer devised isolated the fundamental types of visual stimulation. The images of texture and contour need to be understood before the images of point-groupings and drawn outlines are studied. Wertheimer's drawings were nonsense patterns of the extreme type, far removed from the images of a material world. His laws are applicable, therefore, to some kinds of abstract drawings and paintings, as Kepes in *The Language of Vision* recognizes (63), but not so much to ordinary visual stimulation.

11

Meaning

Meaningful Perception . . . . How Much Perception is Learned? . . . . The Possibility of Spatial Meanings . . . . Detachable Significance . . . . Are There Unlearned Meanings? . . . . The Alteration of Spatial Perceptions by Meanings . . . . The Literal Visual World and the Schematic Visual World . . . . Conclusions

Let us consider the vision of one of our very early ancestors some five or ten million years ago on the plains of Asia. He was no longer living in trees and, although we know little about him, he would probably be classified as a member of the *genus homo*. His eyes were probably just like ours and, if this is true, so were his retinal images. The shades, borders, and gradients of light which composed these images were specific to his environment and were the immediate cause of his seeing the environment. He was probably a sharp-eyed creature for, as we know, he survived a rather risky existence. Since he probably had little or no language, he had no names for things and we shall never know what his ideas or his conscious experiences were. But we do know this. He discriminated among the variations of his retinal images and could therefore react differentially to the objects of his environment.

We can hardly say that he was conscious of geometrical space but we can be sure he discriminated the various distances of an object. For example, one conceivable object to which he must have been sensitive was a sabre-toothed tiger or some beast of equal ferocity. His conduct must have been rather nicely adjusted to distance when he encountered one in open country, varying as the retinal image varied in a precise way. To the tiger at a mile he could react by going about his business. To the tiger at 400 yards he should have reacted by going in another direction. To the tiger at 10 yards he must have reacted (if he *was* one of our ancestors) by running like the wind. His behavior was graded in relation to a variation of his retinal images.

Moreover his behavior was specific to the contour, shape, size, color, and motion of objects. He did not confuse zebras with tigers for, we may conjecture, he pur-

sued the one and fled from the other. The differences in visual stimulation went with differences in his behavior, and we can therefore be sure that he could identify things. The animals that he could eat or that could eat him may not have been named but he could react appropriately to such light-reflecting objects, and they must have aroused at least a primitive kind of meaning. It was, in fact, on account of their meaning that he needed to discriminate them. The color, shape, motion, and distance of things were of no interest to him in themselves. These abstractions were merely the identifying features, often slight and subtle, of objects which invited or compelled action.

Judging from his probable behavior, primitive man discriminated the solidity, separateness, and spacing of things with great accuracy. In his place, we would say that we saw a visual world. But he also behaved toward things with circumspection, for he saw a world of meanings. Speculative as all such accounts must be, it is reasonably certain that our primitive ancestor got about in his environment, and knew one object from another. His behavior was based on locomotion and recognition: it was adjusted to space, and at the same time it consisted of reactions to objects. Vision provided him both guidance for his actions and cues for his actions. Presumably, then, an object like the sabre-toothed tiger was both localized and meaningful in his experience since he reacted to both its distance and its significance. Heretofore we have been mainly concerned with the question how he could localize the tiger. In this chapter we must turn our attention to the question how he could know the tiger.

Meaningful Perception

Our own experience of the visual world can be described as extended in distance and modelled in depth; as upright, motionless as a whole, and unbounded; as colored, textured, shadowed, and illuminated; as filled with surfaces, edges, shapes, and interspaces. But this description leaves out the fact that the surfaces are familiar and the shapes are useful. No less than our primitive ancestor, we apprehend their uses and dangers, their satisfying or annoying possibilities, and the consequences of action centering on them. Surfaces and shapes are in actuality perceived as ice, apples, fur, fences, clouds, shoes, people, and so on. Furthermore, our world is enlarged and complicated as compared with that of our ancestor, by the inclusion of certain forms, lines, and man-made patterns which we know as pictures, symbols, and printed words. The visual world, in short, is meaningful as well as concrete: it is significant as well as literal.

The psychology of meaning is a large subject. One difficulty is that there are so many levels or kinds of meaning. For example, there is first of all the possibility of a sort of primitive concrete meaning which either results from the infant's active exploration of his physical environment or is evidenced by such action. He fingers and manipulates things, and later he gets about among obstacles and goes for moveable objects, and pulls and pushes and upsets things. Examples of such

concrete meanings for the adult would be the way things look as if they were capable of being grasped or pushed or walked on. Second, there are all the simple use-meanings or meanings for the satisfaction of needs such as are embodied in food-objects, play-objects, tool-objects, dangerous objects, and what Freud called love-objects, the parents being the first instances of the latter. For example, food looks eatable, shoes look wearable, and fire looks hot. Third, there are the meanings of instruments, devices, constructions, and machines. Fourth, there are the values or emotional meanings of things which make the shapes of the world attractive or repulsive in a vast variety of ways. Fifth, there is the kind of meaning exemplified in signs, by virtue of which one object or event suggests another not physically present. Clouds are said to be a sign of rain. The red light means stop. Significance may also be defined broadly, if one wishes to do so, to include the simpler kinds of perceptual meaning referred to above. Sixth, there is the particularly human kind of meaning embodied in symbols. Such meanings are said to be abstract. Names mean things or persons, but their principal advantage is that they can also stand for classes, variations, and properties of these things. Carriers of symbolic meaning like money, flags, words – the latter above all – are in common use among people who interact with one another; they are completely determined by culture. Symbolic meanings are the most complex and the most momentous of the list. They mediate knowledge, as distinguished from perception, and they are the basis for reasoning, creative imagination, invention, and discovery. The world of symbolic meanings stands at a far extreme from the world of surfaces, edges, and shapes with which this book is primarily concerned. The former induces thinking, the latter only sitting, standing, walking, and grasping. Nevertheless, they are both the same world. Things must be substantial before they can be significant or symbolic. A man must find a place to sit before he can sit down to think.

The kinds of meaning listed are not exhaustive. There are also the unspoken meanings which go with visual motion such as the expansion, deformation, and transposition described in Chapter 7, and these need special consideration. There are the meanings of perceived events and sequences. There are meanings when one surface touches another, or collides with another, or when one object produces an action in another. There is also the whole range of social meanings, facial expressions, gestures, persons and actions between persons (92, 52). The visual world is saturated with many kinds of meaning, and it seems to get fuller with meaning as we live from year to year.

How Much of Perception is Learned?

We can be sure that the meaningful world is fully achieved only by means of learning. The unsettled question is, how much of it can be achieved without learning? Are the primitive concrete meanings unlearned? Are the fundamental spatial impressions unlearned? Is there an embryonic unlearned meaning for every perception? The issue has remained alive ever since John Locke asserted in 1690 that all knowledge

of the world is learned, the mind at birth being a complete blank. The parallel question, of course, is how much behavior is achieved by learning and how much arises spontaneously with the growth of the organism.

Having discarded the doctrine that men do not learn their sensations of color but learn their perceptions of everything else, what shall we say to this question? The simplest theory to fit all that has gone before might be to suppose *that the visual world is an unlearned experience, that it is meaningless when seen for the first time, and that what one learns is to see the meanings of things.*

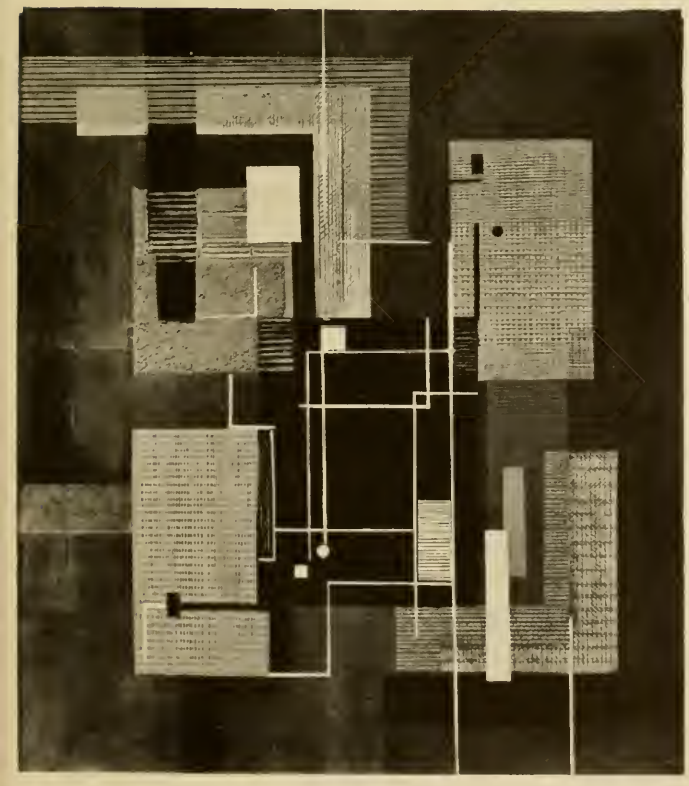
As a general formula, this is consistent with a good part of the evidence. It is at least an improvement over the sensation-perception theory, and as an approximate or working hypothesis it is worth adopting. But, as will be evident, it makes a number of simplifications which are at best imperfect. Moreover, it fails to fit with another general formula which seems to be valid for all the studies that have been made of instincts, habits, and capacities in men and animals, namely that no activity is ever either wholly learned or wholly unmodified by learning.

The first oversimplification is the assumption that the constituents of the visual world such as colors, surfaces, shapes, edges, and interspaces are in themselves meaningless. The second oversimplification is the assumption that the meaning of anything is detachable from its concrete spatial qualities; that one can separate things and events from their meanings by introspection. A third oversimplification is that all meaning is learned – that there are no unlearned meanings. And a fourth is that when meaning is added to things it does not substantially modify their concrete spatial qualities – color, size, shape, motion, and all the rest remaining unaffected.

The evidence regarding learning and meaning in visual perception needs to be gone over if we are to judge the adequacy of these four generalizations. We may consider them in order.

The Possibility of Spatial Meanings

The introspective observations of Chapter 3 suggested that the characteristics of the visual world could be described without reference to the practical meanings of the particular environment observed. But very possibly the practical meanings which make things look ordinary and useful are not the only kind. The world may have a residual and important sort of meaning even when viewed with the purest attitude of contemplation. Even what was termed the visual field – the array of colors divorced from objective character – may not be wholly meaningless if painters are to be believed. Colors, they say, have their own meanings, whatever the doctrine of sensation claimed, and an array of colors has a meaning appropriate to the array, even when it does not compose a recognizable scene. There are distinguished painters who are thoroughly convinced that an array which does not yield familiar objects on a level ground can have a more interesting meaning than one which does! Pictures of this sort are not projections of existing or physically possible environments and are therefore non-representative. Abstract or non-objective paintings do not have use-



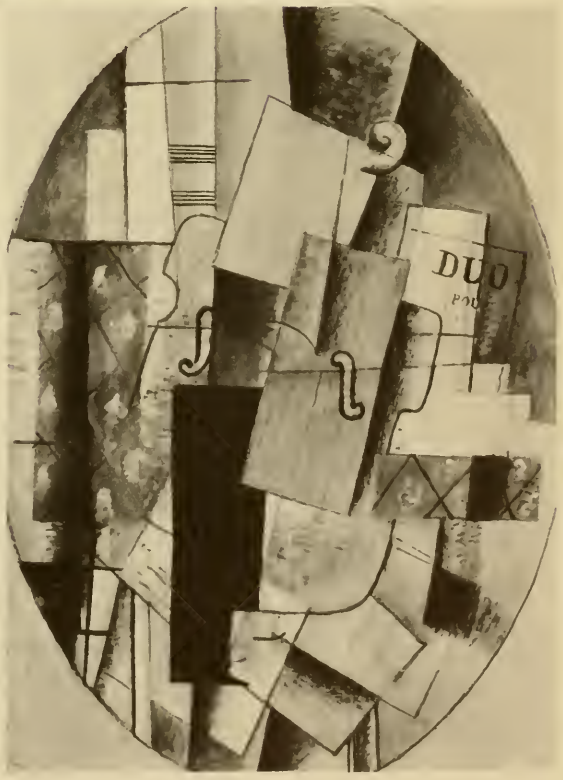


FIGURE 77

Examples of Abstract Paintings

Courtesy of The Museum of Modern Art, New York

meaning, and their forms are not symbols (as defined); but it is profitless to deny that they have any meaning, for the argument reduces to a conventional definition of the word "meaning." The definition of this term has never been agreed upon.

For linguistic meaning, it is true, systematic definitions of signs and symbols can be worked out on the basis of logical and psychological theory. This has recently been done by Morris (84). But visual meaning has so far defied systematic analysis and the whole subject, including art-criticism, is notoriously speculative.

The common-sense opinion and the view of tough-minded psychologists is that colors and shapes which do not produce in perception remembered objects or things which the onlooker knows what to do about should be termed meaningless. For the psychologist, meaning originates in adaptive response. Drawings not resembling anything familiar are called nonsense forms and, along with nonsense syllables, are employed in experiments on memorizing. They are also shown to observers under conditions which impoverish the retinal image, for example a very brief exposure, in experiments on the accuracy of perception. Nonsense forms are much more difficult to reproduce or discriminate than familiar or conventional forms. The fact is, however, that the results of these experiments suggest that nonsense forms are only relatively meaningless. Meanings which are peculiar to the individual observer are always reported and are, in fact, what make the forms memorable. According to E. Gibson, nonsense syllables or nonsense forms can be memorized only insofar as they have been differentiated from one another; memorizing depends on the formation of a unique or identifying response to each nonsense item, and this is the next thing to the formation of a meaning for each item (33, 34). The experiments on drawing nonsense forms from memory, summarized by Woodworth (121, Chapter 4), point in the same direction: a senseless form must gain sense in order to be recalled. Research on learning to recognize aircraft (39, Chapter 7) strongly suggests that when a nonsense form becomes identifiable it also becomes meaningful. The common-sense or toughminded view, therefore, is not strictly in accordance with the facts. Non-objective forms lack namable or socially agreed-upon meaning, but it is not true to say that they lack all meaning.

A visual scene which has modelling, color, texture, surfaces, shapes, and interspaces is ordinarily caused by a physical environment, and constitutes what we have called a visual world. Almost invariably it has a fundamental surface with the meaning of a terrain or a floor, since human animals live as they do. On rare occasions, however, men peer into microscopes or look under the sea through diver's helmets, they stare at flames and clouds, and they sometimes gaze at abstract paintings. No terrain or floor is induced in these perceptions. Here, the modelling, color, texture, and the rest of the variables of the visual world do not produce a visual world of the everyday sort. Take the abstract paintings, for instance. The surfaces and shapes induced by them do not, of course, correspond to any physically existing surfaces and shapes. The personal meanings they arouse could never be verified, and,

consequently, they are not practical. The surfaces and shapes seen by the diver and the biologist, in contrast, are real, but the environment is not a human environment. The world in the microscope is novel; the space of the cloud-patterns is useless; the things in the painting are imaginary; but they are composed of the same stimulus variables which enable us to get around in the environment we inhabit. A contemporary artist and teacher, Kepes, has called these variables "the language of vision," referring to the painter's use of them as abstract constituents of perception (63). They have meaning, and the experimental combination of them creates new meanings, but they are essentially spatial and it is pretentious to suppose that they enable the painter to represent some realm of non-physical reality. The claims of some abstract painters that by the use of these plastic devices they can represent high and hidden truth, however, should not force us to the opposite extreme of declaring that surfaces, edges, and shapes with color are without meaning, for this would be false. They have at least the meaning of surfaces, edges, and shapes.

The implication seems to be that spatial impressions can be largely but not wholly meaningless. If spatial meanings are inseparable from visual objects, the first generalization of the formula being tested is too simple. Perhaps these meanings are to some extent unlearned. Conceivably, then, only the learned meanings are separate from spatial qualities and are what get attached to the visual objects. What is the evidence for this possibility?

Detachable Significance

Titchener's context theory implied that meaning accrues to impressions, that is, is added to them. Meaning is attached by association. A bell acquires the meaning of dinner, or someone at the door, or a streetcar at one's back, depending on the context. Stimuli come to arouse specific expectations, in more recent terminology, and the perceptions thereby acquire significance.

Qualities of touch, temperature, and muscular feeling can unquestionably be added to a purely visual perception by experience in manipulating and using the object in question. Presumably this is why fur looks soft, ice looks cold, and a book looks openable instead of solid and box-like. Manipulative meanings are often added to visual perceptions during adult life. The change can then be noticed introspectively as, for instance, in the case of a new kind of hand-operated tool encountered for the first time. After its use is understood the object looks different. The perception now has properties it did not have before and – here is the point – the properties are not directly aroused by present retinal stimulation. The association theory ascribed them to memory-images clustering around the visual shape, and the latter is said to redintegrate one's past experience with the object. It is only logical to suppose that the coldness of a piece of ice or the squeezableness of a pair of pliers is not given by stimulation and must be given by memory. But these qualities are not like the memory images of recollection. Meaning is not literally recall. Squeezableness is something which seems to be located in the object, not in the hand,

and it refers to the present, not to the past. Visual objects appear to have soaked up such qualities and to be fairly saturated with them, the use of the object and the shape of the object being almost indistinguishable. Nevertheless they are distinguishable to introspection and they are separable when the use is learned.

An even clearer instance of attachable meaning is furnished by symbols. Printed letters and words are merely shapes to the illiterate, to foreigners, and to the very young. (The sound patterns and the vocal patterns of words were also meaningless in the beginning.) To most of us, on the contrary, the words we read are merely meanings and we are scarcely aware of their shapes. If the reader will stare fixedly at a word like "abyss" however, he will no-

abyss evil

tice almost immediately that it begins to look like a mere shape. A familiar word like "evil" usually keeps its meaning longer when fixated, but after an interval the visual appearance becomes prominent, the meaning becomes separated off somewhere, and the word disintegrates into e-vil or e-VI-l. Eventually even the letters may begin to look unfamiliar and the word can become completely geometrized. If you repeat the word to yourself rapidly, the sound also becomes meaningless and the pronunciation may tend to disintegrate. The phenomenon has been called loss or lapse of meaning (7, 27, 96), but it is better to say that the meaning recedes from the word, since it frequently comes back spontaneously, and even when it is gone the observer usually has a feeling that he could get it back if necessary.

My own observation indicates – and it is confirmed by experiments performed by my students – that this recession of meaning can be noticed with prolonged attention to objects other than words. If you stare at a paper match long enough, its familiarity disappears, although it does not break up into sub-units as a word does. The same thing happens when a piece of sandpaper is rubbed monotonously with the fingers. Perhaps any scene begins to appear strange when the eyes are fixated long enough. The attempt to observe one's visual field leads in this direction as does, in a certain sense, the painter's intent view of something which interests him.

There is a third kind of evidence for the separability of meaning and spatial impressions – detachable meanings – and this is derived from studying the perceptions of patients with brain injuries. With injury to the occipital lobe of one of the hemispheres there often results a peculiar visual incapacity without apparent incapacity of other kinds. This defect is sometimes called psychic blindness. The retinal surfaces are projected by neurones on the surfaces of the brain in such a way that the right half of the visual field (the left half of each retina) is represented on the cortex of the left occipital lobe. Damage to the nerve-cells in this area produces a sort of a half-blindness called hemianopsia. The patient is unable to make any visual discriminations to the right of his line of sight, as is proved by mapping his visual field with the eyes fixated on the midpoint of a screen. The

visual field has been constricted to half of the ordinary oval.[1]

Simple right hemianopsia in right-handed persons is often accompanied by word blindness or alexia. The patient can, by ordinary tests, see things fairly well, but print has become wholly or partly meaningless. The letters are visible enough, but they are no longer words. In one test, for instance, he may be able to take a set of children's alphabetical blocks and turn each block to make the letter right side up without being able to name a single letter or arrange the blocks in alphabetical order, let alone read print (117, page 288). Or such a patient may be able to trace letters with his fingers without knowing what they mean. The forms as such are discriminable but they are no longer phonetic symbols.

Along with word blindness there usually goes some degree of what has been called psychic blindness or visual agnosia. This is a failure to recognize or know the names of objects, persons, pictures, and places. The evidence suggests that the spatial aspects of things – distance, size, shape, location, movement – are relatively undisturbed, but that their meaning is reduced. Verbal and conceptual meanings are affected most, concrete everyday meanings least, and the most primitive meanings seem to be retained.

On the whole, the conclusion is that meaning becomes less intrinsic to and more detachable from spatial impressions the more it approaches high order concepts. At least this seems to be true as meaning becomes verbalized. Spatial meanings are tied into their perceptions, relatively speaking. Verbal meanings and civilized use-meanings are relatively detachable.

Are There Unlearned Meanings?

The fact that men differ in the meaning which things have for them is a truism. Men of different training, interests, and convictions do not, as we say, see the same world. A human body is perceived by an anatomist differently from the way in which it is by the rest of us. A nickel is not the same thing to an adult as it is to a child. An industrial machine is not the same thing to its operator as it is to the plant owner. The specialized perceptions of the connoisseur, the photographer, the doctor, the woodsman, and the engineer are all different. Moreover the apprehension of the environment differs in a systematic way

[1] A significant fact, incidentally, is that his visual world is not halved – how could it be? The patient usually complains only that he cannot see things as definitely as he used to. The visual world is not a point-to-point projection of the retinas. The world can still be integrated out of successive fixations of half-fields. The old theory would say that perception fills in the blind area. The point is, however, that a localized area of injury in the visual brain produces a localized area of blindness in the visual field but not a disappearance of objects in a localized area of the phenomenal environment. The brain maps the momentary retinal images of the world but it does not map the world we see. The brain and the retina are in spatial correspondence with one another, whereas the brain and the phenomenal world can only be in a spatio-temporal correspondence – that is, a correspondence other than one of simple point-to-point projection. This means that one cannot expect to find any simple localization in the brain of what is ordinarily called space perception. Consequently it is only natural, not paradoxical as neurologists assume (8), that a localized brain lesion should cause the loss of a piece of the field of view (a blind spot) and at the same time the loss of a feature of the perception of the world (for example, distorted size or defective meaning).

among peoples as well as among individuals. The traditions and culture of an uncivilized tribe consist not merely in strange ways of behaving, such as used to be described, but also in strange ways of perceiving and apprehending things. Accounts by anthropologists of these foreign modes of apprehension, in fact, begin to make the customs intelligible to us. The Trobrianders, it has been reported, recognize that a child resembles its father but never its mother. The former kind of resemblance is proper and to be expected; the latter is inadmissible. The perceptions of similarity are part and parcel of a set of customs regarding kinship. Klineberg suggests that the effect is a failure to note any resemblance rather than an actual interference with the sensory perception (65), but this is hard to be certain about. In any event, it is a perceptual custom which is foreign to our ways of seeing people.

All these facts seem to point to the general conclusion that each man learns the meaning of the world for himself, within the framework of his upbringing and the society in which he lives. Significance and value are formed partly by the cultural background and partly by the individual's unique experience, but in any event are learned. This is the conclusion of what has been called empiricism. It is consistent with the emphasis on learning and education and with the scientific approach to human nature and culture in the history of modern thought. The conviction that all meanings are learned implies that human beings are plastic rather than cast in rigid molds. It is such a valuable hypothesis that exceptions to it should be carefully scrutinized.

Nevertheless there is reason to believe that some meanings, or some components of meaning, are not learned. When McDougall applied instincts to human psychology at the beginning of this century he included in his definition of an instinct "an innate disposition to perceive and pay attention to objects of a certain class" (80). He meant that there were intrinsically fearful objects, intrinsically good-to-eat objects, objects which intrinsically invited acquisition, mating, curiosity, self-assertion, and all the rest of the list. The emotional excitement, the impulse to action, and the purposive striving of an instinct were all contingent upon an innate perceptual inlet. The instinct theory led to absurdities in social psychology and encouraged a tendency to name various aspects of human behavior instead of describing them. Instincts fell into disrepute among those interested in the growing science of human learning. Stimuli, reflexes, and chained reflexes were more exact terms with which the experimenter could work. But the neutral term stimulus is not adequate to explain why behavior is a function of objects, and the theory of patterned stimulation or *Gestalten* arose to reintroduce the notion of intrinsic meaning in a new fashion. McDougall was surely wrong about intrinsically fearful or eatable objects in perception, but perhaps his successors have overstated the case for learning.

Among animals, certainly, there is good evidence for innately meaningful perceptions. Lashley has stated the case convincingly in the following way (71). If

you put the question, "What is an egg to a nesting bird?" a perfectly sensible answer can be given. The answer is, "any object which causes the bird to retrieve it, to clean its surface, and to sit on it." Such an object is discriminated or recognized as an egg, or we might say, has the meaning of an egg. The experimental fact is that such behavior in nesting birds is aroused only by rounded objects within definite limits of roundedness and certain limits of size, and with a texture approaching a specific sort. Objects with these qualities are eggs; objects without them, when put in the nest, are not eggs in the sense that they are pushed out of the nest. *This sort of behavior appears during the first nesting season without opportunity to learn.* Fine discriminations are not made between eggs and non-eggs but gross discriminations are. Lashley points out that discriminative behavior of this sort is as frequent as are instincts, for the two go together.

The human animal matures slowly and learns more than the other species. He can eventually make very fine discriminations – among birds' eggs for example – and these are learned. In the use of language he has acquired a special method for learning and retaining fine discriminations. But apart from language and apart from differentiated adult meanings is it possible that infants have wordless and crude primitive meanings for some of their earliest visual impressions?

The facts are hard to interpret and not conclusive. Most observations of babies are biased. The smiling response is a good instance. If we could be sure that babies identify a human face in the visual field when they first smile while looking at a doting adult, it would suggest an innate meaning for the pattern of a human face – an instinctive response nicely adapted to induce or maintain the doting attitude of the adult. To be sure of this one would have to prove that the supposedly instinctive smile was not a random but a visual response; that it always occurred to a pattern similar to a face; and that it did not occur to such patterns as a milk bottle, a lampshade, or a square of cardboard. It is not asserted here that babies instinctively recognize their mothers; it would be enough to show that they have an unlearned response which would make their mothers think they do.

Just this seems to have been demonstrated recently by Spitz and Wolfe (99). Babies between 2 and 6 months of age, they report, almost universally fixate and smile at any face-like object if it moves. Any face with any expression will serve, or a crude mask, or a dummy, as illustrated in Figure 78. A vicious leer is as good as a benevolent smile; the baby is not imitating, or even discriminating, facial expressions. It must be the pattern of a full face; a profile will not do. Too small a face will not do, such as that of a doll. A completely motionless face will not do, nor will a milk bottle, or a toy, or other objects. These are fixated but not smiled at. After 6 months the baby's smile becomes more discriminating. He no longer smiles at a stranger or a dummy; he recognizes his parents; he reacts to an expression of disapproval; he begins to smile when smiled at. All these responses can be learned. But the original smile is apparently not. The effective stimulus for

FIGURE 78. Patterns which Evoke the Infant's Smile

Courtesy of Dr. Rene A. Spitz. From "The Smiling Response", *by Spitz and Wolfe (Genetic Psych. Monog., Vol. 34, 1946)*

it is an indeterminate pattern with only the elementary characteristics of a human face. The wordless meaning of this crude pattern is equally primitive, but it nevertheless is a discriminable pattern and it has the meaning of something-to-be-smiled-at.

These facts, if they are correct, imply that the human infant does not begin to learn meanings at a zero level. They show the falsity of the notion that we are born with a set of meanings ready-made or a set of innate ideas, but at the same time they contradict the notion that all meaning is acquired. That generalization is too easy. There is probably an embryonic meaning which goes with an embryonic visual perception.

The Alteration of Spatial Perceptions by Meanings

The formula being tested – that the perception of space is unlearned, meaningless, and acquires its meaning in the course of experience – makes an additional simplification which is far from perfect. It supposes that when a certain shape, let us say, gains a new meaning, the shape remains just what it was before. The texture, slant, color, contour, and other constituents of a thing are supposed to be unaltered by the gain in meaning. It may be apprehended differently but it is sensed the same as before, to use the older terminology.

There are few generalizations in psychology which have been refuted as often as this one. It can easily be disproved; what is harder to understand is why nevertheless it must have some amended validity. Consider first the evidence for its untruth.

The color of anything, it might seem, should be unaffected by the mere fact that

it resembles something of another color. Duncker, however, has proved that under certain conditions a piece of cloth in the shape of a leaf is judged as noticeably greener than an equivalent piece of cloth in the shape of a donkey, when the light reflected from both was only faintly greenish. There is an effect of what is called memory-color, that is, the color associated with leaves as contrasted with the color associated with donkeys (29). He also found that a piece of brown chocolate had a stronger chocolate taste than a piece of white chocolate when the taster could see them, but not when he was blindfolded. This result failed, however, if the subject came to suspect the purpose of the experiment.

The size of a thing, likewise, ought not to be affected by its meaning. Bruner and Goodman, however, found that ten-year-olds perceived coins as larger than equivalent cardboard disks by about 25 per cent (14). Presumably the value of the coin influenced its size. There was a striking tendency for poor children to see the coins larger than did rich children. This latter result is the dramatic feature of the experiment – the suggestion that need of money and infrequent possession of it do more than make the child stand in awe of a coin; that indeed they modify the very spatial structure of the child's world, and that the substance and dimensions of the environment for a poor child are different from what they are for a rich child. Economic class may affect even the sensations of things. It is worth noting, however, that when Carter and Schooler repeated the experiment (20) only remembered or imagined coins were exaggerated in size; actual coins were perceived by children, rich or poor, in close relationship to their real size.

The outline of a thing ought not to suffer alteration just because one knows what it is good for, nor should a geometrical form change because it looks like a familiar object. Nevertheless, outline and form are modified by meaning, and the fact has been demonstrated over and over again. The writer once found that the nonsense form on the left in Figure 79 might be interpreted by one observer as a woman's

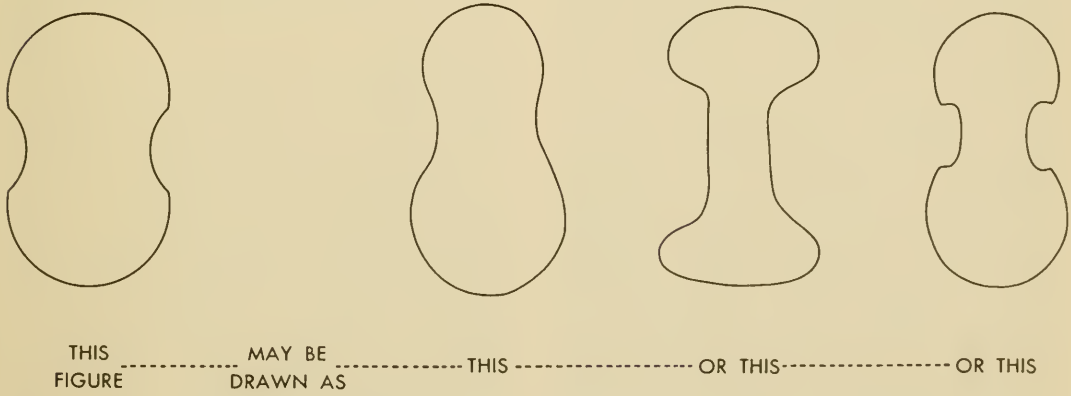


FIGURE 79. The Effect of Meaning on Visual Form

torso, by another as a dumbbell, and by a third as a violin (35). The drawings of the form made by each observer are shown on the right. It might be guessed that the three observers had somewhat different interests in life. It should be noted that the nonsense form had been shown along with others and was reproduced after an interval of time.

The same kind of result has been obtained by many experimenters (121, Chapter 4). The errors made by observers in reproducing visual forms constitute a fascinating subject. Among many hypothetical ways of accounting for these errors (habits, conventions, dynamical organizations, norms, standards, concepts, the law of Prägnanz, or simply the association of ideas), perhaps the most comprehensive is that of Bartlett, who suggests that a pattern tends to be schematized during perceiving and remembering (5). Bartlett himself had investigated story-telling from memory as well as form-perception, and he discovered, as had earlier students of the psychology of testimony, that a story gets retold so as to express the meaning it has for the teller. It is altered in accordance with a schema which is characteristic of the individual, his interests, and his culture.

Color, size, form, sequence, and still other qualities of perception may unquestionably be affected by the past experience and attitudes of the observer. William James once remarked that perception was of probable things. The experiments suggest that perception is also of familiar things, of expected things, of known, typical, average, or normal things, of namable things, of specific or precise things, of valued things, of attractive or repulsive things, and of customary things or things that other people also see. Perception, in Bartlett's term, tends to be schematic.

The study of this tendency in perceptis valuable, for it illuminates a set of problems of the very greatest importance. How do social stereotypes arise – those schematic and usually distorted perceptions of race, nation, religion, and class? Such perceptions tend to be caricatures with respect to physical qualities as well as being simplified with respect to meaning. Why are persons judged as types rather than as persons? Why do rumors and myths persist? How do words exercise their tyranny over thinking?

The Literal Visual World and the Schematic Visual World

It is easy, however, to misinterpret all the evidence for a schematic trend in perception. It is tempting to conclude that all apprehension is selective and distorted; that perception is inevitably a constructive process which creates the world to suit the perceiver; that we see things not as they are but as we are. Any such general conclusion is unwarranted, for it neglects the existence of what we shall call in this book literal perception.

Perception can be studied scientifically by either of two general methods. The first is the psychophysical method of the laboratory in which it is expected that the subject will make the best observations of which he is capable, and conditions are arranged to facilitate them. This method yields literal perception. The average

errors for color, size, and form are of known small amounts. The second is the method of impoverished, ambiguous, or equivocal stimulation in which the experimenter arranges conditions so that errors will occur (40). The experimenter uses dim light, or a flash exposure, or presents more items than are perceptible at once, or makes his test after an interval of time. He may present undifferentiated or nonsense configurations, such as ink blots, which have unusual shapes and suggest many different interpretations. Or he presents reversible figures or patterns with equivocal contours which have more than one meaning. This method yields schematic perception. Impoverished, ambiguous, or equivocal stimulation is necessary because the observer in a psychological experiment, if he is an adult, is usually wary. He will look with attention and perceive literally when conditions permit. The fact is that the evidence for the schematic trend in perception has all been obtained either by this method of presentation or with relaxed attention on the part of the observer. The alterations or distortions might have been eliminated if the conditions for observation had been different.

It is true, to be sure, that the perceiving of everyday life is often a matter of glances and faint or ill-remembered impressions, and the results of impoverished or ambiguous presentations are therefore truer to life than the results of optimal presentations. The perception of everyday life *is* very often schematic. In common speech, a man tends to "see things in his own way." In the course of practical behavior, perception is no more literal with respect to color, size, shape, and sequence than is necessary, since literal perception takes time and effort. The percept is reduced to a cue for action. But perception can become literal whenever the observer needs to discriminate. Under favorable conditions it can be surprisingly exact, as the experiments of the laboratory demonstrate. One can always look at a thing carefully if there is reason to do so.

Perception is not always or necessarily distorted by needs or affected by purposes (18). It is not fated to be stereotyped or assimilated to social norms (97). Misperception is not a consequence of sensory organization but of the inattention of the perceiver or the weakness of physical stimulation. It is perfectly true that perception can be fluid, subjective, creative, and inexact, but it can also be literal. It can be literal with respect to fine differences and complex qualities, as the wine-taster, the artist and the scientific observer prove. The student of human nature and society needs to remember this when he is in danger of assuming that men are the passive victims of their stereotypes and perceptual customs. The detection of witches by the citizens of Salem, Massachusetts, is a case of gross misperception, but it does not always happen.

The world of visual experience which this book is about might now be defined as the literal visual world – the world of qualities as they appear to attentive observation. It is a world of color and space in combination; it is not the world of color-sensations, since that is a myth of seventeenth century psychology. It resembles the world of everyday experience more than traditional sensations ever did. But, on

the other hand, it is not simply the world of everyday behavior either, since that is a schematic world of cues and signs from which many of the qualities of color and space have dropped away. Only in an unfamiliar environment or a problem situation do we become fully aware of the literal visual world. One has to pause and look in order to see it.

The world of visual experience with which this single chapter deals is the schematic visual world. Its richness and complexity cannot be described in one chapter — or even in one volume. In a sense the real study of perception begins with it. Schematic perception, however, is an even more intricate human act than literal perception, and we cannot hope to discuss it sensibly without coming first to an agreement about fundamentals.[2]

In the thinking of many artists, for example, there exists a confusion between the seeing of space and the seeing of symbols. They are not clear where the difference lies between representing and symbolizing. They need to agree upon the first problem before they can undertake to deal with the second. Philosophers and logicians have been concerned for centuries over how to define true knowledge. They need to know what part sensory stimulation contributes to knowledge before they can discover where error creeps in. In the study of social psychology (and the other social sciences which rest upon it), there is an enthusiastic interest in perceptual customs, the social norms which make a man's world what it seems. Study the values of a culture, the formula has it, and you will understand why its members behave as they do. These enthusiasts need to be reminded that all human beings, everywhere, probably see the ground and the sky in the same way. They need to know the basic perceptual capacities of the human species before they can hope to describe the private worlds of persons, classes, races, or nations. It is important to understand that a poor child remembers a fifty-cent piece as much bigger than it really is. But this fact can never be comprehended without understanding how size is judged as a feature of literal space.

Conclusions

1. The classical formula of empiricism, that two-dimensional color-sensations are innate while all other perception depends on learning, fails completely to meet the facts.

2. The formula that space is innate but that meaning is learned meets more of the facts but it, too, is inadequate.

3. Meanings and spatial properties are not entirely separable from one another; meaning is not wholly detachable from color, form, and texture. Symbolic meanings, however, seem to be detachable from their objects and are presumably learned.

4. There is some evidence that, in animals and infants, embryonic meanings

[2] For a summary of the evidence about schematic perception and a consistent presentation of the theory that perception depends mainly on constructive processes within the organism rather than on stimulation, the reader is referred to Vernon's *Visual Perception* (115). Vernon's entire emphasis is on the degree to which percepts are *not* in correspondence with stimulation and are *not* referable to physical objects outside the observer.

do not have to be learned and consequently that all kinds of meaning are not learned.

5. There is overwhelming evidence to show that meanings react upon their perceptions to select or modify the spatial properties (color, size, outline) and that these properties therefore depend upon the personality and the culture of the perceiver. This evidence applies, however, to schematic perception, not to literal perception. The properties of the literal visual visual world, insofar as conditions are optimal, approach an exact correspondence with variables of physical stimulation.

6. The correspondence of perception to stimulation does not have to be a wholly innate correspondence. A psychophysical relation may be shown to exist without our having to decide how it came to exist. Conceivably every such relation is partly innate and partly acquired. Even the relation of color to wave length may be refined and made more exact by past experience with colors.

12

Learning

In What Sense is Behavior Mediated by Perception? . . . . In What Sense Do We Learn to See?

A generation ago psychologists were debating whether their principal concern was consciousness or behavior, and they called themselves introspectionists or behaviorists according to which side they favored. Today that dispute does not seem as important as it once did. Of much greater importance is an understanding of the process of learning in animals and men. A similar issue in a new form arises, however, which might be put in this way: is learned behavior mediated by perception or is perception only an incidental accompaniment of learned behavior? In other words do we adjust to the world because we see it or is our seeing of the world the result of our adjusting to it? Is learning a matter of insight or does insight follow upon learning? This issue is not merely a verbal dispute, for differing opinions yield quite different experiments. Neither is it trivial, for it involves a choice of the direction in which a science shall move. How is perception related to learning?

Two separate questions are implicit here which a theory of perception should try to answer. The first is, in what sense is learned behavior mediated by perception? The second is, in what sense is it true that one learns to perceive? It might be the case that perception is a prerequisite to learning in one sense and a result of learning in quite another sense.

The first question asks what the relation is between perception and performance. The investigators of learning in animals and children have tended to locate the learning process in the performance of their subjects. But the Gestalt theorists and others have tended to locate the learning process in perception. The former emphasize that learning is an alteration of behavior, the latter that learning is comprehension, cognition, expectation, or insight (55). The argument of the former is that behavior is what counts, and they are content merely to specify the "stimuli" or the "cues" which evoke behavior. The argument of the latter is that the perceived environment of an organism, the "field," determines its behavior, and that learning is best understood as a "reorganization" of the field. Learning is conceived to be the understanding of the values of things, of "what leads to what," of where things are, and of what to expect from a given event. Learning is thought to be perceptual and spatial or, in Tolman's words, to

consist of "cognitive maps" (110). The real issue between these theorists reduces to whether it is enough to say that stimuli evoke behavior or whether it is necessary to suppose that perception evokes behavior. Perhaps the issue can be resolved.

The second question asks what the relation is between perception and learning. Granting that our perception of the world is not simply constructed out of unlearned sensations plus images, in what sense do we learn to perceive space and meaning? The simple formula discussed in the previous chapter proved inadequate, and the question remains unanswered.

In What Sense is Behavior Mediated by Perception?

The term "stimulus" is used very loosely in the biological sciences, and the absence of agreement upon a definition has been at the root of much fruitless discussion.[1] In psychology the term is currently used in two different ways. It may mean a variation of energy to which there corresponds a variation of either an experience or a response, or it may mean the occasion for a namable experience or the occasion for an identifiable act. The first usage, the more precise one, is employed in psychophysical experiments; the second usage is employed in experiments on learning and behavior. There is a great difference between these two. The stimulus in a psychophysical experiment is only one instance of a systematically varied set of objects or events – an instance selected from a series of experimental variations of their physical properties made by an experimenter. The stimulus in a learning experiment is often taken to be simply an object or event – a junction of routes in a maze, a bell, a flash of light, a black box, or a printed word. Strictly speaking, these latter are not stimuli at all. They are usually called stimulus-objects or stimulus-situations. It was maintained in Chapter 5 that these terms cloak our ignorance. Behavior is, of course, a specific function of objects and situations, but we need to know why this is so. The stimulus-energies delivered to the receptors have been assumed in the past to be poor representations of the objects and situations to which they correspond.

The classical theory of perception as a special process of inner learning was invented to account for the discrepancy between the stimulus-energies and the perceived objects. So was the more recent theory of sensory organization propounded by the Gestalt psychologists. When a behaviorist, therefore, maintains that he need not concern himself with a rat's perceptions, when he denies that behavior is necessarily mediated by perception, and assumes that his rat responds to the alley of a maze as a "stimulus-object," he is glossing over a difficulty and concealing a problem. He may use the term "cue" instead, as Miller and Dollard have recently done in explaining Hull's theory of learning (83), but this expedient begs the ques-

[1] The definition given in Chapter 5 is worth repeating here: a type of variable physical energy, falling within certain limiting thresholds, which excites a receptor-cell differentially; also the differential excitation of different receptors with respect to the adjacent and successive order of differences over the array of cells.

tion. How do stimulus-energies become cues?

The situation is different if one assumes that the stimulus-energies delivered to the receptors are precise mathematical transformations of the objects and situations to which they correspond. Inner perceptual inference and sensory organization become gratuitous. Behavior is a specific function of ordinal stimulation. The definition of "stimulus" must be revised, it is true, and the convenient notion that a plurality of discharging receptor-cells represents a plurality of stimuli must be wholly rejected. But it becomes possible to refer to the stimulus-correlate of an object and to understand how a response can be a constant function of an object.

The kind of perception with which this book deals – spatial perception or literal perception – is not taken to be a special mental process. It is not something intermediate between stimulation and response. Both perception and behavior may be activities of the organism specific to ordinal stimulation. Behavior, then, is evoked or mediated by perception only in the special sense in which perception is the study of ordinal stimulation. The so-called stimulus-objects for behavior are the stimulus-correlates of the literal visual world. The so-called cues for behavior are certain invariants of stimulation which yield objects with color-constancy, shape-constancy, and size-constancy. Physical objects must be specific in stimulation if they are to be specifically responded to, but it does not follow that they must necessarily be known if they are to be specifically responded to.

The theory that learning is a perceptual process is mistaken if it implies that the only way to be sure an animal has learned is to intuit the animal's perception. It is useful to suppose that a rat has insights and hypotheses and expectations, that he makes simple inferences, and that he remembers a map of a maze, but there is no compulsion to suppose this. Sign-learning may be described as the accrual of meaning to a percept, and it may also be described as an altered tendency to react to the stimulus-correlate of an object. Which formula is more revealing we do not yet know.

In What Sense Do We Learn to See?

The 19th century controversy over nativism and empiricism in perception, discussed in Chapter 2, proved to be indecisive. The Gestalt theory of sensory organization did not clarify the issue, for its critics called it a kind of nativism, but its adherents denied that it was. The formula advanced in the preceding chapter was inadequate, inasmuch as the perception of literal space is not wholly unlearned and the perception of schematic meaning is not wholly learned. The simple solutions fail. Nevertheless there must be some respects in which perceiving develops spontaneously and other respects in which it depends upon practice, experience, or training. The difficulty is to state in what sense we learn to see.

The First Perceptions of Cataract Patients Blind from Birth. The method of determining whether or not a function is learned is to deprive a young individual of all opportunity to exercise the function, at the same time providing him with all the

conditions of life and growth which might enable the function to mature. The function of vision is much too valuable in the human subject to allow such experimentation, but the circumstances required for the experiment sometimes occur naturally. A few individuals develop cataracts in both eyes at, or soon after, birth. In this disease the lens becomes semi-opaque, so that the individual is, for all practical purposes, blind. Some diffused light can enter the eye, and such patients can usually distinguish night from day, but their impression of the world is at best no more than a vague gray fog, brighter at some times and darker at others, or possibly brighter in one region and darker in another when facing strongly contrasting surfaces. When the patient is a child or an adult one or both cataracts can be surgically removed, and after this happens the eye can form a differentiated image on the retina (although glasses must be worn to bring the image into perfect focus). After the bandages are removed, the operated patient is stimulated for the first time by a projection of his environment. The question is, what does he see?

The study of these first perceptions has interested scientists and physicians for many years. Sixty-six cases have been collected by von Senden and the facts brought together in one volume (94). Although the tests of perception were imperfect, and although the reports were obscure and sometimes contradictory, the snatches of evidence they yield are illuminating.

The evidence consists of what the patients said when they opened their eyes, their answers to the doctor's questions, descriptions of what they could do, and their responses to various simple tests. The foremost fact is that the patients were bewildered and confused by the new visual impressions, especially by the continuous, unrelenting flow of these experiences and by their enormous variety. It must be remembered that, in common-sense terms, they did not know what anything *looked like*. They knew only what things *felt like* – and only those things which could be touched. A subject would complain that there were "so many new things he could not comprehend." On a city street, "so many different things, and the rapid movement of the mass, confused his sight to such an extent that finally he could no longer see anything, the latest thing having not yet faded when the next impressed itself" (94, page 148). The patients all had the use of language but they found it difficult or impossible to describe what they saw or to apply words to it. The question, "Are things projected in space?" simply did not mean anything to them.

The use of visual perception seemed to require an extended period of development – weeks, months, or even longer. Too much seems to have been expected of the patients by relatives and friends, who could not understand why they were not now normal persons, and the task of learning to see sometimes overwhelmed them with its difficulty. Some, discouraged and depressed, quit trying to see. These unfortunates were treated by establishing for them a rigid program of exercises and drill in perceiving. The conclusion of one investigator was that "to give sight to a blind-born person is more the task of an educator than a surgeon" (page 145).

Nevertheless there is evidence of spontaneous unlearned visual capacity in these patients. Fixation of objects and the scanning of the environment was possible from the outset. The patients could follow a moving object with their eyes and head. Grasping an object, in which the hand is guided by the visual images of the hand and the object (as contrasted with the groping of the blind), seemed to be possible, although the act was slow and unsure. Watching the movements of the hands in front of the eyes was reported to be an action of great apparent interest to them. Looking at brightly colored surfaces was reported to be pleasurable, although the colors could not, of course, be named.

Most of the investigators had preconceived ideas of what the patients might see. Believing as they did in sensations of color and perceptions of space, they assumed that there were three alternatives: the observer might see things at the eye, that is, touching it, or he might see things in a flat plane in front of the eye, or else he must see space with things in it. To the eager questions about such perceptions the patient could not respond. He did not understand them, although he might fall in with the unconscious coaching of the questioner. On reading the answers now, the obvious conclusion is that the patient saw none of these possibilities. There are indications that he saw more nearly a visual world than a flat visual field, but something less determinate and less specific than the literal world which we see.

The reports make it obvious, again and again, that the newly seeing person had no words for the features of his environment. The patient could not assign to his impressions terms like black and white, moving or still, far or near. He could not immediately say whether there were two or three black spots on a piece of paper. He could not even point on request to an edge or a corner. He could not, in fact, say anything about his visual impressions. *He had these terms in his vocabulary but they referred to tactual and muscular impressions only.* Although this inability to describe anything made it almost impossible to discover what the patient saw, his behavior indicated that he did have visual impressions and that some of them differed from others.

There are perfectly clear cases in which the patient could use the words "same" or "different" (94, p. 155-157). For example, two strips of cardboard, 10 cm. and 20 cm. long, were seen as different but the patient could not say that one was "longer." That word meant something quite different from what it does to us. A silver pencil and a large key appeared different as they lay side by side on the table, but they could not be identified or named. When they were put in the patient's hand, however, they could be named at once. This happened with other pairs of objects such as knives and forks, or cubes and spheres. They could be told apart but not recognized by sight, even though they were things which could be recognized when felt. At the outset, the patient could not even say longer or shorter, curved or straight, square or round, thick or thin, wide or narrow, much less words like cardboard, key, or fork. He could react to the difference, apparently, before he was

able either to specify the difference or to name the objects.

The histories suggest that some comparative terms were applied to the visual world much more rapidly than others. Colors were soon named; visual motion was quickly called what it is; large and small together with far and near seemed to be used appropriately without much delay. The chief difficulty seemed to come in learning to use the innumerable terms for comparing and naming visual shapes. Perhaps this is not surprising in view of the subtlety and variety of those spatial features by which we distinguish objects from one another. According to the accounts it took many weeks or months to be able to name common shapes, that is, to identify the objects, places, events, people, and animals of the environment, and still longer to learn the signs and symbols. The patient had to undertake the task of revising radically the meanings of the words he used.

Kinds of Visual Learning. Since a person who sees for the first time literally does not know what anything looks like, although he knows what anything he could touch feels like, he must learn. A considerable part of his task is to name the new visual impressions using the words previously aroused by the old tactile and muscular impressions. This kind of naming is analogous to what is called paired-associates learning in the psychological laboratory when new cue-items are substituted for old ones. Such experiments would imply that transfer of learning should occur (47). It should be noted, however, that visual stimuli and tactile-muscular stimuli are dissimilar and should therefore manifest a minimum of stimulus-generalization. There is nothing to show in any of the cases described that the two modes of sensory impressions for the same object were in the least connected at the beginning. There was no evidence that an object intrinsically looks the way it feels. One might suppose, for instance, that the visual and the kinesthetic quality of "vertical" are originally alike. A case is described by von Senden (page 158) in which the patient was asked, some weeks after the operation, to identify a horizontal and a vertical line drawn on a card. He could do so only after tracing them with his finger. Presumably he had learned by this time to trace, that is, to correlate the sight and the feeling of his moving finger. This phenomenon will be further considered in the next chapter.

Another task faced by the patient is the naming of new visual impressions which, unlike manipulable objects, never had any tactile-muscular counterparts, for instance, the many kinds of forms, places, events, signs and symbols. Such impressions are very numerous and many of them are quite similar to one another. An extreme example would be learning faces or, more precisely, learning to identify people by the shape of their faces. With this kind of learning all the operated patients had extreme difficulty and made slow progress. There are analogous experiments on identifying nonsense forms (34), learning a code (62), and recognizing aircraft (39). It could be predicted from these experiments, and from E. Gibson's theory based on "intra-list generalization" (33), that the more items there are to be identified, and the more similar they are to one

another, the more difficult will be the learning. It would also be predicted that there should occur mis-identification of similar items, and that the learning of new items should cause forgetting of similar old items. All these phenomena are reported in the learning of the cataract patients.

It is obvious that these individuals had to learn to see the world. Putting words to their impressions had to be learned, and perhaps this is no small part of seeing. The accounts do not suggest, however, that surfaces, edges, slants, and shapes — that is to say space — were at first invisible and later became visible. They suggest, instead, that these variables were not at first, or not completely, identifiable. The impressions were, however, discriminable, and they might have been demonstrated to be such had the investigators used more ingenuity in testing.

Unlearned Visual Identifications in Animals. Things, places, and events may be identified without necessarily being named. Animals other than man are limited to this kind of identification. When an organism responds in some unique way to a fact of the physical environment it may be said to have made an identifying response. When this reaction depends on a retinal image it is a visual identification. Naming things is only one type of visual identification, although an extremely important one for human learning. Our hypothetical pre-human ancestor who could pursue zebras but fled from sabre-toothed tigers was identifying things.

Instinctive behavior involves not only unlearned motor activities but unlearned identifications. Birds can identify egg-like objects at the first nesting season inasmuch as they retrieve and care for such objects only (page 207). Geese show fear at the sight of a moving hawk-like silhouette in the sky but not at an otherwise identical goose-like silhouette (107). If Spitz is correct, human infants retain a trace of instinctive behavior in their tendency to smile at face-like objects and at these only (page 208). It would appear, then, that identifying responses are not necessarily learned in their entirety, possibly not even by men. In the terminology of the previous chapter, there may be embryonic meanings in the first visual impressions.

The question of whether or not space-perception in animals is learned is a poor question. What can be asked sensibly is whether animals with no opportunity for practice will react appropriately to spatial situations at the first occasion. One of the best of such experiments is Lashley and Russell's study of the jumping behavior of rats reared in complete darkness (72). Thirteen rats were allowed to grow to maturity without seeing even the walls of their box except for a few seconds every other day when food was inserted. At 100 days of age each rat was brought into the light and placed on a high platform with a gap between this and another platform containing food. The rat was first allowed to step across a 5 cm. gap five times, and then was confronted with a 20 cm. gap. Twelve jumped the distance successfully on the first trial, and all could do so very consistently over ten trials.

Subsequently each rat was given three trials at 40 cm., three at 20, and three

more at 40. Most of them failed at 40 cm. or refused to jump, but the significant fact is that they increased the force of their jumps with the increased distance. The force of the jump was recorded automatically by the jumping platform. Finally they were allowed nine jumps with the landing platform at varying distances. The force proved to be graded in proportion to the distance. These rats could not jump as successfully as normal rats reared in the light (they fell more often) but the gradations of force were nearly as accurate. The implication is that they could discriminate the distance nearly as well.

In the terminology of this book the rats were probably responding to depth-at-an-edge, which is the essential feature of a gap in the floor. Not only were they reacting differentially to different abrupt changes in optical stimulus gradients: they were with some success identifying a fact of the physical environment by an appropriate act. The identifying act seems to have been innate.

Other experiments with animals reared in darkness are being continued in Lashley's laboratory using chimpanzees, whose perceptions are probably more like man's. The first tests for visual perception indicated that these animals were blind (90), or at least that they failed to use vision in guiding their behavior, but no conclusions can be drawn since it is possible that the chimpanzee's retina does not mature normally in darkness.[2]

[2]Since this passage was written, the writer has been informed by Dr. A. Riesen that this explanation is indeed the most probable one. The visual defect seems to be due to abnormal maturation, and it cannot be ascribed wholly to the absence of learning.

The Study of Visual Identification. The nature of identifying responses, including naming, is not very well understood. It is becoming clear, however, that the discriminating of stimulus-variables is an essential component of all learning. An organism cannot learn a reaction without identifying the cue for it. The object must be reacted to as the same object on different days. Similar objects and events are said to be equivalent for behavior unless the organism discriminates among them; a stimulus to which a response has been conditioned is said to show generalization in that the response will occur to a whole class of stimuli.

The ordinary discrimination experiment isolates a pair of objects or situations differing in only one respect. The psychophysical experiment in its simplest form does the same thing. The former elicits differential reactions while the latter elicits a discriminative judgment of "more" or "less," but there is a basic similarity between them. Objects and events in daily life, however, do not come in neatly controlled pairs. In order to understand the nature of cues and of learning a different kind of experiment is necessary. It would be an experiment which requires the subject to react in a unique way to each of a whole set of objects or events, more or less similar. These items should differ in many dimensions of variation and should fall into classes and sub-classes. Attention is centered on the errors of identification and how they are eliminated. As a type it might be called the identification experiment. Research of this type is beginning to appear (34, 62, 39).

Learning to See. Considering the evidence, in what sense do we learn to see and in what sense do we not? Clearly, we do not learn to accommodate, converge, fixate, and move the eyes, although it is very possible that practice improves these functions. We do not, in other words, learn to produce optimal visual images of speckled light having borders, cycles, shades, and gradients of intensity and frequency-mixtures. We do not learn the space-values of separate retinal points. We do not learn to associate retinal points so as to see form. We do not learn to interpret color and form sensations so as to see the third dimension. What we do learn preeminently is to identify the features of visual stimulation which correspond to the important features of the physical environment.

Words facilitate but are not essential to this process of identification. Whatever it may prove to be in detail, it is sure to involve the discriminating of complex variables discoverable in retinal images. There must occur, in other words, a difference in response along with a difference in stimulation – either a discriminative reaction or a discriminative judgment. The discriminating of variables is necessary for the identifying of things, grading into one another as they do and innumerable as they are.

As things become identifiable, and as we learn to notice the differences between them, our perceptions of the world become differentiated. Formerly indefinite qualities become definite. Shapes and textures and surfaces and colors become specific. Indeterminate movements, locations, sizes, and distances become determinable. Properties like inside and outside, congruency, symmetry, opposition, and continuity are elaborated. Objects, events, and situations are recognized. In the case of human beings, things are named. The qualities of, or differences between, objects are also named. This enables us to name classes of objects. Once this process is started it builds upon itself; new differences emerge, new similarities become visible, and more general classes are named. At the same time more and more objects are identified. The traditional way of putting it is to say that things have meaning and that we have abstract ideas about them. But the progress of learning is from indefinite to definite, not from sensation to perception. We do not learn to have percepts but to differentiate them. It is this sense in which we learn to see.

13

Spatial Perception and Spatial Behavior

The Motor Theory of Space-Perception . . . . Visual Kinesthesis . . . . The Co-variation of Visual and Muscular Motion . . . . The Ego in Perception . . . . The Impression of Distance from Here to There . . . . Orientation

Ever since Berkeley's *New Theory of Vision* in 1709, it has appeared plausible that the seeing of space depends, in some fundamental way, on exploring and manipulating the environment. Seeing things, Berkeley argued, could always be verified by touching things, and hence it was possible that the solidity and depth of the visual world were originally not visible but only tangible. Vision might get its spatial character from the tactile and muscular impressions which always accompany it. We learn to trust our vision of the table as being there, for instance, because we can always go over and touch it.

The Motor Theory of Space Perception

We know that the infant and young child ceaselessly explores his environment as his vision develops. Is it not likely that his visual impressions get their solidity and depth from their association with these movements? This is the argument for a motor theory of perception.

This theory has such great persuasiveness that it must embody a truth even if it does not state one. Why does it seem to be consistent with the observations of so many people? Introspectively, it is a fact that seeing is almost inseparable from acting. Spatial behavior is intimately connected with spatial perception. Things do look capable of being grasped, or pushed, or touched on all sides, or of resisting these actions. The floor does look capable of being walked on whereas the walls do not, and neither does a gap or a hole in the floor. An edge looks capable of being traced with the finger. A slant looks as if you could angle the palm of your hand to it, or climb it. Thus, the visual world has the appearance of inviting many kinds of behavior.

It is one thing to say that the visual world has motor meanings, however, and quite another to maintain that it gets its spatial qualities from these meanings. It is possible to agree that we can almost feel the visual scene without concluding that we do not really see it but only remember how it feels. Perhaps what is wrong with a motor theory of perception is its one-sidedness. The infant explores

his environment and feels the things he sees, true enough, but he also sees the things he feels. If visual impressions acquire motor meanings, is it not just as inevitable that motor impressions should acquire visual meanings? Perhaps the trouble with a motor theory of space perception is that it needs to be supplemented with a visual theory of muscular impressions. For every seeing individual there is a co-variation of retinal stimulation and muscular-tactile stimulation during behavior. Neither kind has to be taken as primary.

Visual Kinesthesis

The term kinesthesis, meaning sensitivity to motion, is usually applied to the muscle-sense. Actually, there are receptors in the muscles, tendons, and joints which yield impressions of both the motion and position of our limbs. There are also receptors in the inner ear which yield impressions of the motion and position of our head and body. Moreover, the touch receptors anywhere on the skin can yield impressions of motion, and those in the hands and feet notably do so when we manipulate things and walk about. Impressions of our own movements are known to depend on all these sources, but there is another and usually unrecognized source in the retina. The reasons for using the term visual kinesthesis have already been given in Chapter 7. A tabulation of the kinds of retinal motion and their correspondence to the movements of external objects and of the body is given on page 132. The point to be emphasized is that there are normally retinal stimuli which are precisely co-ordinated with both locomotion and manipulation.

In cases of what is called locomotor ataxia, the patient has no kinesthetic sense in the lower part of his body. Damage to nerve centers in the spinal cord has blocked the transmission of impulses from the receptors in the muscles and joints and from the soles of the feet. He cannot feel the position or motion of his legs. He walks with a very peculiar gait, as if he had to throw his legs forward, and when he is blindfolded he cannot walk at all, or even stand up. The thing to note, however, is that he *can* stand and walk, after a fashion, if he looks at the ground and his feet while he does so. Without bodily sensitivity the use of the legs is impossible unless visual sensitivity fulfills the necessary functions. It would seem that there are two forms of the kinesthetic sense; one based on the known list of proprioceptors, and one based on vision.

The term sense is almost as misleading as the concept of sensation. Kinesthesis is not one of the senses; there are not just so and so many departments of sense. Kinesthesis is mediated by a number of types of receptors. It probably should be conceived very broadly, but its basic function is to adjust and set the pace for muscular action. It guides manipulation and the using of tools. It also guides locomotion of the ordinary sort which is driven by muscular action. In a more complex form it guides locomotion of the sort driven by gasoline engines, propellers, and jets, and in the case of the automobile and the airplane, locomotion requires complex manipulation. In such

advanced types of spatial behavior, visual kinesthesis plays an essential role.

The Co-variation of Visual and Muscular Motion

In as apparently simple an act as picking up a pencil both visual and muscular impressions are involved in the control of the performance. The contracting image of the hand projected on the retina as it coincides with the image of the pencil is paralleled by the feeling of the hand being extended and touching the pencil. This is called an eye-hand co-ordination. The retinal motion and the muscular impression vary together, both being controlling stimuli for the pace or flow of the act.

Animals stimulate themselves as they act and this stimulation affects the action. The process is circular; it has recently been compared to the feedback mechanism of electronics (120). In the case of spatial behavior a visual component must be added to the circle, so that the complete process is like that diagrammed in Figure 80.

The diagram helps to explain why visual impressions and motor impressions imply each other when we introspect on our experience. Retinal motion is automatically linked to bodily action from birth onward, so long as the eyes are open and there is light to see by. Bishop Berkeley was correct in asserting that to walk over and touch the table is to confirm visual space; what he did not understand is that the expanding visual field also confirms muscular-tactile space.

Spatial behavior and spatial perception are coordinate with one another. We have neither to see space before we can behave nor to make spatial responses before we can see. A man who gets about his environment without collisions and manipulates objects to the satisfaction of his wants is exercising a function in which his behavior is spatial and his space is behavioral.

The Ego in Perception

Perceiving the world has an obverse aspect, perceiving oneself. Observers have often pointed out that one's own

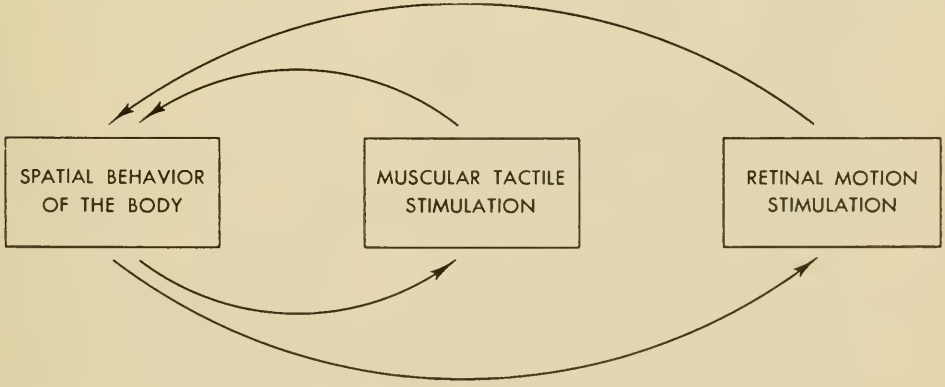


FIGURE 80. The Coordination of Spatial Impressions and Spatial Behavior

body is represented in the visual field, and it has been argued that the ego is therefore an object in the field of experience like any other object (67, page 319-331). A more satisfactory statement, however, is that perceiving the environment includes the ego as part of the total process. In order to localize any object there must be a point of reference. An impression of "there" implies an impression of "here", and neither could exist without the other.

The definition of the ego is a problem with which psychologists and philosophers have struggled without much success. The concept of a self, by whatever term it is called, is necessary for any scientific theory of personality, of social behavior, of abnormal behavior, or of ethical behavior. A clarification of the social ego (98) might be possible if a few solid facts could be established about the biological ego, that is, the ego manifested in maintaining equilibrium and posture, in locomotion and manipulation, and in literal visual perception.

What are the stimulus-correlates for the perception of oneself? There must exist multiple correlates, or what used to be called "co-operation of the senses," rather than a single correlate. Tentatively, the following can be listed:

1. The tensions of the skeletal muscles which maintain equilibrium and regulate the posture of the legs, trunk, head, and eyes.

2. The contact stimulation of the skin against the surfaces on which the body rests.

3. The movement stimulation aroused by action of the muscles and joints, and also by acceleration of the head affecting the inner ear.

4. The movement stimulation aroused by changing contact of the skin with supporting surfaces and resisting surfaces.

5. The so-called "boundary of the visual field," i.e., the peripheral retinal images of the nose and other parts of the body. This includes images of the hands and feet which protrude into the visual field from its lower margin.

6. The retinal displacement of these bounding images. They shift in a specific way when the eyes turn; they also shift in a different way when the head turns. The shift is "concomitant and reciprocal" with the turn.

7. The deformation of the whole retinal image. The visual field expands when one goes forward and contracts when one goes backward. The maximum velocity of this motion is reached at the boundary – at the images corresponding to the parts of the body.

8. Finally, several types of more or less continuous organic stimulation known generally as *somaesthesia*, the self-stimulation involved in breathing for instance, and also in eating, drinking, and sexual activity, together with the stimulation underlying bodily discomfort or comfort.

These are at least some of the forms of stimulation which might yield a primitive ego in perception. The notable thing about them is that they all co-vary with action. The process is circular. An overt performance is controlled by various types of kinesthetic stimulation, including the visual. A posture, on which any overt performance is based, is controlled by stimuli of the first seven types listed

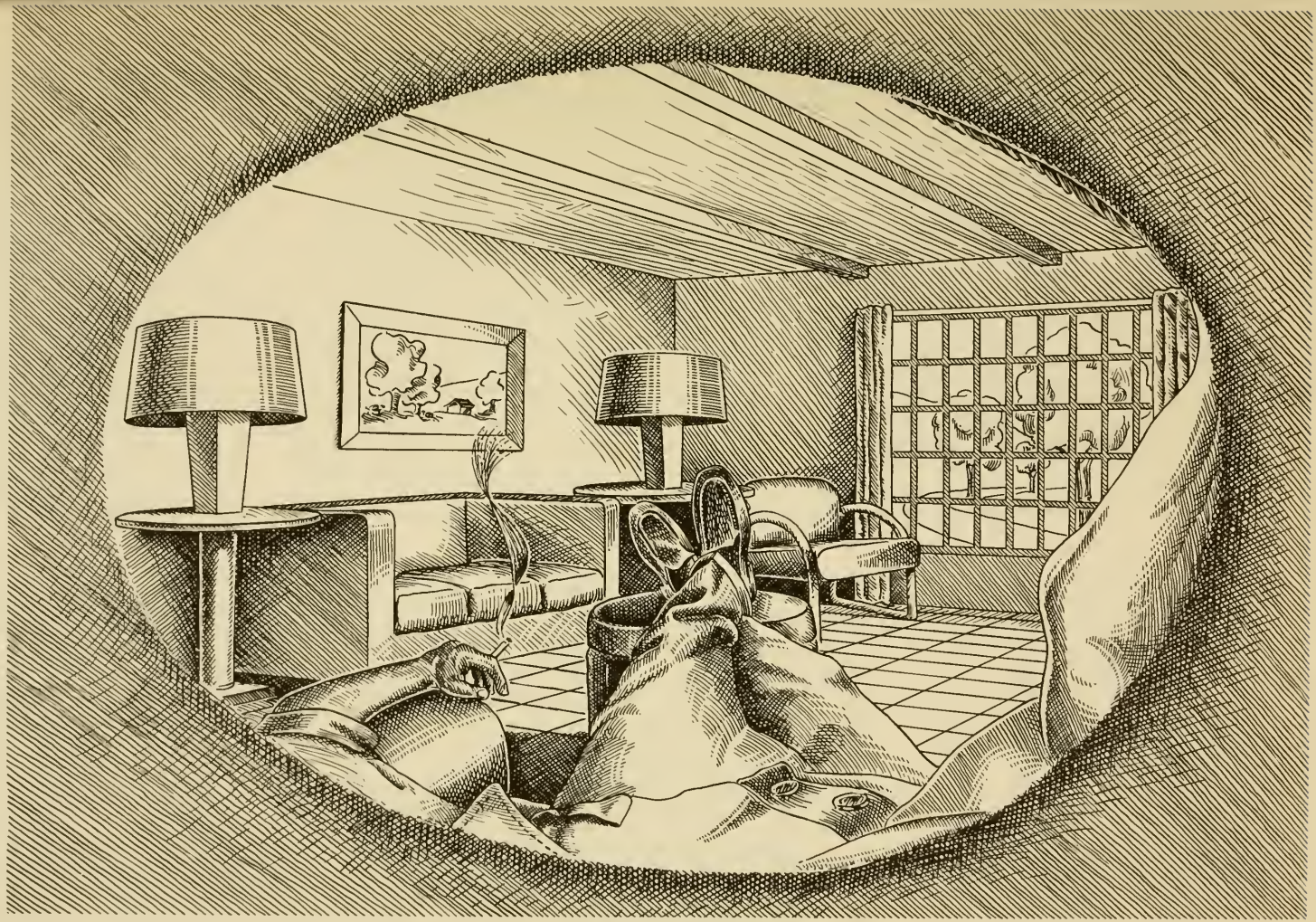


FIGURE 81. The Modern Visual Ego

This drawing, like the one made by Ernst Mach in the 1880's (Fig. 5, ch. 3) represents the visual field of the left eye. The nose, lip, and part of the cheek are visible on the right; the body lies at the bottom of the field; and the environment appears "out there".

above. A motive or need, from which any performance takes its origin, is controlled by stimulation of the eighth type. Whatever else the ego may be, it includes elements of performance, posture, and need. Perhaps these are the fundamental components of the biological ego – the sense of one's own body and the impression of something that is uniquely and continually "here."

The Impression of Distance from Here to There

The central hypothesis of this book asserts that space is constituted of the same variables as things. It holds that surfaces and margins are what we see, not air. Space must be filled to be visible; empty space is an abstraction. Against this hypothesis the objection can be made that we *do* have visual impressions of empty space and we have them all the time. We see the distance from here to there; we "look out" upon the world, and the space between one's eyes and the nearest object is plainly empty. The objection is convincing, but it can be overcome.

One kind of empty space, the distance between an object and the surface behind it, has already been accounted for in terms of depth-at-a-contour. The visual superposition or overlapping of surfaces, it was argued, is an important type of depth-perception, not a cue for depth-perception. The explanation was found in steps, as contrasted with gradients, of the main stimulus-variables for distance – a step in the density of texture, a step in the rate of deformation of texture, and a step in the binocular disparity of texture. The steps are proportional in amount to the physical difference in depth between the two surfaces in question.

The same kind of reasoning may be extended to the impression of the distance from here to there – the space between oneself and the ground or the distance to the wall. The boundary of the visual field, particularly the image of the nose, incorporates all three of these abrupt steps in ordinal stimulation. The nose is given as a very large area in the visual field. (In the writer's case, the line from the tip to the nostril coincides with a foot rule held at arm's length, but he may be especially well endowed.) It is also a crossed double-image. Moreover the edge of the nose has a marked visual velocity when the head moves or turns. The nose, in fact, projects the maximum possible degree of crossed retinal disparity and should, therefore, according to the present theory, represent the greatest possible degree of visual nearness (see page 106). Likewise, at the margin formed between the nose and external reflecting surfaces, the greatest possible velocity of parallactic motion is reached in the retinal image, and this also should represent the maximum impression of nearness. *These facts establish an absolute zero of distance in stimulation and this, in turn, makes possible an absolute scale of distance in experience.*

There is nothing mysterious, then, in the impression that a person "looks out" upon the world, that he sees distance or space continuously extending from "here"

to "there" and that he, himself, is "here." This impression is in correspondence with retinal stimulation. The perception of the world and the perception of oneself are both, figuratively speaking, cut from the same cloth.[1]

There is no necessity for the assumption that a person projects his sensations outward from the eyes. All the ingenuity that has been devoted to explaining such an event has been wasted. The steps of distance at the edges of things and the great first step of distance at the windows of the eyes themselves are all to be found within the eyes.

Orientation

All organisms from the lowest to the highest can orient themselves to certain forces of the physical environment. Plants can orient to the direction of gravity and the direction of light. Animals possessing vision and locomotion react to finer variations and can orient themselves, not only to the sun and the earth, but to things. The human species has a still more differentiated environment and can orient to very complex and unique features of it. The human habitat, consisting of rooms, houses, streets, cities, and countries, contains millions of things to which a man can find his way – a cigarette, a cup of tea, a tax receipt, a racetrack, or a sweetheart in Texas.

The study of locomotion is in its infancy. Although a good deal is known about the movements of walking, swimming, and flying, surprisingly little is known about how animals and men find their way about (88). The explanation is probably that we do not fully understand the process of orientation. There is one special form of locomotion, it is true, which has been very thoroughly investigated – the behavior of white rats in running through mazes. But maze-learning seems to be a very complicated sort of behavior, and, it does not represent a simple form of locomotion. How the rat comes to be oriented to the goal-box of a maze is a disputed issue, as the last chapter indicated (58, 109).

How, exactly, is oriented locomotion to be defined? What would be a fundamental experiment on the process of getting about? Lewin and his students (74) made a start on these questions fifteen years ago, but they became so interested in using locomotion as an analogy for higher forms of behavior that they never got down to studying the literal process. The simplest kind to study would probably be the act of going to a visible object or place, the goal. Locomotion of this sort is oriented directly toward the goal. The body movement is a function of optical stimulation which yields the perception of a visual world with the goal-object in it. Body movement is modified only by the necessity of avoiding obstacles, or direct-

[1] So important is the impression of being "here" to having a sense of self, and so important is the visual image of the nose to the impression of being "here," that nose-perception must be a prominent factor in the awareness of the ego. There have been a few students of nose-perception such as Cyrano and Durante, and they have clearly understood the relation between a man and his nose, but the subject deserves more investigation.

ing the movement into the field of safe travel (41).

A more advanced form of locomotion would be the act of going to an object or place beyond the range of vision. This might be called the destination. "Being oriented," in the popular sense of the term, refers to this ability. In common sense terms, one must know both where he is going and where he is now. It requires, over and above the visual world, a frame of reference (93) or a topographical schema (46). The individual must perceive the space which surrounds him on all sides, as described in Chapter 3, and must also apprehend the world beyond the visible scene – the layout of the building, of the city and its streets, of the region, and of the country with its highways and cities. He is then said to be oriented in space – actually, in a series of more and more inclusive spaces of which the most general is the astronomical universe.

The conception of an objective world, independent of the standpoint of any observer, rests upon this type of orientation. The ability to take the position of another person, to see from his point of view, depends on being oriented in space. Orientation is inseparable from locomotion, for, only because an observer gets a different visual field at every different standpoint, does he perceive a single integrated world (Chapter 8). That is, because the visual field changes systematically with change in the position of the head, an ordinal network is established among the diverse fields yielding a visual world independent of any standpoint. The facts of spatial perception and of spatial behavior are united in the fact of the visual ego.

References

1. Ames, A. (1946), "Binocular Vision as Affected by Relations Between Uniocular Stimulus-Patterns in Commonplace Environments," *Amer. J. Psychol.*, 59, pp. 333-357.

2. Asch, S.E., and Witkin, H.A. (1948), "Studies in Space-orientation. I Perception of the Upright With Displaced Visual Fields," *J. Exper. Psychol.*, 38, pp. 325-337.

3. Asch, S.E., and Witkin, H.A. (1948), "Studies in Space Orientation. II Perception of the Upright with Displaced Visual Fields and with Body Tilted," *J. Exper. Psychol.*, 38, pp. 455-477.

4. Asch, S.E., and Witkin, H.A. (1948), "Studies in Space Orientation. III Perception of the Upright in the Absence of a Visual Field," *J. Exper. Psychol.*, 38, pp. 603-614.

5. Bartlett, F.C. (1932), *Remembering, A Study in Experimental and Social Psychology*. Cambridge: Cambridge Univ. Press.

6. Bartley, S.H. (1941), *Vision, A Study of Its Basis*. New York: Van Nostrand.

7. Bassett, M.F., and Worne, C.J. (1919), "On Lapse of Verbal Meaning With Repetition," *Amer. J. Psychol.*, 30, pp. 415-418.

8. Bender, M.D., and Teuber, H.L. (1949), "Disturbances in Visual Perception Following Cerebral Lesions," *Jour. Physiol.*, 28, pp. 223-233.

9. Berkeley, G. (1922), *A New Theory of Vision and Other Writings*. (Everymans Library ed.) New York: Dutton.

10. Boring, E.G. (1933), *The Physical Dimensions of Consciousness*. New York: Century.

11. Boring, E.G. (1935), "The Relation of the Attributes of Sensation to the Dimensions of the Stimulus," *Philos. Sci.*, 2, pp. 236-245.

12. Boring, E.G. (1942), *Sensation and Perception in the History of Experimental Psychology*. New York: Appleton-Century.

13. Bourdon, B. (1902), *La Perception Visuelle de l'Espace*. Paris: Schleicher Freres.

14. Bruner, J.S., and Goodman, C.C. (1947), "Value and Need As Organizing Factors in Perception," *J. Abn. Soc. Psychol.*, 42, pp. 33-44.

15. Brunswik, E. (1944), "Distal Focusing of Perception: Size-Constancy in a Representative Sample of Situations," *Psychol. Monog.*, 56, No. 254, pp. 49.

16. Buhler, K. (1922), *Handbuch der Psychologie. I Teil. Die Struktur der Wahrnehmungen. I Heft. Die Erscheinungsweisen der Farben*. Jena.

17. Buswell, G.T. (1935), *How People Look at Pictures.* Chicago: Univ. Chicago Press.

18. Cantril, H. (1948), "The Nature of Social Perception," *Trans. N.Y. Acad. Sci.*, 10, pp. 142-153.

19. Carr, H.A. (1935), *An Introduction to Space Perception.* New York: Longmans Green.

20. Carter, L.F. and Schooler, K. (1949), "Value, Need, and Other Factors in Perception," *Psychol. Rev.*, 56, pp. 200-207.

21. Cassirer, E. (1944), "The Concept of Group and the Theory of Perception, *Philos. and Phenom. Res.*, 5, pp. 1-35.

22. Child, C.M. (1941), *Patterns and Problems of Development.* Chicago: Univ. Chicago Press.

23. Collins, F.T. (1922), *Arboreal Life and the Evolution of the Human Eye.* Philadelphia: Lea and Febiger.

24. Conant, J.B. (1947), "On Understanding Science," *Amer. Scientist*, 35, pp. 33-55.

25. Courant, R., and Robbins, H. (1941), *What Is Mathematics?* New York: Oxford Univ. Press.

26. Dodge, R. (1902), "Five Types of Eye Movement," *Amer. J. Physiol.*, 8, pp. 307-329.

27. Don, V.J., and Weld, H.P. (1924), "Lapse of Meaning With Visual Fixation," *Amer. J. Psychol.*, 35, pp. 446-450.

28. Duncker, K. (1929), "Über -Induzierte Bewegung. Ein Beiträg zur Theorie Optischer Wahrgenommener Bewegung," *Psychol. Forsch.*, 12, pp. 180-259.

29. Duncker, K. (1939), "The Influence of Past Experience Upon Perceptual Properties," *Amer. J. Psychol.*, 52, pp. 255-265.

30. Duke-Elder, W.S. (1938), *Textbook of Ophthalmology*, Vol. II. London: Henry Kimpton.

31. Eisenstein, S. (1942), *The Film Sense* (Tr. by J. Leyda) New York: Harcourt Brace.

32. Ellis, W.D. (1938), *A Source Book of Gestalt Psychology.* New York: Harcourt Brace.

33. Gibson, E.J. (1949), "A Systematic Application of the Concepts of Generalization and Differentiation to Verbal Learning," *Psychol. Rev.*, 47, pp. 196-229.

34. Gibson, E.J. (1942), "Intra-List Generalization As A Factor in Verbal Learning," *J. Exper. Psychol.*, 30, pp. 185-200.

35. Gibson, J.J. (1929), "The Reproduction of Visually Perceived Forms," *J. Exper. Psychol.*, 12, pp. 1-39.

36. Gibson, J.J. (1933), "Adaptation, Aftereffect, and Contrast in the Perception of Curved Lines," *J. Exper. Psychol.*, 16, pp. 1-31.

37. Gibson, J.J. (1937), "Adaptation, Aftereffect, and Contrast in the Perception of Tilted Lines. II Simultaneous Contrast and the Areal Restriction of the Aftereffect, *J. Exper. Psychol.*," 20, pp. 553-569.

38. Gibson, J.J. (1937), "Adaptation With Negative After-Effect," *Psychol. Rev.*, 44, pp. 222-244.

39. Gibson, J.J. Ed. (1947), *Motion Picture Testing and Research.* (Report No. 7, AAF Aviation Psychology Research Reports). Washington: U.S. Gov't. Printing Office.

40. Gibson, J.J. (1948), "Studying Perceptual Phenomena". In T. Andrews, Ed., *Methods of Psychology*, New York: Wiley, pp. 158-188.

41. Gibson, J.J., and Crooks, L.E. (1938), "A Theoretical Field-Analysis of Automobile-Driving," *Amer. J. Psychol.*, 51, pp. 453-471.

42. Gibson, J.J., and Mowrer, O.H. (1938), "Determinants of the Perceived Vertical and Horizontal," *Psychol. Rev.*, 45, pp. 300-323.

43. Gibson, J.J., and Radner, M. (1937), "Adaptation, After-Effect, and Contrast in the Perception of Tilted Lines. I Quantitative Studies, *J. Exper. Psychol.*, 20, pp. 453-467.

44. Grant, V.W. (1938), *Psychological Optics* Chicago: Professional Press.

45. Graybiel, A., Clark, B., MacCorquodale, K., and Hupp, D.I. (1946), "Role of Vestibular Nystagmus in the Visual Perception of a Moving Target in the Dark," *Amer. J. Psychol.*, 59, pp. 259-266.

46. Griffin, D.R. (1948), "Topographical Orientation," (Ch. 17 in E.G. Boring, H.S. Langfeld, and H.P. Weld, *Foundations of Psychology)* New York: Wiley.

47. Hamilton, R.J. (1943), "Retroactive Facilitation As a Function of Degree of Generalization Between Tasks," *J. Exper. Psychol.*, 32 pp. 363-376.

48. Harper, R.S., and Boring, E.G. (1948), "Cues," *Amer. J. Psychol.*, 61, pp. 119-123.

49. Hebb, D.O. (1937), "The Innate Organization of Visual Activity. I Perception of Figures by Rats Reared in Total Darkness," *J. Genet. Psychol.*, 51, pp. 101-126.

50. Hebb, D.O. (1949), *The Organization of Behavior.* New York: Wiley.

51. Heider, F. (1927), "Ding und Medium," *Symposion*, 1, pp. 109-157.

52. Heider, F. (1944), "Social Perception and Phenomenal Causality," *Psychol. Rev.*, 51, pp. 358-374.

53. Helmholtz, H. (1925), *Physiological Optics*, Vol. 3 (Ed., J.P.C. Southall). Optical Soc. Amer.

54. Henneman, R.H. (1935), "A Photometric Study of the Perception of Object-Color," *Arch. Psychol.*, 179, pp. 88.

55. Hilgard, E.R. (1948), *Theories of Learning.* New York: Appleton-Century-Crofts.

56. Holt, E.B. (1903), "Eye-Movement and Central Anaesthesia," *Psychol. Monog.*, 4, 17, pp. 3-46.

57. Holway, A.H., and Boring, E.G. (1941), "Determinants of Apparent Visual Size With Distance Variant," *Amer. J. Psychol.*, 54, pp. 21-37.

58. Hull, C.L. (1943), *Principles of Behavior.* New York: Appleton-Century-Crofts.

59. Ivins, W.M. (1946), *Art and Geometry.* Cambridge: Harvard Univ. Press.

60. James, W. (1890), *The Principles of Psychology*, Vol. II. New York: Holt.

61. Katz, D. (1935), *The World of Colour.* (Translated by R.B. MacLeod and C.W. Fox) London: Kegan Paul, Trench, Trubner and Co.

62. Keller, F.S., and Taubman, R.E. (1943), "Studies in International Morse Code. II Errors Made in Code Reception," *J. Appl. Psychol.*, 27, pp. 504-509.

63. Kepes, G. (1944), *The Language of Vision*, Chicago: George Theobald.

64. Klemm, O. (1914), *A History of Psychology.* (Tr. by E.C. Wilm and R. Pintner) New York: Scribner.

65. Klineberg, O. (1940), *Social Psychology.* New York: Holt.

66. Klüver, H. ed. (1942), *Visual Mechanisms.* Lancaster, Pa.: Cattell Press. (Vol. VII of *Biological Symposia).*

67. Koffka, K. (1935), *Principles of Gestalt Psychology.* New York: Harcourt-Brace.

68. Köhler, W. (1940), *Dynamics in Psychology.* New York: Liveright.

69. Köhler, W. and Wallach, H. (1944), "Figural After-Effects. An Investigation of Visual Processes," *Proc. Amer. Philos. Soc.*, 88, pp. 269-357.

70. Langeweische, W. (1944), *Stick and Rudder.* New York: McGraw-Hill.

71. Lashley, K.S. (1938), "Experimental Analysis of Instinctive Behavior," *Psychol. Rev.*, 45, pp. 445-471.

72. Lashley, K.S. and Russell, J.T. (1934), "The Mechanism of Vision. XI. A Preliminary Test of Innate Organization," *J. Genet. Psychol.*, 45, pp. 136-144.

73. Le Conte, J. (1897), *Sight: An Exposition of the Principles of Monocular and Binocular Vision.* New York: D. Appleton and Co.

74. Lewin, K. (1936), *Principles of Topological Psychology.* New York: McGraw-Hill.

75. Lewin, K. (1931), "The Conflict Between Aristotelian And Galileian Modes of Thought in Contemporary Psychology," *J. Gen. Psychol.* 5, pp. 141-177.

76. Luneberg, R.K. (1947), *Mathematical Analysis of Binocular Vision.* Princeton Univ. Press.

77. Mach, E. (1897), *Contributions to the Analysis of the Sensations* (Tr. by C.M. Williams) Chicago: Open Court.

78. MacLeod, R.B. (1947), "The Effects of Artificial Penumbrae on the Brightness of Included Areas," *Miscellanea Psychologica Albert Michotte*, Louvain: Institut Supérieur de Philosophie, pp. 138-154.

79. MacLeod, R.B. (1932), "An Experimental Investigation of Brightness Constancy," *Arch. Psychol.*, 135; pp. 102.

80. McDougall, W. (1921), *An Introduction to Social Psychology.* Boston: Luce and Co.

81. Metzer, W. (1930), "Optische Untersuchungen am Ganzfeld. II Zur Phänomenologie des Homogenen Ganzfelds," *Psychol. Forsch., 13, pp. 6-29.*

82. Metzger, W. (1936), *Gesetze des Sehens.* Frankfurt am Main: Kramer.

83. Miller, N.E., and Dollard, J. (1941), *Social Learning and Imitation.* New Haven: Yale Univ. Press.

84. Morris, C. (1946), *Signs, Language and Behavior.* New York: Prentice-Hall.

85. Nicolson, M.H. (1946), *Newton Demands the Muse.* Princeton: Princeton Univ. Press.

86. Ogle, K.N. (1938), "Induced Size Effect. I. A New Phenomenon in Binocular Space-Perception Associated With the Relative Sizes of the Images of the Two Eyes," *Arch. Ophthal.*, 20, pp. 604-623.

87. Pfaffmann, C. (1948), "Aircraft Landings Without Binocular Cues: A Study Based on Observations Made in Flight," *Amer. J. Psychol.*, 61, pp. 323-335.

88. Robaud, E. (1928), *How Animals Find Their Way About.* (Tr. by I.H. Myers) New York: Harcourt Brace.

89. Ratoosh, P. (1949), "On Interposition As A Cue For the Perception of Distance," *Proc. Nat. Acad. Sci.*, 35, pp. 257-259.

90. Riesen, A. H. (1947), "The Development of Visual Perception in Man and Chimpanzee," *Science*, 106, pp. 107-108.

91. Rubin, E. (1921), *Visuell Wahrgenommene Figuren.* Copenhagen: Gyldendalske Boghandel.

92. Ryan, T.A. (1938), "Dynamic, Physiognomic, and Other Neglected Properties of Perceived Objects: A New Approach to Comprehending," *Amer. J. Psychol.*, 51, pp. 629-650.

93. Ryan, T.A., and Ryan, M.S. (1940), "Geographic Orientation," *Amer. J. Psychol.*, 53, pp. 204-215.

94. Senden, M. von (1932), *Raum-und Gestaltauffassung bei operierten Blindgeborenen Vor und nach der Operation.* Leipzig: Barth.

95. Senders, V.L. (1948), "The Physiological Basis of Visual Acuity," *Psychol. Bull.*, 45, pp. 465-490.

96. Severance, E., and Washburn, M.F. (1907), "The Loss of Associative Power in Words After Long Fixation," *Amer. J. Psychol.*, 18, pp. 182-186.

97. Sherif, M. (1936), *The Psychology of Social Norms*. New York: Harper.

98. Sherif, M., and Cantril, H. (1947), *The Psychology of Ego-Involvements*. New York: Wiley.

99. Spitz, R.A., and Wolfe, K.M. (1946), "The Smiling Response: A Contribution to the Ontogensis of Social Relations," *Genet. Psychol. Monog.*, 34, pp. 57-125.

100. Stavrianos, B.K. (1945), "The Relation of Shape-Perception to Explicit Judgments of Inclination," *Arch. Psychol.*, 296, pp. 94.

101. Stevens, S.S. (1946), "On the Theory of Scales of Measurement, *Science*, 103, pp. 677-680.

102. Stevens, S.S. (1935), "The Operational Definition of Psychological Concepts," *Psychol. Rev.*, 42, pp. 517-527.

103. *Studies in Visual Acuity* (1948), Staff, Personnel Research Section, the Adjutant General's Office (PRS Report No. 742) Washington: U.S. Gov't. Printing Office.

104. Ternus, J. (1926), "Experimentelle Untersuchungen Über Phänomenale Identität," *Psychol. Forsch*, 7, pp. 81-136.

105. Thompson, D.W. (1942), *Growth and Form*. New York: Macmillan.

106. Thouless, R.H. (1931), "Phenomenal Regression to the Real Object I," *Brit. J. Psychol.*, 21, pp. 339-359. Also, "Phenomenal Regression to the Real Object II," *Brit. J. Psychol.*, 22, pp. 1-30.

107. Tinbergen, N. (1948), "Social Releases and the Experimental Method Required For Their Study," *Wilson Bull.*, 60, pp. 6-51.

108. Titchener, E.B. (1910), *Textbook of Psychology*. New York: Macmillan.

109. Tolman, E.C. (1932), *Purposive Behavior in Animals and Men*. New York: Appleton-Century.

110. Tolman, E.C. (1948), "Cognitive Maps in Rats and Man," *Psychol. Rev.*, 55, pp. 189-208.

111. Troland, L.T. (1929), *The Principles of Psychophysiology*, I *The Problems of Psychology and Perception*. New York: Van Nostrand.

112. Troland, L.T., (1930), *The Principles of Psychophysiology*, II *Sensation*. New York: Van Nostrand.

113. Troland, L.T. (1932), *The Principles of Psychophysiology*, III *Cerebration and Action*. New York: Van Nostrand.

114. Vernon, M.D. (1937), "The Perception of Distance," *Brit. J. Psychol.*, 28, pp. 1-11, 115-149.

115. Vernon, M.D. (1937), *Visual Perception*. Cambridge: Cambridge Univ. Press.

116. Ware, W.R. (1900), *Modern Perspective*. New York: MacMillan.

117. Weisenberg, T., and McBride, K.E. (1935), *Aphasia*. New York: The Commonwealth Fund.

118. Wertheimer, M. (1912), "Experimentelle Studien über das Sehen von Bewegung," *Zsch. f. Psychol.* 61, pp. 161-265.

119. Wertheimer, M. (1923), "Untersuchungen zur Lehre von der Gestalt II," *Psychol. Forsch.*, 4, pp. 301-350.

120. Wiener, N. (1949), *Cybernetics*. New York: Wiley.

121. Woodworth, R.S. (1938), *Experimental Psychology*. New York: Henry Holt.

Index

Index

Abstract paintings, 196, 200, 202, 203
Accommodation, sensations of, 20; 21, degree of, 72; as cue for distance, 111; reflex of, 111; 137, 141, 157, 177, 222
Acuity, 9, 108, 112, relation of to texture, 110; dependence on illumination, 113; 175
Adjacent Order, of retinal image 9; 134, 154
Aerial perspective, 72, 114, 116, 137, 141
After images, 31, 32, 58, of color, 133; 147, 159
Aircraft, cues for landing, 128, 129
Aircraft recognition, 91, 202, 219
Air light, 5
Air theory, of space, 6; 7
Alexia, 205
Alley experiments, 190
Ames, 104
Aniseikonia, 106
Aniseikonic glasses, 104, 105
Aperture colors, 5
Asch, 150
Astigmatism, 111
Attached shadow, 94
Autokinetic phenomenon, 127
Automobile driving, 128, 133, 135
Aviation, night, 150
Aviation psychology, 59, 107, 127, 131, 180

Bartlett, 210
Bartley, 65, 73, 113
Berkeley, Bishop, 19, 20, 145, 223, 225
Binocular, cues for depth, 21, 22; disparity, 20, 21, 72, 100, 106, 107, 177; fusion, 103; parallax, 117, 118; perspective, 139; vision, importance of, 107, 108
Blindness, psychic, 204, 205
Blind spot, 48, 49
Blur, gradient of, 112; perspective, 141
Boring, 14, 19, 75
Brain localization, 205
Brightness, as a cue for distance, 137; discrimination, 110
Bruner, 209
Brunswik, 29, 150
Buhler, 5
Bullis, G., 18
Buswell, 155

Carmichael, Dr. Leonard, 65
Carr, 100
Carter, 209
Cartesian, coordinates, 149; geometry, 19
Cassirer, 153, 193
Cast shadow, 94
Cataract patients, 216, 220
Child, C. M., 73
Circle of confusion, 112
Color, sensation, 15, 17; extensity, 17; constancy, 166; film, 166; surface, 166; memory, 209
Conant, 7
Concurrent stimuli, 108
Conscious present, 158, 159
Constancy, hypothesis, 62; of perceived objects, 169; index of, 171; of size at distance, 183; index of, 186
Continuity, of outline, 143
Contour, depth of, 106
Convergence, 20, 21, 30, degree of, 72; as cue for distance, 111; 137, 157, 177, 222
Correspondence, projective, 47, 48, 53; psychophysical, shift in, 195, 196; 153, psychophysical, 57; 61, 62, 75, 104, 165
Courant, 153
Crossed disparity, 104
Cues, theory of, 19; for depth, 21, 22, 59; 69, 71, secondary, 72; primary, 72; 90, 103, 111, for aircraft landing, 128, 129; for depth, 137, 138; conflicting, 149; reciprocal, 149; for depth, 150; reliability of for space perception, 150; shading, 168; for depth, 171; 177, for behavior, 214, 215; for behavior, 221

Dartmouth Eye Institute, 104, 106
Da Vinci, Leonardo, 114
Depth, in vision, 15; of field, 112; test for perception of, 130
Depth at a contour, 93, 106, 107, 137, 138, 141, 165, 177, 178, 221, 228
Depth cues, 21, 22, 137, 138, 150, 171

Depth perception, 6, 7, test for 107, 108
Depth shape, 34, 54, 91,94;
Disorientation, of man in visual space, 5, 6
Discrimination experiment, 221
Discriminative judgment, 221
Distance, in space, 15
Distant stimulus, 63
Dodge, 155
Dollard, 215
Double images, 72, 73
Duke Elder, 116
Duncker, 209

Edge, impressions on visual world, 8
Ego, phenomenal as explained by Mach 27, 28
Eidetic image, 159
Eisenstein, 160
Elementary sensations, 22, 107, 108
Emmert's law, 32
Empiricism, 14, 15, 16, 17, 24, 206, 212, 216
Entoptic phenomena, 31, 32
Epistemology, 13, 24, 212
Equilibrium, and gravity, 33
Euclid, 36, 42, 54, 58, (Euclidean geometry), 122, 188, 190; (non-Euclidian geometry), 122, 189; 188, 189
Euclidean space, 14, 166
Extensity, of space, 15, 16; in color, 17; 169
Eye, defects as an optical instrument, 116
Eye hand co-ordination, 225
Eye movements, pursuit, 32; 57, 58, effect on retinal gradient of motion, 124; compensary, 136; and pictorial composition, 155, 156; saccadic, inhibition of vision during, 146, 147; 155, 156, 159, 160

Facial expressions, 207
Farsightedness, 111
Feedback mechanism, and stimulation, 225
Field theory, 23, 29, 74
Figure ground phenomenon, 8, 196
Film, color, 5, 75, 65, 166; editing, 159; montage, 160
Foreshortening, 83, 172, 183
Form, or shape in vision, 15; or shape in two dimensions, 16, 17; transposability of, 18, 19; 99, (Form perception), past experience in, 192; (Form quality), 19, 22; 151
Frame of reference, geographical, 230
Freud, 199
Frontal surfaces, 66, 70, 76, 84

Ganzfeld, 5
Geometrical space, 15, 188
Gesell, A., 18
Gestalt Theory, 2, 9, 10, 19, 22, 25, 39, 57, 62, 74, 100, 151, 158, 191, 196, 216
Gibson, E., 202, 219

Goodman, 209
Gradient, definition of, 73; of blur, 112; physiological, 73; of density of texture, 78, 80, 172
Gravity, influence on perceptions, 6; 33
Graybiel, 136
Ground theory of space, 6, 7, 59, 60

Haunted swing, 150
Heider, 63
Helmholtz, Hermann Ludwig von, 19, 21, 44, 59, 118, 119, 145, 147, 148, 157
Hemianopsia, 205
Henneman, R. H., 168, 183
Homogeneous stimulation, 4, 5, 6, 53, 64, 108
Horizontal and vertical axes, 148, 150
Horopter, 60, 100
Hull, 215

Identification experiment, 221
Identifying response, 202
Illusory perceptions, 14
Image, of imagination, 158; eidetic, 159; primary memory, 158, 159, 160; of the nose, 228
Indentation, 95, 104, of modelling, 143
Ink blots, and stimulation, 211
Inhomogeneity, 9
Insights, and learning, 216
Instinct, theory of, 206; and discriminative behavior, 207, 220
Invariant properties, 153-154
Ives grating, 109

James, William, 15, 210
Johnson, Samuel, 145

Katz, 5, 169
Kepes, 196, 203
Klineberg, 206
Koffka, 9, 23, 25, 28, 29, 40, 63, 64, 127, 148, 150, 169, 171, 176, 191, 196

Landolt ring, 109
Langewiesche, 129
Lashley, 206, 207, 220, 221
Law, of Pragnanz, 210
Law of the Visual Angle, 83, 175
Laws of sensory organization, 149, 191, 192
Lewin, 229
Linear perspective, 35, 36, 69, 71, 82, 85, 86, 138, 139, 175
Linear projections, 190
Linguistic meaning, 202
Literal perception, 9, 10, 210, 211
Local sign theory, 16, 64, 146, 147
Localization, of points in space, 15, theory of, 133
Location, of points, 15; of after-images, 31, 32
Locke, John, 12, 199
Locomotion, point of aim for, 123, 124; 127, 128, 131, perception of, 135; passive, 135, 136; 153, passive, 154; and retinal stimuli, 224; and perception, 226; study of, 229
Locomotor ataxia, 224
Longitudinal surfaces, 66, 70, 76, 84
Luneberg, Rudolf, 189, 190

Mach, Ernst, 27, 113
Mach rings, 114
MacLeod, 65, 73, 168, 169

Macrogradient, 73
Material world, perception of, 9
Maze learning, 229
McDougall, 206
Meaning, psychology of, 198, 199, 200, 202; emotional, 199; social, 199; (use-meaning) 200, 202; linguistic, 202; context theory of, 203, lapse of, 204; innate, 205, 206.
Meanings, spatial, 200, 202
Memory, immediate, 57; primary, 158, 159; 165
Memory color, 209
Memory trace, 159
Metzger, 5, 52
Michotte, A., 135
Microgradients, 73, 113, 114
Microstructure, 5, 65, 113, 172
Miller, 215
Modelling, perception of, 92; and grades of illumination, 94; 143
Monocular cues for depth, 21
Monocular field, 27, 28
Morris, 202
Motion, types of retinal, 131
Motion parallax, 71, 124, 117, 118
Motion perspective, 119, 120, 124, 140, 141
Motion pictures, 119, 130, 134, sequence and scene, 159
Motion, theory of relativity of, 127
Movement, pursuit, 132; saccadic, 132; stroboscopic, 134
Movements, eye, 136, 146, 147
Mowrer, 149, 150
Muscle sense, 134, 135, 224

Narrow angle lens, 157
Nativism, 14, 15, 24, 216
Nearsightedness, 111
Negative after image of motion, 133.
Negative after effect of tilt and curvature, 195, 196
Newton, Sir Isaac, 14, 15, 189
Nonsense forms, 202, 210
Nose, image of, 106
Nystagmus, 136

Oculogyral illusion, 136
Ogle, 105
Ophthalmology, 147, 148
Optic nerve, anatomy of, 49, 50
Optical range-finder, 177
Order, definition of, 64; successive, 134; adjacent, 134, 151, 154; successive, 154; 161
Ordinal stimulation, 56, 63, 64, 152, 168, 216
Organization, physiological, 25; sensory, 215
Orientation, to physical environment, 229
Outlines, 86

Painting, abstract, 196, 200, 203
Panoramic vision, 108
Peephole vision, 178
Perception, relation to sensation, 12, 13; of external world, 13, first visual, 17, 18; constancy in, 33, 34; 43, 44, test for depth, 107, 108; depth, test for, 130; of acceleration, 134, 135; and sensation, 138; form, past experience in, 192; meaningful, 198, 199; literal, 210, 211; schematic, 211, 212; space, motor theory of, 223, 224.
Perceptual constancy, 23
Perspective, aerial, 72; 114, 116, 137, 141
Perspective, linear, 70, 82, 85, 86, 137, 138, 139, 175
Perspective, binocular, 139
Perspective geometry, 82
Perspective illusion, 181
Perspective, motion, 124, 140, 141
Perspective, of blur, 141
Perspective, size, 138
Perspective, texture, 138
Pfaffmann, 129
Phenomenal ego, 225, 226, stimuli for, 226
Phi phenomenon, 134, 161
Physiological gradients, 73
Pictorial vision, 181
Picture plane, 54, 55, relation to retinal projection, 79, 104, 122, 153, 218
Pinhole camera, experiment of, 47
Plato, 152
Point of aim, 127. 128
Point sensations, 15, 22, 59, 114, 133, 191
Post, Wiley, 108
Posture, relation to vision, 32, 33; and perception, 226; 228
Postural equilibrium, 150
Postural stimulation, 149, 150
Pragnanz, law of, 210
Primary cues for depth, 21, 72
Primary memory, 158, 159, image, 160, 165
Primitive man, vision of, 198
Projected form, definition of, 16; 188, 190, shape, 34, 35; and transformation, 169; 171, 174
Projective correspondence, 47, 48, 534, 153, geometry, 153
Projective transformation, 153
Protuberance, 95, 104
Protuberances, of modelling, 143
Proximal stimulus, 63
Psychic blindness, 204, 205

Psychomotor behavior, 135
Psychophysical experimentation 9, correspondence, to visual perception, 8, 25, 51, 57, 61, 62, 75; shift in, 75; 104, 165, 195, 196; shift in, 195, 196
Pursuit movements, 126, 127, 132

Ratoosh, 143
Reaction, identifying, 220
Reference - axes, 149
Relativity of motion, theory of, 126, 165, 189
Relief, perception of, 92
Resolution, definition of, 112
Response, identifying, 202; smiling in the infant, 207; smiling in infants, 220; identifying, learned, 221

Retina, description of, 46, 56
Retinal image, differentiation of, 8, 9; 33, 46, 50, 52, 57, 67, 61, 62, 79, 117, 118, motion, types of, 131
Retinal receptors, density of, 48, 49; distribution of, 55
Reversal of relief, 99
Reversible figures, and stimulation, 211
Riesen, Dr. A., 221
Robbins, 153
Rubin, 39
Russell, 220

Saccadic eye movements, 29, 30, 126, 132, 146, 147, 155, 159, 160
Scanning in visual world, 29, 30, 31, 126
Schematic perception, 9, 211, 212
Schooler, 209
Secondary cues for depth, 21, 72
Selective perception, 9
Senders, 113
Sensation, relation to perception, 12, 13; 43, 44; and perception, 128; color, 15, 16; 15, 16; 16, 17
Sensations, elementary, 19, 22, 107, 108; innate, 15, 163, 171, 218; point, 15, 22, 133, 191; of accommodations, 19
Sensory organization, 22, 23, 215, 187; laws of, 191, 192; external forces of, 196
Shading, 72, 94, 137, 143, 144; as a cue for depth, 168
Shape, or form in vision, 15; or form in two dimensions, 16, 17; constancy of perceived objects, 169; apparent, 171, 172, projected, 174.
Signs, in meaning, 199; local, 222
Size, constancy of, 174
Size - perspective, 138
Snellen, 109
Social meanings, 199
Social norms, 212; relation to perception, 211
Social stereotypes, 210
Somaesthesia, 226, 227
Space, geometrical, 183, 188; sensation of, 8; perception, analysis of, 4; perception, motor theory of, 223, 224
Spatial meanings, 200, 202
Spatial world, perception of, 9
Spectroscope, 166
Spitz, 207, 220
Stereoscope, purpose of, 6, 20, 104, 107
Stereoscopic acuity, 110
Stereoscopic impressions, 4, photographs, 107
Stereotypes, social, 210; 211
Stevens, 75
Stimulation, homogeneous, 53; ordinal, 63, 64; homogeneous, 64, 108, impoverished 151, 211, ordinal, 216
Stimuli, concurrent, 108
Stimulus, method of impoverishing, 9, 10; definition of, 63; distant, 63; proximal, 63; definition of, 151, 152, 215; for a psychophysical experiment, 215
Stimulus generalization, 219
Stimulus - objects, 215, 216
Stimulus - variable, relation to visual space, 8, 51, 52.
Stravrianos, 172
Stroboscopic movement, 134, 161
Successive order, 134, 161
Surface - color, 6, 166
Surface, impressions on visual world, 8
Superposition, 71, 137, 142, 177; type of depth-perception, 228
Symbols, 164, 204, 205, 212; in meaning, 199

Tachistoscope, 9, 211
Television camera, likeness to retina, 48; 56
Ternus, 57
Test for depth perception, 130
Texture, compression of, 173; gradient of density of, 66, 67; relation of to acuity, 110; visual, elements and gaps of, 80, visual, 81
Texture - perspective, 138
Theory, context (of meaning), 203; instinct, 206; motor, of space perception, 223, 224
Thompson, D'Arcy, 153
Thouless, 168
Titchener, 203
Topology, 153
Transfer of learning, 219
Transformation, 104, of images, 132; of retinal pattern, 152; projective, 153, 169; groups, 193; projective, 153; 191
Transposability, of form, 18, 19, 55, 56, 151; of retinal image, 153
Triangulation, 19
Troland, 100

Unconscious inference, 19, 119, 145, 146
Uncrossed disparity, 104
Use - meanings, 199, 200, 202

Vernier acuity, 109
Vernon, 212
Vertical and horizontal axes, 148, 149, 150
Visual acuity, 108, 109, 110, 111, 112, 175
Visual agnosia, 205
Visual field, boundaries of, 27, deformation of, 40; boundaries of, 45, 46, 155, 226, 228
Visual kinesthesis, 124, 224
Visual motion, stimulus for, 160, 161
Visual perception, evolution of, 60
Visual texture, 52, 65; elements and gaps of, 80, 81
Visual world, properties of, 3
Volkmann, John, 181
Von Senden, 217, 219

Wertheimer, 149, 150, 160, 161, 196
Witkin, 150
Wolfe, 207
Woodworth, 202
Word blindness, 205
Wundt, 15
Wheatstone, 20, 107, 111
Wide - angle lens, 157

Zeno, 134