Psychological Bulletin
1974, Vol. 81, No. 8, 471–497

PICTURE PERCEPTION:
TOWARD A THEORETICAL MODEL

MARGARET A. HAGEN [1]

Boston University

J. J. Gibson's new theory of picture perception is described, and a program of research within his framework is outlined. An analysis of pictorial information is proposed in which a systematic investigation of the structural components of pictures and their varying effects on perception is seen as preliminary to the postulation of hypothetical pickup mechanisms. The basic components of pictures are described, and literature is reviewed in the problem areas of distorted and impoverished information, observation from the wrong station point, coexisting flatness and depth information, and the ambiguity of the source of a single projection. The feasibility of the Gibsonian enterprise is demonstrated, and further avenues for research into a structural analysis of pictorial information are pointed out.

A *picture* is a delimited surface with markings on it that represent something. This article is concerned with pictures in the Western post-Renaissance mode; namely, that attempt to represent, by means of structural equivalence of some order, the layout of surfaces in the world. Alternative modes of representation, perhaps requiring alternative analyses, are beyond the scope of this article.

By what means can a picture be said to represent, to bring clearly before the mind, the segment of the world pictured? The aim of this article is to demonstrate the feasibility of extending J. J. Gibson's theory of perception via motion-generated information to an adequate theory of frozen pictorial information and to explicate the problems that must be dealt with in any comprehensive theory of the perception of pictures. The problems to be dealt with herein are the use by artists of modified linear perspective, caricature, impoverishment of information in outline drawings and black-and-white photographs, the coexistence of flatness and depth information, observation from the wrong station point, and the ambiguity of a single projection (i.e., the same projected form can arise from an infinite variety of shapes).

THE GIBSONIAN MODEL OF INFORMATION

For Gibson the essential condition of resemblance between pictures and the world was the correspondence of information contained in the structured light to the eye, coming either from a picture or from the real scene. Gibson (1971) stated his position quite succinctly in the following passage.

> A picture is a surface so treated that a delimited optic array to a point of observation is made available which contains the same kind of information that is found in the ambient optic arrays of an ordinary environment [p. 31].

Central to Gibson's theory is the assertion that picking up information about the persistent properties of objects and layouts of surfaces in the world is the grasping of invariant structure. In order to consider the implications of this view, it is necessary to explicate more fully the ideas of structure and invariant information.

THEORY OF OPTICAL STRUCTURE AND INFORMATION

Gibson (1966) stated that ambient light must have structure if it is to carry *information about* (which means specification of) the environment. So whence comes the structure of the ambient light in an illuminated environment? The ambient light reverberating in the medium affords many possible station points, or points of observation. The ambient light converging on any one of these points becomes the potential light to the eye. This

---

[1] The author wishes to thank William M. Mace and Herbert L. Pick, Jr. for their conceptual assistance in the preparation of the original manuscript.

Requests for reprints should be sent to Margaret A. Hagen, Department of Psychology, Boston University, 64 Cummington Street, Boston, Massachusetts 02215.

light is structured, modeled if you will, by the structure of the physical surface layout of the environment and as such can convey specific information about it. The surface layout consists of the faces (with respect to the observer) of surfaces and objects and the smaller facets within them.

Boundaries between faces determine boundaries in the light to the eye just as the form and slant with respect to the observer determine the angular projection of the light. Gibson argued that boundaries in the optic array, in the light converging on a station point, are due to the different amounts and colors of light reflected from the surfaces of objects to that station point. Boundaries, or changes in intensity, occur when two adjacent faces are at different angles of inclination to the light source. Surfaces project different intensities if they have different composition or pigmentation and thus produce borders in the optic array via their different reflectance. These differences in reflectance, Gibson (1966) wrote, are further reinforced usually by differences in color, doubly guaranteeing borders in the light to the eye.

> In short, an array is structured, i.e. caused to have borders within it, by (1) the physical inclinations of the faces and facets of surfaces, (2) the reflectance of the substances, and (3) the *spectral* reflectance of the substances, or chromatic reflectance, and (4) shadows [p. 194].

So slant, composition, color, and shadow determine borders in the light to the eye and therefore specify borders of environmental faces and facets.

We can now assert a lawful relationship between the light to the eye and the surface layout of the environment. The presence of environmental information in reflected light is defined, Gibson argued, by the univocal relationship of a property of the stimulus to a property of the object. This is what he meant by the conveying of environmental information. But determinate information (information specific to the persistent properties of a particular layout) is invariant information; that is, only through invariance (nonchange amidst change) can we arrive at specificity.

So far our postulated environment is stationary and our observer, motionless. Gibson (1966) pointed out that "the stationary perspective of light does not specify the gross physical layout of its source. The information in the static array . . . is ambiguous in this respect [p. 199]." Movement of the observer is critical in resolving the ambiguities inherent in a static array because it provides opportunity for the detection of invariant properties of the array through the systematic perspective changes occasioned by motion. Motion perspective is the regular perspective change or optical flow of the texture elements (faces or facets) as the observer moves in the environment.

Slants, edges, corners, composition, and various separations of surfaces can be specified in gradients of texture flow velocities or by specific types of discontinuities in the optical flow. For example, Gibson, Olum, and Rosenblatt (1955) have shown that the location, slant, and shape of surfaces as well as the movement of the observer are specified in the structure of optical texture flow for the case of straight line motion. Hay (1966) showed that the initial orientation and shape of a flat surface is specified in the sequence of projections of rigid motions of that surface. These distinctive characteristics of the optical flow as the observer moves are specifically determined by particular characteristics of the physical environment and as such constitute information for that environment. It is only through transformations of the optic array, through motion of the observer, that the invariants in the optical flow specific to a particular surface layout can be detected.

PICTURES AS CARRIERS OF OPTICAL INFORMATION

If the perception of the persistent properties of objects and surface layouts is dependent on the motion-generated information previously described, then successful picture perception seems to provide a serious embarrassment to Gibson's theory of perception. Pictures as static, isolated, perspective instantiations of three-dimensional scenes cannot, apparently, act as occasions of such motion-generated information, so how then do they function as carriers of information about such scenes? Evidence is subsequently reviewed to suggest that pictorial perception is determinate not only with full-color, veridical, angle-for-angle, trompe-l'oeil representations but also with pictures suffering from the problems of distorted or impoverished information, observation from the wrong station point, coexisting surface and depth information, and ambiguity of the source of the plane projection.

How do we deal with this evidence in the building of a coherent theory of picture perception? The field has gone in two directions. The more frequented avenue has been to build hypothetical structures of the mind capable of perceiving the identity and location of pictured objects through the addition or interpretation of information gained through past experience. This avenue has been followed by Goodman (1968), with his theory of pictorial structure as an arbitrary, conventional language that must be learned as a skill, like reading; Hochberg (1968), who has hypothesized stored schematic maps evoked by feature pickup to account for picture identification; Gregory (1970), with his theory of evoked stored object hypotheses used in picture recognition; and many others.

The alternative avenue, in the absence of motion-generated information, is to look not for hypothetical mental structures but for the structure in pictures; that is, for invariance in pictures, for what fails to change as the observer moves or the slant of the picture is changed. We can look for arrangements of, or relations between, lines and forms and textures that remain invariant under transformation and hence provide the structural basis for determinate pictorial perception. We can look for the relations among pictured lines, forms, and textures that duplicate invariant information in the real scenes pictured. Frozen three-dimensional information is equivocal in part because its source may be a picture and not a real scene; hence, such real world information may be captured in a picture (Pirenne, 1970; Smith, 1958).

Gibson (1971) has argued that pictures perform their representative function by means of just such an informational equivalence to the scenes they represent. Purdy (1960) and Gibson (1950) have suggested that one may look for pictorial structure in the static rates of change (e.g., texture gradients) in the absence of the rates of change of motion perspective information. It is thus possible to separate descriptions of pickup mechanisms from descriptions of what is "pickupable," and indeed, given the state of the field, it seems most advisable to do so. Until we arrive at a comprehensive, systematic analysis and description of the structures of different pictures and their varying effects on perception, hypothetical models of the structure of the pickup mechanism seem premature.

In the review of the different types of pictorial materials to be discussed next, it is hoped that the feasibility of the Gibsonian enterprise is demonstrated and further avenues for research into a structural analysis of pictorial information pointed out. Since the data are meager and a systematic analysis of the problems of picture perception has not yet been undertaken, a precise empirically based model does not come out of this review, but the complex aspects of pictures subsequently discussed must, in the future, be embraced and resolved by any comprehensive theory of the structure of pictures and pictorial perception.

BASIC PICTORIAL COMPONENTS IN WESTERN POST-RENAISSANCE ART

According to Gibson, a picture succeeds as a representation of a real object or scene because it reflects the same structural information in the light to the eye as the scene represented. Light carries information because it is structured, or caused to have borders within it, by the slant, composition, color, and shadows of surfaces in the world. Linear perspective is a necessary by-product of this structuring of light. It is a function of the properties of light and the distance from the light-reflecting object to the point of observation. Light travels more or less in straight lines. As distance between object and observer increases, visual angle (size of retinal image) decreases in a lawful gradient of projected size. This diminishing projected size is true of all receding textured surfaces. Linear perspective is simply a special case of this texture size gradient; it specifies the edges of surfaces, the borders in the optic array. Texture gradients and therefore linear and size perspective are information for distance, since they are determined by the distance between observer and observed and the properties of point projection. The consideration of the pupil and lens system does not invalidate the information character of perspective.

Linear perspective has served as the primary basis of depth information in Western post-Renaissance art. It is employed to represent correct layout of the surfaces of objects and correct spatial relations between objects in depth. The success of linear perspective as a source of information about the depicted world rests, in part, on the assumption that the single perspective view chosen must show the best or most characteristic aspect of the object or scene. When such an aspect is not depicted, successful representation may depend on contextual or redundant information.

Linear perspective is not, however, the *only* Western post-Renaissance discovery for translating the three-dimensional world to the two-dimensional picture plane. Relative size has also been employed as a depth cue in its role as a special case of a texture gradient. Superposition (overlapping or occlusion) is also information for relative depth. Aerial perspective may be employed to give relative depth via a gradient of clarity of outline and warm-to-cold color gradations. Depth is further clarified and information for flatness minimized by the addition of shadows. Shadows are used to indicate differential orientation of surfaces to the source of illumination and to mold the volume of a single surface (Arnheim, 1969; Gombrich, 1971; Wolfflin, 1950).

These traditional cues or formulae comprise the model of representation that served as the aim of Western art from the Renaissance to the nineteenth century French revolution in painting. They still produce, for the average Western adult, pictures that look "real" or lifelike. These formulae induce good pictorial depth perception because as rules, as principles, they capture sources of depth information that are structurally similar to the sources of depth information in the scenes represented. The classic cues of perspective (linear, size, texture, and aerial), superposition, and shadowing do not represent simply the limiting cases of motion-generated information, nor are they arbitrary conventional symbols invented by artists. Rather, they may be considered in their own right as informational components of pictures, bearing an obviously present (if unanalyzed) relation to three-dimensional reality, contributing to determinate picture perception, and as such requiring careful investigation.

The Western post-Renaissance model assumes the static situation. It assumes a frozen world and frozen viewer at a single station point (which must theoretically be the point of projection for the picture). The model is also burdened with a variety of viewing restrictions made explicit by Gibson (1971) and discussed in the following section.

DISTORTED INFORMATION

Modified Linear Perspective

Not everyone concerned with the making of pictures agrees with Gibson's theory of structural equivalence as the essential aspect of the resemblance between the picture and the depicted. Much of the argument about the nature of the resemblance revolves around the status of linear perspective. Goodman (1968) argued:

> Representational customs, which govern realism, also tend to generate resemblance. That a picture looks like nature often means only that it looks the way nature is usually painted [p. 39].

He stated that pictures in perspective, like any others, have to be read; and the ability to read has to be acquired.

Gibson (1971) attacked the Goodman language view of pictorial art by attacking the arbitrary conventionality of linear perspective. He argued that the artist must, of necessity and not arbitrarily, use perspective geometry to transcribe the three-dimensional world onto a two-dimensional surface. However, Goodman wrote:

> In diametric contradiction to what Gibson says, the artist who wants to produce a spatial representation that the present-day Western eye will accept as faithful must defy the "laws of geometry" [1968, p. 16].

Goodman argued that not only are the conditions of observation of linear perspective cumbersome and unnatural, but linear perspective itself, as traditionally used in art, is unfaithful to the true perspective of light rays converging on a point. He stated that artists and cameras systematically alter certain aspects of true perspective to produce the artistic linear perspective, which functions as a depth symbol in the Western mode of representation.

For Gibson the conflict resolved into an inadequate consideration on Goodman's part of the conditions of observation that must accompany the successful use of perspective in art. According to Gibson the adequacy and validity of the laws of perspective have not been challenged. For him the laws of perspective are not conventions in any sense; they are a geometry, a property of light, and an optical and logical necessity. When the laws of perspective are employed in picture making, however, they are indeed saddled with conventions.

Gibson formulated rules for observing the picture surface: (*a*) It should be seen with one eye; (*b*) it should be upright and perpendicular to the line of sight, instead of slanted, and its distance must be just such that the visual solid angle from the picture is the same as would the visual solid angle be from the thing pictured; and (*c*) there should be an aperture in front of the eye hiding everything but the picture itself. He noted that these viewing prescriptions are almost never followed in practice. He stated that the *system* of perspective projection, the optical geometry, must be distinguished from the *practice* of perspective. "Perspective distortion" is produced by inadequate attention to the conditions of observation and not by the infidelity to the real world of the laws of perspective. Since viewers could not be made to abide by the prescribed conditions of observation, the system was subsequently modified to reduce the resulting perspective distortion. The fidelity of the original system of linear perspective, under proper viewing conditions, remains unchallenged.

The modified artistic perspective system, however, raises serious questions for those who would argue that there exists a nonarbitrary structural resemblance between pictures and the world, between linear perspective portraying depth in a picture and depth information in the real world. Can observers tell the difference between modified and true perspective under the *conventional* conditions of observation—that is, at the proper station point, with one eye, and with a window occluding the frame? Can they tell the difference under *normal* conditions of observation—that is, walking around in a gallery? Are there strict constraints on the degree of distortion required to produce an apparently faithful picture observed in a free situation? It may well be that after an exhaustive analysis of pictorial information the explanation for picture perception may require reference to some sort of pictorial mode of perceiving, learned through experience with pictures as representations of reality, as well as to Gibson's theory of the structural–informational equivalence between the picture and the depicted.

Caricature

Caricature has heretofore provided a special problem in picture perception because of its nonconformity with point-to-point correspondence theories on the nature of the structural resemblance between the picture and the depicted. A caricature makes no attempt to duplicate the light rays from the figure of the person represented, according to Renaissance principles of art, but succeeds, nevertheless, in conveying information about the person portrayed. Gibson dealt with the problem in his new definition of pictures as carriers of optical information about the represented persons or scenes.

> But the definition is broad enough also to admit the case of a caricature, where the contrasts of luminous energy are quite different, and even the forms are different, but where the high-order information to specify a particular person is common to both arrays. In short the optic array from a picture and the optic array from a world can provide the same information without providing the same stimulation. Hence an artist can capture the *information* about something without replicating its *sensations* [1971, p. 31].

Presumably the high-order information specifying an individual's face consists of such relations as the length or sharpness of the forehead curve relative to the length or sharpness of the nose curve, the width of the eyes relative to the length of the nose, and so on. Kennedy (1971) quoted Gibson as saying that the information or invariants "are not to be found in the light rays and color patches, and not even in the *forms* of these elements, but rather in the *forms* of form [p. 27]." That is to say, the structural equivalence between a caricature and the person caricatured depends on the arrangement of forms and not on angular congruence or point-to-point correspondence. If perception consists in part of the grasping of such information as "forehead curve is very much greater than nose curve" or even "nostrils are very wide relative to most other nostrils," and if it is this sort of relation information that specifies a particular face, then it follows that a drawing or caricature that preserves these relations in portrayal will specify, even after exaggeration, the same face.

There are certain constraints on caricature, however. For instance, if an individual's small nose and wide mouth are more distinctive than his small nose and wide eyes, then a caricature exaggerating the former will be more successful than one exaggerating the latter. There may also be limits on how much a relation can be exaggerated before it is treated as a new relation no longer specific to the subject of the caricature. The distinctive nose–mouth relation must be preserved under the caricature transformation. The problem of which facial relation is most distinctive, most specific to a particular individual, remains currently in the eyes of the caricaturist.

There is presently little or no research on the subject of caricature. A careful developmental investigation of the means and limits of caricature distortion would give us invaluable information about facial perception and the sensitivity of the untutored to distinctive features versus photographic detail. Shaw, McIntyre, and Mace (in press) have achieved promising success in the determination of strain and shearing transformations as specifying the perceptual invariants of both aging and age levels; they intend to apply a similar analysis to caricature. Ryan and Schwartz (1956), in a very different approach, investigated cartoons, the cousins of caricature. They ran an experiment to test the effects of different media on the rapid pickup of pictorial information for such items as the posture of hands, the position of a switch, etc. They compared a black-and-white photograph, an ink-and-wash-shaded drawing, a high-fidelity line drawing, and cartoons. The order of difficulty (easiest to hardest) was as follows: (*a*) cartoons, (*b*) photograph and shaded drawings, (*c*) line drawings. It may well be that the rapid pickup of information in a picture is best in pictures capturing the least amount of information not relevant to the specific task demands. Further research is needed, especially of a developmental nature. Gibson's theory may well solve the problem of caricature but it currently does so mainly by assertion.

Impoverished Information

Outline drawings present a different problem from caricature in that the information they present is impoverished, not distorted. Kennedy (1971) has defined outline figures in his thesis.

> A tentative definition of these is that they are line-figures formed by continuous lines with no areas of "shading," i.e. there are no large areas, usually marked off by lines, whose width is several times larger than the lines in the drawing, which have been covered or "filled in" with pigment. In outline figures the *length, direction* and *shape* of every line is significant [p. 2].

Gibson (1951) made the point that outline forms geometrically represent the margins of a solid form or the edges of a surface form and that they have two margins instead of the single margin given by the edge of an object. Line drawings preserve only the information for the edges of objects, for boundaries in the optic array. These boundaries show discontinuities of optic layout due to changes of color or slant or texture of surfaces or the presence of shadows. These sources determine contrasts in the optic array, but any contrast can be due to any one of these sources. Kennedy noted,

> So any particular contrast, considered on its own, is ambiguous with respect to its environmental origins. However, there may be information in the optic array, in terms of the *arrangement of contrasts* rather than the isolated contrast, to specify the origin of any particular contrast [1971, p. 31].

Kennedy, however, stated that we do not yet know enough to discuss in any detail the structure or arrangement of contrasts. We know they exist. We don't know how to describe them and we don't know how they operate in perception. Their role in resolving pictorial ambiguity in outline drawings must be crucial.

If contrasts in the optic array are determined by changes in slant, reflectance, composition, or shadowing, how is it that the change itself can be specified by a line on paper independent of any portrayal of the surfaces, reflectances, or shadows determining the change, except by means of that same line on paper and its position or function in an arrangement of lines? Kennedy simply remarked, "It can be said that lines have two contours, and so give rise to two contrasts in projected optic arrays [1971, p. 32]." Two contrasts specify two changes, that is, at least three sources of differential intensity in the optic array. It doesn't necessarily follow from the presence of areas of change in the optic array that a line drawn on paper will give rise to a perception of that change. However, granting the possibility, Kennedy's argument continues along the following lines. Because a contrast in the optic array may be projected by any number of sources of change in surface layout or shadowing, so too, perhaps, may a line in a line drawing represent any one of those features. Since a contrast in the optic array presumably gives rise to a determinate perception of change through its position in an array or arrangement of contrasts, it is feasible to undertake research on the types of information that can be picked up in line drawings. It should be noted, however, that the sources of determinate perception in line drawings need not be the same sources as those that give rise to determinate perception in a colored, textured picture or a three-dimensional scene.

Kennedy then, with Gibson, was arguing that edge, or boundary, information is fundamental to perception of surface layout. If a picture captures that edge information, then it should serve as a source of information for surface layout. The perception of surfaces in the world, however, is multiply determined both by motion-generated information and the co-occurrence of sources of contrast with a single surface change. For example, as the eye moves from observing one object to observing another, it is unlikely in the natural world that the change of surface will occur without an accompanying change in color or texture or slant. Nonoccurrence of such simultaneous change (e.g., a change in slant without accompanying changes in texture and color) indicates the several surfaces of a single object. Outline drawings lack this crucial redundancy of information about surface change, but Kennedy argued that the resolution of the contrast ambiguity lies in the arrangements of the lines specifying those contrasts. He suggested a program of research investigating responses to arrangements of lines to achieve a theory of that resolution. Kennedy's own thesis question along such lines was, "Can line-segments in line-drawings represent basic features of surface layout [p. 4]?" His answer was mostly in terms of demonstrations for the adult reader, similar to Gibson's (1969a) observations in "Three Kinds of Equivocal Information in Line Drawings." Such demonstrations do *not* provide answers to the basic questions of the nature of the structural relation between outline drawings and their subject matter. Is boundary information in a line drawing a sufficient source of information for the untutored observer, a very young child, or a member of a culture with no outline drawings? Must the relation between outline drawings and the world they picture be learned; must the relation between outline drawings and other pictures be learned? Do they really preserve enough information for unambiguous perception by the naive observer? Much of the developmental research on picture perception makes use of the impoverished information in outline drawings or black-and-white photographs to test the ability of the untutored to recognize pictures of familiar objects, to isolate or complete the contour of single objects, and to see depth or the layout of surfaces in pictures. In the brief review that follows, the impoverishment of information in black-and-white photographs is clarified.

Recognition of Pictured Objects

For the simple recognition of pictured objects, data from infants are far from conclusive. There is, however, one study of a 19-month-old child explicitly on recognition of familiar objects in pictures. Hochberg and Brooks (1962) reared their child until the age of 19 months with extremely restricted exposure to pictures and no exposure to picture-plus-naming experiences. At 19 months he was able to successfully identify simple and complex line drawings and photographs of familiar objects. The line drawings always (except on one presentation) preceded the photographs of the same object in the task. Of outline drawings they wrote:

> "Ghost shapes," as Gibson has called them, may be anemic, but they are by no means deceased . . . the complete absence of instruction in the present case (the absence of "association" between picture and represented object) points to *some* irreducible minimum of native ability for pictorial recognition. If it is true also that there are cultures in which this ability is absent, such deficiency will require special explanation; we cannot assert that it is simply a matter of having not yet learned the "language of pictures [p. 628]."

That some African natives have not yet learned the "language of pictures" was the assumption made by Segall, Campbell, and Herskovits (1966). They quoted Herskovits:

> I have had an experience of this kind, similar to that reported from many parts of the world by those who have had occasion to show photographs to persons who had never seen a photograph before. To those of us accustomed to the idiom of the realism of the photographic lens, the degree of conventionalization that inheres in even the clearest, most accurate photograph is something of a shock. For, in truth, even the clearest photograph is a convention; a translation of a three-dimensional subject into two dimensions, with color transmuted into shades of black and white. In the instance to which I refer, a Bush Negro woman turned a photograph of her own son this way and that, in attempting to make sense out of the shadings of grays on the piece of paper she held. It was only when the details of the photograph were pointed out to her that she was able to perceive the subject [p. 32].

Whether or not such "data" are any good, Segall et al. have a point about the impoverishment and conventionality of black-and-white photographs. They are certainly two-dimensional representations of real scenes and as such are sources of both types of information. They are less impoverished than outline drawings, however, because they preserve texture and shading information for the specification of sources of contrast in the optic array. Representation of color changes in the optic array as shades of gray is of course a convention, although there certainly exists a predictable relationship between the shades of gray and the color changes. Such a relationship may have to be learned, although Hochberg and Brooks (1962) indicated that it does not.

Segall et al. offered an anecdote on picture perception in relatively pictureless environments.

> It is interesting that the experience of anthropologists shows that motion pictures are almost universally perceived without trouble and that colored prints are also—although here the naivete of the respondents may be questioned in more recent field experience [1966, p. 33].

However, evidence contradictory to this statement was found by Nadel (1937). He compared the picture perception of the Nupe, a people with imageless art, with the Yoruba, a people with art rich in images, using black-and-white photographs and found no failures of picture detection in either group, although identifications were frequently rather idiosyncratic or culture bound.

Deregowski (1968) cited Brimble (1963), who submitted men and women from Barotseland to an identification test consisting of two sheets of "simple line drawings," 58 in all. Correct identification responses ranged from 94.3% to 98.7%. Deregowski pointed out that it is impossible to tell if the rare errors were due to "detection" errors (failure to recognize the representative nature of the drawing) or "identification" errors (naming the wrong object).

In Deregowski's own work (1968) in rural Zambia, he found that both school children and men were above chance on recognition of photographs of models of familiar animals, but only children were above chance on recognition of pictures of unfamiliar animals. His data suggest that errors were in identification or inability to identify and not due to detection failure. Similarly, Mundy-Castle (1966) in Ghana and Kilbride and Robbins (1968) in Uganda found many "misidentifications" of outline drawings of animals and a road but no failures to perceive the outline drawings as representations of three-dimensional objects and surfaces.

As a last note on "primitive" art, the work of Dennis (1960) deserves mention. He studied the human figure drawings of illiterate Bedouins living in the Syrian desert. They have no native forms of representational art and only minimal exposure to Western art forms. They could draw pictures of men both in outline and filled in but usually in a rather sticklike form. Although there is no clear relation between people's drawings and their perceptions of drawings, it is interesting that a people with almost no experience with representative art of any sort could accept and produce a variety of "conventions" in drawing human forms. Dennis stressed the impoverishment of their drawings, the lack of facial features, clothing, and other detail, but this very impoverishment, this schematization, is intriguing. There seems to be a spontaneous acceptance of the evocative power of lines and outlines, of the representation of a large object on a small piece of paper, of a round object on a flat surface. Such acceptance does not fit well with Segall et al.'s assertion of African inability to recognize the representative character of black-and-white photographs. Of course, the Bedouins may be an unusual group, but their relative innocence of representation was well established. More work needs to be done as Segall et al. (1966) noted:

> Here is an area in which systematic research should be done to support anecdote, for naive respondents on whom to try such experiments may now or will soon be lacking [p. 35–36].

Davenport and Rogers (1971) have gone even further than the African work in their investigation of the ability of the untutored to recognize black-and-white photographs of objects. Two chimpanzees and an orangutan had previously been trained to match a visual sample with a haptic object. In this study they were required to match a photograph with a haptic object. They had no familiarity with photographs. Correct choices with color photographs ranged from 80%–100% across 40 different pictures. With black-and-white photographs, choices were 60%–100% correct. "We can now conclude that apes can perceive a photograph of an object *at first sight* [p. 320]." If this finding is replicable, it certainly suggests a fundamentally evocative character for photographs, despite their multiple dissimilarities to reality, and argues strongly against a theory of pictures as an arbitrary language whose rules must be learned through experience.

Overlapped, Embedded, and Fragmented Contour

Bower (1966) used infants 50–60 days old who had never seen a triangle before except as they occur in the real world. He conditioned them to a wire triangle crossed with an iron bar. Infants generalized most to a plain uninterrupted triangle. He repeated the experiment with slides and there was no preference for one test figure over any other.

> As in the case of space perception, the infants' performance appeared to depend not on static retinal cues but rather on the information contained in variables, such as motion parallax, that are available only to a mobile organism viewing a three-dimensional array [p. 90].

Interpreting baby studies is always difficult, but Bower's data seem to suggest that motion parallax information is necessary to determine the *identity* of an object across the superposition transformation. Motion parallax appears to be needed to establish the identity or unity of the triangle alone when crossed with a bar. For very young infants, overlapped pictured objects seem to be seen as flat, not layered in depth. Bower's data suggest that it is possible to ignore or fail to perceive depth information in a picture while attending to the information for pictures as flat objects (single surfaces).

Ghent (1956) studied the performance of children 4–13 years old on overlapping and embedded figures. She found that young children exhibit little difficulty in analyzing overlapping figures, regardless of whether they are realistic or geometric, and rather great difficulty in analyzing embedded figures. She suggested that

> when a figure was so "hidden" by other forms that the boundaries of the added forms *coincided* with those of the original figure, the figure would be harder to find than when the boundaries of the added forms *intersected* with those of the original figure...it could be said that the improvement with age reflects an increase in the capacity to perceive a boundary as belonging to more than one figure [p. 587].

Since her figures were outline drawings with no background, it is extremely difficult to generalize her findings to other types of pictures, more complete representations. Coincidence and intersection of boundaries, however, certainly occur in the real world and in the world pictured. Her findings with respect to embedded figures are an interesting comment on Kennedy's (1971) assertion that lines have two contours and hence give rise to two contrasts in the projected optic array. The younger children's difficulty in analyzing embedded figures suggests that they see only one contour and thus one contrast. It may be that contrast ambiguity in line drawings is a learned perception or a theoretical figment and that the arrangement of lines in line drawings is far more determining of perception than previously supposed.

Bower's (1966) data suggest that a pictured object whose contour is partially overlapped by another object is not completed, or filled in, by infants. Ghent's (1956) data, on the other hand, show that children as young as four years have little trouble perceiving pictured objects with overlapping contours. This does not, however, reflect a developmental change in the ability to fill in or elaborate impoverished information because in Ghent's case contour was intersected but not disrupted, and the surface of the object was not occluded by another depicted surface as it was in Bower's study. Ghent's overlapping study was a study of sensitivity to transparency rather than to superposition or occlusion information.

There seems to exist no programmatic research on the development of sensitivity to the multiple forms of pictorial information (linear perspective, size perspective, superposition, texture gradients, aerial perspective, shadows and highlights, color contrast, transparency, caricature, and arrangements of forms) be it with high-fidelity, colored, textured pictorial stimuli or impoverished outline drawings and black-and-white photographs. We have, rather, patches for a crazy quilt insufficient for piecing together. Two more such patches on contour fragmentation are mentioned before moving to a brief review of studies on perception of pictorial surface layout with impoverished information.

Gollin (1960) used line drawings to test for age changes in the amount of information needed in identification of outline drawings. His subjects were nursery school and kindergarten children. The younger children required more completion than the older ones for successful recognition, but even their performance was extremely good considering the fragmentation (the contour interruption) of the outlines. Even very young children seem to assume continuity where there is none. A careful structural manipulation of components can determine under what conditions such assumptions occur, what possible configurations of line segments give rise to the unification of the segments into a perceptual whole. The urge to do so must be quite primitive, perhaps because chaos and meaninglessness are unnatural or unperceivable.

Mooney (1957) studied closure ability in children aged 7–13 years using some extremely interesting stimuli. The stimuli were black-and-white drawings producing the effect of strongly lighted high-contrast photographs of heads and faces showing only salient shadows and highlights. He found an increasing ability with age to recognize and classify the faces by age and sex. This may reflect an improved ability to recognize higher order relations of shadows and highlights specifying faces. A more complete analysis of the stimuli and the configurations that facilitate recognition would be most desirable.

Surface Layout and Depth

There have been very few investigations of surface layout and depth in outline pictures and black-and-white photographs, and what work has been done is mainly of one type—studies of African pictorial depth perception, Hudson (1960, 1962a, 1962b, 1967) was the prime mover of a whole minor body of research. He devised the Hudson pictorial depth perception test, which has since been used repeatedly by himself and other investigators. The test consists primarily of outline drawings of three figures: a hunter, an antelope, and an elephant. Depth was depicted variously by relative size, overlap, and linear perspective.

> Responses to the questions whether the hunter was aiming at elephant or antelope or whether elephant or antelope was nearer the hunter were taken as self-evident indications of two-dimensional or three-dimensional pictorial perception [1967, p. 94].

The consistent result seems to have been that subjects did best whose cultural background and home life favored frequent exposure to pictures. Hudson thought the two-dimensional performance was due to misperceptions of the "conventional pictorial depth cues." It is impossible, however, to tell if nondepth perception was due to the inability to perceive depth information, the greater salience of the information for flatness, the general poor quality of the drawings, or the specific terms used in questioning subjects, as suggested by Omari and Cook (1972). A wider sampling of stimuli and testing conditions is needed to check the validity of the task as an index of the ability to see depth in outline drawings.

Two other points should be noted. Given the simplicity of the test, it is odd that no sample, white or black, young or old, produced anything near perfect performance. This result indicates that it may not be the subjects or the pictorial depth cues that are at fault but the depth portrayal in Hudson's pictures. Also, with black high school pupils and graduate teachers Hudson (1960) stated, "Hesitation in responding was noticeable and was particularly pronounced with the graduate teachers, some of whom took as long as one hour per picture to respond [p. 203]." One hour per picture to respond to two simple questions suggests that these black samples did not view the task in quite the same manner as Hudson did. Better communication between the experimenter and the subjects would have given the results greater credibility.

Deregowski (1968) repeated Hudson's test with Bantu adults and children in Lusaka. He also gave them line drawings representing wire figures to be constructed by the subjects. Even where subjects did not make three-dimensional responses to Hudson's pictures, they frequently constructed three-dimensional wire figures from the line drawings. Making the wire figures first improved performance on Hudson's pictures. These apparently contradictory results cast some doubt on the validity of Hudson's test as a measure of pictorial depth perception.

Mundy-Castle (1966) used the Hudson test with Ghanaian children 5–10 years old. Nearly all the subjects almost always gave two-dimensional responses. There were also many misidentifications of objects like the road, the horizon, the hill, and the elephant. Now clearly if the elephant is seen as a "pig, rabbit, goat, sheep, lion, tiger, dog" or other small animal, then all the resultant size–depth relations in the picture are altered. Depth in these pictures is not very determinate even *with* correct identification; *without* it the test is useless as a depth perception measure.

Another point, even granting perfect communication, perfectly drawn pictures, and errorless identification of objects, is that although one *can* draw in perspective, one *need* not. In Hudson's pictures, elephant, man, and antelope are all drawn in direct side views except that one cannot tell the front of the hunter's body from the back. Two converging lines are then inserted to give these cutouts relative depth. The point is not *must* two lines representing perspective always evoke depth, but *can* they? *When* will the subject see such lines as perspective lines, not *does* he always and *must* he regardless of other elements in the picture, to be considered a three-dimensional picture perceiver?

A pictured scene is a unit. Piecemeal manipulation of its parts will almost invariably result in odd distortions of the overall relations. A certain amount of impoverishment of information is required before such manipulation is even possible. (Witness the lack of texture and shading in Hudson's pictures.) Something must determine that the pictured road is a road, that the horizon line does represent the horizon before depth relations become determinate. The Hudson test fails to analyze the conditions for determinate picture relations. Six depictions of three depth cues simply cannot analyze which components or combinations under what conditions will act as sources of depth information. Dismissing the cues as culture-bound conventions does little to further understanding.

All of the investigators using Hudson's tests seem to deny the basic ambiguity of the pictures. For instance, Kilbride and Robbins (1968) used two of Hudson's pictures to test Bagandans' (Uganda) ability to use linear perspective as a pictorial depth cue. Each subject was asked for each road, "What is this?" A correct response (e.g., road, river, path, etc.) was interpreted as use of the linear perspective cue to pictorial depth perception. An incorrect response (e.g., hill, stone, ladder, letter "A", etc.) was interpreted as a failure to use the linear perspective cue for depth. This was *not* a test of Bagandans' ability to use linear perspective as a depth cue. It did indicate their inability to accept unequivocally two lines drawn on a card as the edges of a road. "Correct" response was the experimenter's response, not intrinsic unambiguous truth.

As a concluding note, Hudson stated that school-attending white children acquire adult responses to his pictorial depth perception test gradually between the ages of 6 and 12 years. The credibility of this statement is questionable as previously noted, although the developmental literature, with one exception (Stern, 1924), does indeed suggest a gradual development of the ability to pick up pictorial depth information. Stern (1924) found that his daughter Hilde, one and a half years old, would imitate actions represented in pictures and could interpret real size from the relative sizes of pictured objects. She and other less-than-two-year-olds could also recognize objects drawn in perspective. Bower's work, on the other hand, showed an insensitivity in infants to pictorial information for size-at-a-distance.

Bower (1965) conditioned infants 40–60 days old to respond to a white paper cube. Another group was conditioned to respond to a slide projection of that cube equal to the real cube in visual angle, luminance, and so on. There were three subsequent generalization conditions: a change in size, a change in distance, and a change in size and distance such that the retinal image was equal to that of the training stimulus. Infants trained on the real cube discriminated all three conditions from the training stimulus (or at least responded differentially). Infants trained on the slide responded differently to the change in size and to the change in distance than they did to the training stimulus, but they failed to discriminate the large far cube, projecting a retinal image equal to the training stimulus, from the training stimulus. That is, they responded to measurable area on the projection screen and not to portrayed size. Pictorial cues present but insufficient for making the discrimination were: (*a*) relative height on the projection screen, (*b*) density of table texture at point of cube contact, (*c*) density of wall texture occluded by the cubes, and (*d*) linear perspective of the table to point of cube contact. Bower concluded that motion parallax is the basic source of size-in-depth information for babies and that sensitivity to pictorial cues comes in later. This research does not tell us that an infant cannot recognize a picture of its mother, bottle, or other familiar object, but it does suggest that a mother pictured near is not the same as a mother pictured far, that cube pictured near is not seen as identical to the same cube pictured far.

Wohlwill (1962) tested children in grades one, four, and eight and adults on line drawing perspective displays. He presented them with rectangles at different pictorially defined depths. The children adjusted the "near" rectangle larger than the "far" rectangle to achieve an apparent size match and vice versa. Even the youngest children were sensitive to perspective depth information. Wilcox and Teghtsoonian (1971) tested children three and nine years old and adults. They placed equal area figures at different pictorially defined depths on slides to produce the illusion of a "large" object "far back," that is, high up on the picture plane. They used everything from outline drawings like the Ponzo illusion to magazine illustrations. Three-year-olds showed no responding above chance to the illusory larger object; nine-year-olds showed 75% responding and adults, 90%. This suggests that insensitivity to pictorial cues such as found by Bower in babies may persist beyond the age of three and still have some effect on adult performance. Wilcox and Teghtsoonian failed, however, to test for the effects of varied motion parallax information, and their data may show not the ineffectiveness of pictorial size cues but the effectiveness of motion parallax as a cue to flatness. They also failed to vary and control for the relative effectiveness of the cues (texture density and regularity, superimposition, and relations among various cues or components) employed in the drawings and photographs. Also, the geometric pairs may have been perceived as floating in space rather than resting on a surface because they were flat figures without three-dimensional form or attached shadows. The failure of adults to respond 100% correctly to pictorially defined size in depth may also suggest that there was something wrong with the slides. Because the figures were pasted on the slides, they may have looked in some way "unnatural."

Summary

The studies of recognition of pictured objects in outline drawings and black-and-white photographs by the untutored overwhelmingly suggest an unlearned responsiveness to impoverished representations. This supports Gibson's and Kennedy's arguments on the fundamental importance of edge information for perception of surface layout in the natural world and in pictures. The evocative nature of line drawings and black-and-white photographs of objects has been demonstrated, but we still lack a systematic attack on the sources of the evocations, on the arrangements of lines and contrasts that give rise to determinate perceptions of specific representations.

For overlapped, embedded, and fragmented figures the picture is more confusing. Infants do not seem to complete interrupted contour with surface occlusion. Nursery school children are quite good at completing interrupted contour without surface occlusion. Four-year-olds successfully perceive figures through overlapping transparencies (intersected, noninterrupted contour without surface occlusion), and between the ages of 7 and 13 there is an increasing ability to employ shadow and highlight information specifying faces. The need for programmed research is obvious. The preceding research gives only the barest hints about the evocative nature of various line arrangements and the development of responsiveness to such arrangements.

Research on perception of surface layout and depth with outline drawings and black-and-white photographs similarly gives little ground for a firm theoretical stand on the evocative character of the line. The African work, in general, suggests a gradual development of the ability to see depth in such pictures, but the test and testing conditions as previously discussed are too questionable for either theory testing or building. The Western data suggest that sensitivity to linear and size perspective information is absent in infants, present in children less than two, absent in three-year-olds, and thereafter present but improving to 90% sensitivity in adults. Such conclusions are hardly warranted, however, given the state of the field, and as long as we lack a comprehensive task analysis and programmatic developmental research the answers will not be forthcoming.

Gibson treated the *developmental* problems of picture perception within the framework of naive versus pictorial attitudes in perception (1969b, 1971). Gibson stated that there is evidence to suggest that young children, animals, and prepictorial men do not notice the appearance of an object (an aspect or perspective at a single stationary point of view) or the perspectives of the environment. Children do not notice the appearance of the environment as a frozen patchwork of flat colors confined by the boundaries of the temporary field of view. "Instead, they notice only the crude distinguishing features of objects that are given by invariants of transformation in time [1969b, p. 9]." By the information carried in the light from the structured world, children see whole objects, not aspects or appearances of objects, and their relation to one another in the surface layout of the visual world. This is the naive attitude.

The pictorial attitude, the ability to see the appearances of a frozen world, perspective on a two-dimensional plane (not depth in a three-dimensional world) is a later development for children and for man. When man began to draw his world he had to learn to see its two-dimensional appearance; likewise, when a child begins to draw he must *learn* to draw, learn how to see his world pictorially as a flat bounded surface.

If all of the above is true, as Gibson would have us believe, what are the developmental implications of such a theory? If it is assumed for the moment that the child must learn to see appearance, to see the world as flat and stationary, in order to draw or paint that world and that adults can switch between the naive and perspective attitudes once taught the latter, what does this say to us about children's ability to pick up information in pictures for surface layout *before* they have learned to draw? What man must do to create pictures may say very little about what he must do to perceive them. Gibson's theory is not, however, a model for information pickup. His chief endeavor has been to describe what is "pickupable," not how it is done. That is, Gibson was concerned with describing the structure of the light to the eye, but the extension of his theory into models of the pickup mechanism involves consideration of the structure of the head. Because Gibson himself has not taken that step, the problems arising from it do not lie at his door. Nevertheless, there are some problems and they deserve brief mention because they must eventually be considered by any coherent theory of picture perception.

The theory of structural equivalence between the picture and the depicted is quite promising (although preliminary) in its attempt to describe and account for the basis of adult picture perception but is less successful in describing the development of such perception in the child. Once the invariants have been detected, once the distinctive features of real objects have been discriminated, then the basis of pictorial perception has been established; but how does the child acquire sensitivity to such information or features? How many times and from how many different perspectives must a child see a particular cube, not to mention how many different cubes, before he can recognize any member of the family of cubes from any perspective in a picture? What are the constraints on the ability to do so and do such constraints change with age? Are distinctive features of classes of objects hierarchically salient? Is there a developmental change in the position of features in such a hierarchy? Are depth cues hierarchically informative, that is, can they be ranked according to the strength with which they determine the perception of depth relations in a picture? Are their effects in combination different from their effects when depicted singly? Are redundancy and contextual determination diminishing needs with experience? How do these two factors interact with the necessity of representing an object or scene in its best or most characteristic aspect? Is it true that any representation can be disambiguated with sufficient pictorial context, perception-determining arrangements of lines and forms?

We may also need to posit some sort of nested hierarchy of invariant relations. That is, there must exist "knowledge" at some level of how the invariant relations specifying cube differ from those specifying rectangle, sphere, and so on. When do such discriminations and their interrelationships become clear to children? In order to argue that the untutored perceive representative art successfully because of its structural equivalence to reality, it is necessary to show some basis for assuming their sensitivity to the prerequisite structural information in the world. We need carefully controlled developmental studies of sensitivity to structural features of the natural world via motion-generated information, sensitivity to those same real world features in the static case, and sensitivity to that static information isolated and manipulated in pictures. Some preliminary work along these lines is now reviewed.

The Coexistence of Surface and Depth Information

The problems of distorted and impoverished information are of a particular order. They deal with the *content* of a picture and its relation to the world. They attempt to answer the question of how the observer knows what is being represented via the content of the vehicle of representation. At this level the central issue is the equivalence of the structural information in the light to the eye coming from the world and from *within* the picture. But the problem of the ways and means of representation presupposes a prior level of the act of representation. To say that a picture is a framed piece of canvas or similar surface with paint or similar markings on it that represent something presupposes the knowledge, on some level, that A *can* represent B, that a picture *can* represent the world, regardless of the means of representation. When and how does the child come to know that object A is a particular object representing another object B (or an arrangement of objects, a segment of the world)? When is a picture accepted as anything other than a framed surface with a flat array of colors and forms? It is clear that at this point the two levels of the representative act, (*a*) knowledge that A can represent B and (*b*) A represents B by means of similar information, become so entwined that the explication of one is dependent on the explication of the other, at least for pictorial representation.

Perception of the visual world occurs by means of information. If pictorial representation occurs by means of that same information, in what sense must knowledge of the possibility of pictorial representation occur prior to the successful perception of the content of pictures? In other words, if the two sources of information (the world and pictures) are identical, then no act of perception of representation is necessary, but only another act of perception of the real world. In this sense, a knowledge of the possibility of representation need not be postulated prior to the perception of pictorial content because no such knowledge is necessary for perception of the real world, and the two types of perception occur by means of the same information. There is, however, one major difficulty with this dismissal of the problem of representation. The formulation only holds for the perfect picture observed under perfect conditions of observation. We can modify the demand for perfection a little, as Smith showed with his photomural of a corridor in 1958, and the result is almost the same. Observers perceive not the picture but the pictured world. No knowledge of the *possibility* of representation is required because the *presence* of representation cannot be detected. In such a case, from the point of view of the observer, presentation, not representation, occurs. Thus we need posit no prior knowledge of the possibility of representation. Gibson (1951) stated,

> When it is carefully arranged that the picture is seen through an aperture so that the frame is invisible, the head is motionless, and only one eye is used, *the resulting perception may lose its representational character* [p. 406].

Difficulty arises, however, because this case, the perfect picture under perfect conditions of observation, is the nonexistent case except in experimental laboratories. Thus, its usefulness is somewhat limited in explaining other cases of picture perception. The perfect case does, however, highlight the intertwining of the two levels of representation, possibility and means. Under the perfect case, the means obviated the necessity of postulating prior knowledge of possibility.

In less-than-perfect cases, the less-than-perfect picture viewed under less-than-perfect conditions of observation, the means of representation become far more complex than straightforward duplication of real world information. As previously noted, many components of pictures bear an obvious structural relationship to the world they represent—for instance, color, texture gradients, linear perspective, aerial perspective, superposition, relative size, and arrangements of light and shadows, not to mention the little understood relationships among the several components of a single optic array. Our observer, however, is now holding the picture in his hands, turning it over, and viewing it with two eyes from the wrong station point. Real world information must in some sense be duplicated in this picture, but there are a good many other sources of information specifying the picture only as a rather flat, perhaps framed, surface with splotches of color on it. As Gibson (1951) stated, "Pictures *stand for* substantial objects in addition to *being* substantial objects [p. 409]."

How pictures stand for substantial objects, how they carry information for the real world, has already been discussed extensively. Information for a picture being a substantial object comes from many sources. Motion parallax tells the observer the picture is flat thus: For a picture held upright in the frontal–parallel plane, the entire picture translates at the outside edge of the frame across the extrapictorial background as a unit, and there are no differential rates of translation of objects across the background within the picture (no gradients or discontinuities of optical flow). Stereopsis and information for the surface of the plane of projection (e.g., the texture of the canvas) also tell the observer the picture is only a two-dimensional surface with a pattern of varying luminosity or pigmentation. The texture of the canvas, projection screen, and the like, the mismatch of lighting between slides and the surround, and the mismatch of the picture surface with the surrounding surfaces all are sources of information for the flatness of the picture.

Herein lies the central problem of recognizing that a vehicle of representation is possible. Because both information for the picture as a rather flat object like any other object and information for the pictured scene coexist in a picture, can one or the other be ignored or unperceived? If the latter is perceived, then perception of the pictured content takes place; if the former is perceived as well, then perception of a picture as an object that can act as a vehicle of representation also takes place. Picture perception is perceiving that "this is a picture of a lake with swans up close and sailboats far away." It is *not* mistaking the picture for a real lake or seeing only a framed blue surface with white patches on it. If this is true, then picture perception demands response to both sources of information, and such a response entails knowledge of the possibility of representation even if that knowledge first occurs only upon the first presentation of a picture to the naive observer.

Granting, then, the presence of coexisting and conflicting flatness and depth information in pictures, the question of point of observation, or station point, must be raised to rescue the observer from the horns of this dilemma. Pirenne (1970) treated this question explicitly. He observed that most ordinary pictures are almost invariably viewed from the wrong station point, that is, a point of view different from the correct center of projection of the picture or photograph. Nevertheless, observers successfully perceive the pictured objects and relations of objects. In such cases, the linear perspective contained in the picture is no longer duplicating the light coming to the eye from a similar real scene. If perspective remains a valid effective source of information even when the observer is at the wrong station point, it must do so because of hitherto unexplained higher order relations between the views from the right and wrong station points. Besides the theoretical deformations that should result from improperly observed linear perspective, Pirenne also noted that at a station point sufficiently different from the correct one, the picture flattens. The scene depicted no longer appears in three dimensions. His example, however, was a painted church ceiling at a considerable distance from the spectator. With an ordinary picture the station point disparity needed to flatten the scene would necessarily be much smaller. Kennedy (1971) argued that given a knowledge of the correct station point, the transformation of the optic array, its difference from the standard, could be mathematically determined, although the mathematics would not be trivial.

The implications for perception for an observer at the wrong station point are far from clear, yet it is probably true that the wrong station point is the most frequently occupied one. Can such perception really be unlearned? Can it be said to follow naturally from some as yet undetermined relation between the right station point and all of the wrong ones? Can the one-eyed observer even *find* the right station point and if not, why not? Does an observer notice the deformations and flattening noted by Pirenne (1970)? If the validity of linear perspective is explained in terms of the properties of light and the laws of projection, does this explanation hold when a fundamental axiom of the picture projection system is violated (i.e., the observer not at the correct center of projection)?

Smith (1958), using a photomural of a corridor, found that adult subjects could estimate the number of paces from their viewpoint to a specific part of the pictured scene and the number of paces between different places in the scene. The scene was viewed at two distances, two feet and nine and one half feet from the eye of the subject, monocularly.

> When the photograph was viewed at a distance of 9-½ feet, the corridor appeared to be over twice as long on the average as when the photograph was viewed from a distance of 2 feet [p. 81].

Smith did not discuss the relation of these distances to the correct station point, but the magnification-minification effect suggests they were on the normal from the correct station point to the picture surface. Smith's study was the beginning of an answer to the question of what happens to pictured relations when observed from the wrong station point.

Smith and Gruber (1958) attempted a more systematic analysis of the magnification–minification effect of changing the station point on the normal previously defined. They found that when the photomural was viewed from the correct station point, impressions of depth were approximately the same for the photograph and the real corridor. At other viewing distances apparent depth was reduced by approximately one half when the image was magnified two times. That is, as the viewing distance increased, the image grew smaller and the perceived distance greater, in a geometrical relation.

Farber (1972) investigated the effect of angular magnification on sequences of rigid optical motions, and demonstrated that such magnification may produce consequent nonrigid sequences. The simple magnification of information alters the nature of that information, in this case.

We still need information, however, on the effects of wrong station points other than those on the normal from the correct station point to the picture plane. Given a sufficiently disparate station point, pictured forms should appear distorted and the scene depicted should flatten. Yet when most spectators view most ordinary pictures, neither of these events occurs. Pirenne (1970) suggested an "intuitive process of psychological compensation" to account for this. "[Ordinary] pictures produce an illusion of a very particular kind, to which we become accustomed as part of our education [p. 162]."

> Leaving aside such exceptional cases as Pozzo's ceiling, peep-holes, and ordinary trompe-l'oeil paintings (in which invariably the subject matter has very little extension in depth) ordinary pictures viewed binocularly in the usual manner do not give a genuine three-dimensional representation of their subject matter.... In spite of the fact that it may give a very strong *suggestion* of depth, the representation given by ordinary pictures is of a *sui generis* nature, and depends on the spectator's subsidiary awareness of the characteristics of the picture surface [p. 166].

The intuitive process of psychological compensation "is based both on the spectator's awareness of the surface of the picture, and on his preconceived ideas regarding the components of the scene presented [p. 162]." What Pirenne meant by preconceived ideas about the picture's components is unclear, but Polanyi clarified the notion of surface awareness in his introduction to Pirenne's book. Polanyi made a strong distinction between trompe-l'oeil paintings and ordinary paintings. In tromp-l'oeil art there exists no subsidiary awareness of the canvas, of the surface of the painting, no knowledge of representation enters in. The distortions and flattening are easy to observe from the wrong station point. There is little flexibility in the conditions of observation. Because a trompe-l'oeil painting is regarded not as a painting, not as a representation, but as the "real thing," distortions are striking when obtained views violate expectations. With ordinary paintings, there is no such illusion of reality.

> The trompe-l'oeil artist succeeded in producing a compelling illusion of a three-dimensional scene; whereas, however effective the perspective of a normal painting may be, the picture will not be

> mistaken for the sight of its real objects. Its appearance has depth and this is an essential feature of it, but this depth is not deceptive: it is felt to have flatness in it. It has the quality of a mixture of depth and flatness [p. xvii].

The Pirenne–Polanyi thesis is that ordinary pictures provide a very special kind of illusion. They provide information for both flatness and depth, an unusual duality to which we become accustomed through education and experience. Successful perception of the depth relations in a picture is dependent on a subsidiary awareness of the picture's surface, on the flatness information, and on the perception of a representative event.

Developmental Utilization of Coexisting Surface and Depth Information

This definition of picture perception, however, may well describe only an end point and bypass developmental steps on the way. The flatness and the depth information in a picture may be considered to jointly determine the perception of a picture as an object that represents another object or scene, as suggested by Pirenne (1970) and Polanyi (1970), but they may also be thought of as potentially in conflict during the development of picture perception. E. J. Gibson (1969) suggested that the discrepancy between the pictured scene information (in depth) and the information for flatness given by the absence of this motion parallax (gradients and discontinuities of flow velocities) in pictures is ignored by the sophisticated viewer who attends to the former for the perception of the pictured scene and ignores the latter except to perceive a picture not a replica.

All of the previously cited studies failed to take into account the coexistence in a picture of flatness (or surface) and depth information. Yonas and Hagen (1973) investigated children's sensitivity to pictorial depth in a replication and extension of the work of Wilcox and Teghtsoonian (1971). An attempt was made to test E. J. Gibson's contention that the flatness information in a picture must be minimized for successful perception of portrayed depth by the young or untutored. Following E. J. Gibson's suggestions, Yonas and Hagen compared the performance of three-year-olds, seven-year-olds, and adults on judgments of relative size in photographic transparencies, under two different conditions of observation. Half of the subjects viewed the slides in a condition of potentially conflicting information. The slide edges were visible and the subject's head was freely moving; thus, motion parallax information for the flatness of the slide was available. The other half of the subjects viewed the slides under restricted conditions of observation designed to minimize conflict by decreasing the information for flatness available. The slides were viewed through a peephole, thus eliminating head motion, and through a window placed between the peephole and the screen, so that it cut off view of the slide edges. Thus, in this condition there was no motion parallax information for flatness.

Adults and three-year-olds showed no consistent difference between the two conditions. Only seven-year-olds showed a consistently inferior performance in the conflict condition with free head movement. It may be that our manipulations failed to sufficiently diminish the conflict. It is true that even in the nonconflict condition, the slides looked like slides, not real objects in depth. Our results with the nonconflict slides should have paralleled those obtained in the control condition employing real objects viewed through a peephole. They did not. At all ages and with all visual angle relations pictured, performance on the slides was worse than performance on the real object control. There seemed to have been sufficient information for flatness even without motion parallax. It should be noted, however, that although performance on the slides did not depend on the presence or absence of motion parallax information, it did depend very strictly on the visual angle relations portrayed. As the visual angle size relations of the two objects pictured went from equal to the reverse of the absolute size relations, subjects increasingly responded to the visual angle relations rather than to the real size relations. This indicates an increasing tendency to react as if the slides portrayed only a flat surface without depth. Subjects increasingly responded to the amount of area on the screen occupied by the photographed object. Still, no age group's performance ever reached a perfect error rate. Subjects always seemed to respond to the real size differences as well as to the visual angle relations, perhaps vacillating from trial to trial. As the visual angle relations became more drastically the reverse of the real size relations, performance became more and more under the control of the former.

Apparently then in size judgments in slides two things are going on. One is the effect of the visual angle relations portrayed, and the other is the conflict between depth information and flatness information. The two effects can be factored out by comparing performance on slides with that obtained in real object controls, with and without head motion. Our results suggest that it is possible to produce a continuum of responding increasingly controlled by the visual angle relations presented rather than the real size of the portrayed objects. This result may not obtain with binocularity.

A better setup is needed to control for and diminish the amount of information available for flatness in slide viewing. E. J. Gibson (1969) might be right in arguing that one can induce good pictorial depth perception in young children by minimizing flatness information, and we simply failed to reduce the conflict enough to test her hypothesis. On the other hand, Pirenne (1970) might be right. Back-projected slides on a glass screen may offer few surface cues while still providing both flatness and depth information. At the right station point, an observer may not require an intuitive process of psychological compensation for distortions, but he may still need to be in a pictorial or representative mode of perception. If it is necessary to acquire knowledge of the peculiar kind of illusion presented by pictures, then it must be necessary to know when to apply that knowledge. Knowing when may depend on awareness of the picture's surface. That is, a perception of flatness may trigger a compensation for flatness. Back-projected slides may provide an insufficient source for this requisite information, leaving observers torn between the two types of information presented. In Yonas and Hagen (1973), the surface information present was ambiguous. There was too much information to parallel a trompe-l'oeil or a real scene and too little information to trigger consistent perception in a pictorial mode. The surface of a slide looks both insubstantial and present.

The amount of information for surface available may be scaled from least to most in three situations: (*a*) trompe-l'oeil art paralleling real scenes, (*b*) slides viewed under ordinary conditions, and (*c*) photographic prints. E. J. Gibson (1969) and Pirenne (1970) agreed that picture perception is a learned ability. Gibson noted that the sophisticated viewer attends to the information for depth and ignores the flatness information. Thus, for adults in this culture she should predict no difference in performance in situations varying in amount of surface information available. The performance of young children would presumably depend on their level of sophistication with pictures. At very young ages, performance should depend on the amount of discrepant information available: Performance with trompe-l'oeil art should be better than that with ordinary slides, which in turn should be better than performance with photographic prints.

In Yonas and Hagen (1973), however, adult errors reached 25% with certain types of stimuli and 10% with others. Adults were conflicted. They were clearly not ignoring all the flatness information or their slide performance would have been error free, as it was in the real object control condition. E. J. Gibson (1969) overestimated the abilities of the sophisticated adult, at least when the pictures in question are slides. She might be right, however, about the effectiveness of minimizing flatness cues to enhance depth information. Successfully carried out, such manipulations would produce a trompe-l'oeil situation that would be conflict free. It is questionable, however, whether partially successful manipulations would perform a similar function to a lesser degree.

Pirenne (1970) would probably argue that between trompe-l'oeil art and prints or paintings exists a conflicted limbo into which slides would surely fall. Surface awareness is present but in a ghostly form. Spectators respond neither as to trompe-l'oeil nor as to an ordinary painting, both of which courses would lead to perfect size-in-depth relation responding. Rather, observers vacillate between the flatness information and the depth information. Their vulnerability to flatness cues in this conflicted situation presumably decreases as their experience with pictures increases. Thus, children should respond to the flatness information to a greater degree than the adults, as indeed they do. Even adult performance, however, should be conflicted since there is no unambiguous information present in a slide to tell them how to respond. After all, given two present but conflicting sources of information, something must determine which receives a response.

An investigation was designed to test the importance of awareness of the pictorial surface and of position of station point in accordance with the Pirenne–Polanyi theory of artistic perception. As previously argued, photographic transparencies provide a special kind of pictorial depth ambiguity. The projection surface is neither completely imperceptible, as in trompe-l'oeil setups, nor is it obviously perceptible, as it is with photographic prints or paintings normally viewed. The representative quality of slides may be ambiguous. Pirenne and Polanyi would predict that adult performance improves with an obvious picture surface.

Since a fairly regular scaling of surface information from trompe-l'oeil to ordinary slides to photographic prints was desirable, an initial study was designed to test the effectiveness of a trompe-l'oeil slide-viewing apparatus. It was postulated that true trompe-l'oeil responding duplicates responding to a real frozen scene viewed through a peephole. Hence, the study was designed to see which condition in Yonas and Hagen (1973) predicted the results in the proposed trompe-l'oeil apparatus. In Yonas and Hagen adults made no errors when viewing real scenes through a peephole. When viewing slides through a peephole, they made 3%, 12.5%, and 27% errors on the equal visual angle, 80% visual angle, and 70% visual angle slides, respectively.

This study was not a replication of Yonas and Hagen since the stimuli were colored and the background was regular, not random. The stimuli were also more numerous and a greater variety of sizes and distances were used. It was felt, however, that the conditions were sufficiently similar to warrant comparison. In the trompe-l'oeil study, adults made 3.9% errors on the equal visual angle slides, 28.9% errors on the 85% reversed visual angle slides, and 39.8% errors on the 70% reversed visual angle pairs. The results do not in any way suggest the error-free or low-error rate responding expected in a trompe-l'oeil condition. The control by visual angle relations apparent across all ages with slide viewing in Yonas and Hagen is evidently operative in the trompe-l'oeil study. Accordingly, the attempt to devise a trompe-l'oeil apparatus was abandoned for the present, and the attempt to scale surface information available was confined in the next study to a comparison between ordinary back-projected slides and photographic prints.

The task was similar to that used in Yonas and Hagen. Subjects were asked to choose the larger of two objects of similar shape photographed against a textured background. Squares were paired with squares and triangles with triangles. The objects were positioned against the background at distances yielding three types of visual angle relations between members of a pair: (*a*) pairs in which the visual angles of the two objects (large and small) were equal, (*b*) pairs in which the larger object had a visual angle only 85% that of the really smaller object, and (*c*) pairs in which the visual angle of the larger object was only 70% of the visual angle of the really smaller object. There were 16 slides and photographs of each kind of pairing. The subjects were four-year-olds, seven-year-olds, and adults.

The study was designed to test the development of sensitivity to the point of observation and to look for developmental evidence of a pictorial mode of viewing triggered by station point and surface information. If perspective is a valid cue to distance, to what extent is its effectiveness constrained by the artistic rules for observation? If linear perspective can be successfully used irrespective of the correct station point, then its effectiveness must be dependent on the perception of higher order invariants, such as gradients, independent of their direct projection to the

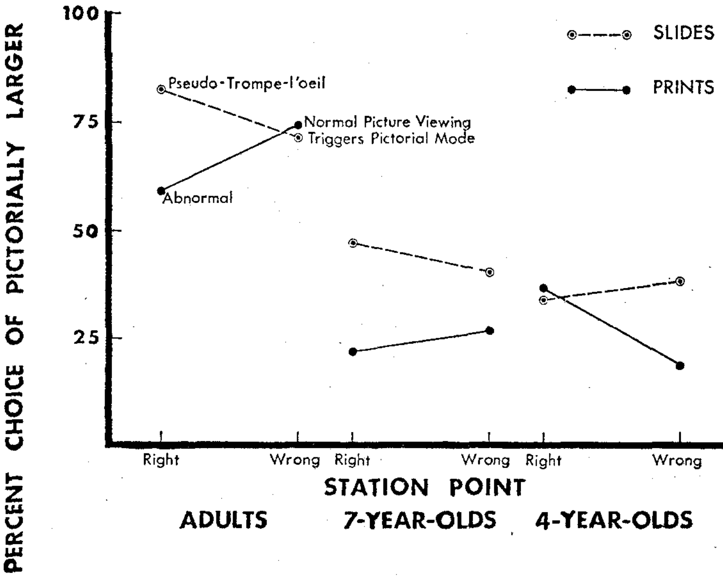

FIGURE 1. The interaction of age with surface information and station point.

eye. This perception must be learned since perspective views of the real world must necessarily be projected to the eye and to no other point. Only in art does a visual scene have a center of projection other than the eye. To test station point sensitivity, the test displays were viewed both from the correct station point and from an incorrect station point forty degrees to the left of the correct station and at the same distance from the screen.

The Age × Surface × Station Point interaction reached significance only at a fairly low level ($0.05 < p < 0.10$), but its trends are highly suggestive of hypotheses for future work (see Figure 1). Adults seem to treat photographic transparencies and photographic prints in very different manners. Slides viewed from the correct station point, although not identical to trompe-l'oeil presentation, nevertheless induce the best pictorial depth responding in adults. The situation seems to provide enough trompe-l'oeil information so that no pictorial mode need be triggered nor flattenings compensated for. This is not, however, the normal experience of viewing pictures. Pictures, both slides and prints, are normally viewed from the wrong station point. When a picture is observed from the wrong station point, the observer may be tripped into a pictorial mode of viewing and the compensation for incorrect station point misprojections triggered. Thus, the data suggest that the adult viewers treat both slides and prints the same way when observing from the wrong (the normal) station point. The data also suggest that the size-in-depth perception of adult viewers in normal pictorial situations (wrong station point) is inferior to that obtaining with pseudo-trompe-l'oeil representations (right station point, minimal surface information). Likewise, prints viewed from the correct station point provide the adult observer with an abnormal situation. Size-in-depth responses are near chance since he is consistently in neither the pictorial mode nor the trompe-l'oeil (real world) mode of perceiving, and his performance reflects his indecision.

Overall, four-year-olds responded to pictured depth only about a quarter of the time and their pattern of responding was very different from that shown by the adults. Four-year-olds responded to the pictorial size-in-depth in slides and prints at the correct station point and in slides at the incorrect station point all at approximately the same low level. Although station point did not affect size-in-depth performance with transparencies, depth responding was considerably diminished when prints were observed from the incorrect station point. Four-year-olds, unlike the adults, cannot use the station point cue and surface information to switch a pictorial mode of viewing and thus are increasingly hampered as more and more flatness information is added by introduction of observation from the wrong station point and of information for the texture of the paper surfaces of photographic prints.

The pattern of results for the seven-year-olds lies between that of the adults and the three-year-olds. There is a suggestion of a developmental change in susceptibility to both flatness of surface information and shift in station point. As with the adults, slides at the correct station point induce in seven-year-olds the greatest number of correct size-in-depth responses, and the seven-year-olds' responses to depth in photographic transparencies is diminished by observation from the incorrect center of projection. Slides may offer insufficient surface information to trigger a compensation for the flattening of the picture observed from the wrong station point. With prints, however, the data suggest the beginning of a pictorial mode of viewing triggered by observation from the wrong, but most frequently occupied, station point. Seven-year-olds may be beginning to treat prints observed from the right station point as the abnormal situation, neither trompe-l'oeil nor in the normal pictorial mode.

Under observation from the wrong station point, projective distortions do occur and are readily observable in photographs of photographs. Why then are they not perceptible in ordinary viewing? Pirenne (1970) has argued that perception of the surface and components of the picture triggers a compensation for the perspective distortions. The present study concentrated on manipulation of available surface information and station point to test Pirenne's hypothesis, but the data are only suggestive of such a compensation mechanism. At least two explanations other than the compensation hypothesis may explain successful picture perception from the wrong station point. First, the observer may fail to notice the pictorial distortions simply because he cannot do so without the interjection of a pictorial medium; that is, pictures of pictures taken from the wrong station point and subsequently observed from the wrong station point. Perhaps there is simply enough elasticity in the visual system so that observers are unable to notice a station point switch. The trends in the data may simply reflect randomly different behaviors in different conditions across different ages. Perhaps, too, size judgments may be rather impervious to such a minor manipulation as a 40-degree station point change. Shape judgments or matching may provide a more sensitive index of the effect of station point through the theoretically present perspective distortions. If the system is sufficiently elastic, observers should be unable to notice even the shape distortions across a fairly wide range of station points.

Second, whereas the degree of elasticity just suggested may not be present in the visual system, it is certainly not unfeasible to argue for the redundancy of information in the components of the picture. It may well be that the perception of a variety of surface relations is so overdetermined that a misperception is unlikely even at the wrong station point. We don't yet know enough to describe the correspondences and codeterminations of the relations in a single optic array, but it seems highly probable that they affect perception in a way unavailable to single cues in isolation.

Summary

The Pirenne–Polanyi hypothesis of a pictorial compensation mechanism has been suggested to account for successful picture perception under the normal conditions of observation from the wrong station point and may, by extension, account for successful perception of modified linear perspective. Goodman's (1968) argument for the necessity of modifying linear perspective to achieve correct perception of pictured relations, however, is somewhat inconsonant with the suggestion of such a learned compensation mechanism. Who compensates? Is it artist, observer, or both? The Gibson–Kennedy argument for the power of "forms of forms," for determinacy in the arrangement of the pictorial components, may well dispense with the need for either postulated compensation for successful pictorial perception. Until the problems of form distortion and depth flattening have been systematically attacked by the pitting of one such theory or postulate against another, no resolution of the questions is possible. Gibson's "forms of forms" sounds, on the surface, simplistic because we presently lack a careful analysis of redundancy and codetermination in the frozen optic array. Yet such an approach may well relieve us of the need to hypothesize compensation mechanisms and bring the problem of pictorial perception back to an analysis of the structure of the light coming to the eye from the picture. Even Pirenne made brief note of the role of the observer's ideas about the components of the picture in overcoming distortions and flattenings. Pirenne's "preconceived ideas" about pictorial components may perhaps be better translated into Gibson's notion of the perception of the several codetermining aspects of the frozen optic array rather than into the syntax of Goodman's learned picture language. The arrangement of forms in an optic array (in a picture) may also ultimately resolve the dilemma of the fundamental ambiguity of any single projection, to be considered next.

Multiple Generation of Frozen Information

All pictures are fundamentally ambiguous because each is a single frozen perspective of an infinitely large possible set of visual scenes. In pictures there is no disambiguating motion parallax, no walking around to the other side to see. Because there is no possibility of seeing the other sides of objects in pictures, the identification of pictured objects is always necessarily ambiguous. The back of a pictured cube may well be concave, in which case it is no longer a cube. This seems a trivial ambiguity, particularly because it is also encountered in the real world, albeit with fewer constraints on checking the perception, but it is not. Successful perception of the form of an object from a single frozen perspective implies that perception can occur on the basis of a "best guess" made with only a subset of the features necessary to distinguish the true form of an object. A careful structural analysis of co-occurring features of objects and a contextual analysis of surface layout would help us account for such perceptual weights as *assumptions of regularity*. It would be helpful to know which aspects of a single perspective view of a shape, both in and out of context, give rise to assumptions about unseen surfaces.

Recall that any contrast in the optic array can be due to any combination of differential facing, reflectance, color, and shadowing. The source of a single contrast is ambiguous. That is, a cast shadow may look like a change in color, a cast shadow may look like a change of facing, a change in color may look like a change in slant, and so on. In the real world, gradients of optical flow velocities, occasioned by the texture of a surface as the observer moves, provide information sufficient to make the extension and slant of a surface determinate. Pictures, however, provide no such source of additional information since the motion perspective occasioned by the observer's movements is irrelevant to the perception of pictured surface relations. Indeed, as the observer moves in front of a picture, the single gradient of texture flow velocities tells him that the relation the pictured surfaces have to one another is one of coplanar adjacencies, a piece of information most unhelpful to the perception of pictured surface layout. How then can an observer know the sources of contrasts in a pictured optic array? As previously suggested, to resolve the problem of impoverished information in outline drawings and black-and-white photographs, we must (even with fully colored, textured pictures) fall back on such difficult-to-grasp, but no less valid, variables as arrangements of contrasts and forms of forms, or, in other words, the components of the picture.

The lack of motion parallax information in pictures also gives rise to theoretically ambiguous slant and relative depth. Kennedy (1971), in his exposition of Gibson, has pointed out that most surfaces are homogeneously textured, and a discontinuity of the textured light indicates a discontinuity of environmental texture, a change of surface.

> Hence, adjacent solid angles of optic texture in the optic array, that lie inside discontinuities of optic texture, will have come from adjacent surface areas of the environment [p. 8].

Thus, we know that a change in texture indicates a change in surface, but we do not know from the configuration of the solid angle alone whether the surface slants or remains at a constant distance from the station point. How can we tell a slanted cube from a truncated pyramid?

> While a particular surface area in the environment will project a particular solid-angle element to a given station point, the same configuration of that solid angle could be projected by different surfaces in the environment [Kennedy, 1971, p. 8].

Kennedy appealed to the gradient of a textured surface occasioned by recession of that surface from the station point to disambiguate slant. If a surface is homogeneously textured, then the farther a texture element is from the eye, the smaller will be its projection by the light to the eye. The texture of the surface as a whole will be packed more densely, and its borders will approach each other, in accord with a lawful gradient, as the surface slants away from the station point (Purdy, 1958). Goodman (1968) has taken issue with the determinacy of this information in an attack on artistic linear perspective. He pointed out that the same texture gradient can be observed and drawn when a surface rises in space or goes down in a hole, as well as when a surface recedes along the *z* axis. A texture gradient in a picture tells us the surface is at a slant to its plane of projection, but nothing in the texture gradient alone can tell the observer the slant with respect to the ground plane or with respect to the observer himself. Even if the observer views the picture from its center of projection, slant remains ambiguous without an assumption of congruence between his orientation and the picture's orientation.

A pictured surface is not just slanted with respect to the observer, resulting in the perception of "this is a surface at a slant with respect to me." Pictorial slants and arrangements of slanted surfaces must be specified as well by the information giving the location of the ground plane within the picture. Texture gradients alone cannot give this sort of useful information about surface arrangements in pictures without recourse to other components of the picture or assumptions carried by the observer. For example, Benson and Yonas (1973) found sensitivity in three-year-olds and adults to shadow and highlight information for concavity and convexity. The three-year-olds, however, showed this sensitivity only when the pictures were shown vertically, upright in the frontal–parallel plane. Adults consistently assumed the direction of illumination to be at the "top" of the picture, be it horizontally farther or vertically higher, whereas the three-year-olds assumed a source of illumination in keeping with a gravitational reference system; that is, from the sun or light above. The adult behavior is an excellent example of a learned assumption held by the observer. On the other hand, in the same study three-year-olds showed great sensitivity to linear perspective information for size-at-a-distance in Gibson's (1950) cylinder drawing, even when the picture was rotated 180 degrees, rendering the surface relations depicted most unlikely.

In the face of the difficulties with textured pictorial surfaces, the next point is rather trivial and has previously been discussed quite extensively. Whatever problems of ambiguity textured drawings suffer, the problems occasioned by the further impoverishment of outline drawings are far more numerous. That human beings so readily accept drawn lines as representing clear borders in the optic array is no less than astonishing. As Gibson (1951) remarked,

> an outline, representing as it does only the edges of a surface, may stand for any *object which projects that particular* outline, including some very queerly shaped surfaces [p. 410].

In the same article, Gibson reported that subjects' responses to simple outline figures almost invariably fell into one of two classes: (*a*) objects for which the outline form was a plan view, and (*b*) objects for which the outline form was a perspective view. Any theory of picture perception would benefit greatly from a theory of the pictorial components determining which type of perception occurs. If it is possible to set up an ambiguous situation, then it should be possible to disambiguate it in an eventually lawful manner. The type of ambiguity referred to here should more properly be called equivocality because the information is not strictly ambiguous but rather calls forth more than one perception.

CONCLUDING REMARKS

The position expressed in this article is basically in agreement with Gibson's new formulation of the order of structural resemblance between pictures and the world they depict.

As reasonable as Gibson's new definition of pictures seems, problems with it come immediately to mind. What is information in the world like? Is it static and frozen so that its replication in pictures is fairly simple, or is it dynamic, dependent on movement of the organism through its environment? It is possible to make a case for either position.

This article assumed, with Gibson, that motion-generated information is fundamental to the perception of surfaces and their layout in the world. If motion-carried information is fundamental, however, where does that leave the position on pictures as environmental information duplicators? One can talk of the static case as the limiting case for motion information, but it is not particularly helpful. One can also construct elaborate models of hypothesized mental structures to account for the pickup of frozen information. This article suggests that a more fruitful pathway for research is the systematic investigation of pictorial information in the Gibsonian terms of ecological optics. It is hoped that the feasibility of this enterprise has been demonstrated. The information analysis approach does not dispense with the problems of distorted or impoverished information, coexisting surface and depth information, or pictorial ambiguity, but it does suggest a specific line of investigation of their effects on perception. A pictorial mode of perceiving patterned after Pirenne's (1970) and Polanyi's (1970) suggestions was briefly considered, but it is hoped that any such hypothetical compensation mechanisms will remain only in the backgrounds of theories of picture perception, until the data from an exhaustive investigation of the information components of pictures demand the use of such constructs.

There seems to be insufficient information in pictures to result in unambiguous identification of pictured objects or perception of pictorial surface layouts. We know, however, that a 19-month-old child can recognize familiar objects in outline drawings and photographs without previous training and with very restricted exposure to pictures. We know that untutored Africans and young chimpanzees can do the same. We have only fragmentary information on the ability of the untutored to perceive pictures despite overlapped, embedded, and fragmented contour.

We know very little about the perception of surface layout and depth in pictures. The African research suggests that it depends on how much one's cultural background and home life favor frequent exposure to pictures. American studies indicated that adults are very good at estimating distances and otherwise responding to the depth relations portrayed. The African data is not such as one would wish to build a theory on, and the American studies have not gone far enough, especially in their failure to systematically study the development of pictorial perception in children. We know very little about adult performance on pictures with an eye to discovering which theoretically present ambiguities and problems influence responding and which fail to do so. We know almost nothing about children's growing responsiveness to pictures or about the development of the ability to recognize and use representations.

This article argued that picture perception is perceiving "this is a picture of a lake with swans up close and sailboats far away," not mistaking the picture for the real thing or seeing only a framed blue-and-white surface. That is, in order for a picture to be perceived as such, not only must it carry information duplicating the information in the real scene, but it must also carry information for itself as a picture. This moves the problem along a different avenue. We no longer want to know if pictures can fool people into thinking they see the real world, but how do pictures come to act as representations of the real world and by what means do they do so?

In order to successfully perceive pictures, the naive observer must accomplish several tasks. (*a*) He must learn to disregard the unnatural stillness of the frozen world presented to him by a picture and make the best use of the information remaining. (*b*) He must learn to make the best guesses about size, shape, and distance in the face of impoverished information or ambiguity, perhaps based on past experience with the world or, as Gibson would have us believe, on the codeterminance of the several components of an optic array (the "forms of forms") providing both specialized and redundant information. (c) The naive observer must learn to understand and to use certain conventions whose relationship to the world they represent is symbolic, without the physical correspondence of two similar structures (in the world and in the picture) modeling the light to the eye in a similar manner. (*d*) He must learn to disregard the station point with its attendant deformations and flattening. (*e*) He must accommodate himself to the coexistance of flatness and depth information in a picture either through suppression of the flatness information or through its use, as Pirenne suggested, as an aid in resolving the ambiguities, distortions, and flattenings of a frozen array observed from an incorrect station point.

REFERENCES

Arnheim, R. *Visual thinking*. Berkeley: University of California Press, 1969.

Benson, C., & Yonas A. Development of sensitivity to static pictorial depth information. *Perception and Psychophysics,* 1973, **13**, 361–366.

Bower, T. G. R. Stimulus variables determining space perception in infants. *Science,* 1965, **149**, 88–89.

Bower, T. G. R. The visual world of infants. *Scientific American,* 1966, **215**, 80–92.

Brimble, A. R. The construction of a non-verbal intelligence test in Northern Rhodesia. *Rhodes-Livingstone Journal,* 1963, **34**, 23–35.

Davenport, R. K., & Rogers, C. M. Perception of photographs by apes. *Behavior,* 1971, **39**, 318–320.

Dennis, W. The human figure drawings of Bedouins. *Journal of Social Psychology*, 1960, **52**, 209–219.

Deregowski, J. B. Difficulties in pictorial depth perception in Africa. *British Journal of Psychology,* 1968, **59**, 195–204.

Farber, J. M. *The effects of angular magnification on the perception of rigid motion.* (Doctoral dissertation, Cornell University) Ann Arbor, Mich.: University Microfilms, 1972, No. 73-7134.

Ghent, L. Perception of overlapping and embedded figures by children of different ages. *American Journal of Psychology,* 1956, **69**, 575–587.

Gibson, E. J. *Principles of perceptual learning and development*. New York: Appleton-Century-Crofts, 1969.

Gibson, J. J. *The visual world*. Boston: Houghton Mifflin, 1950.

Gibson, J. J. What is a form? *Psychological Review,* 1951, **58**, 403–412.

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

Gibson, J. J. Three kinds of equivocal information in line drawings. Ithaca, N.Y.: Author, September 1969. (a) (Memorandum)

Gibson, J. J. The psychology of representation. Ithaca, N.Y.: Author, October 1969. (b) (Memorandum)

Gibson, J. J. The information available in pictures. *Leonardo,* 1971, **4**, 27–35.

Gibson, J. J., Olum, P., & Rosenblatt, F. Parallax and perspective during aircraft landings. *American Journal of Psychology,* 1955, **68**, 372–385.

Gollin, E. S. Developmental studies of visual recognition of incomplete objects. *Perceptual and Motor Skills,* 1960, **11**, 289–298.

Gombrich, E. H. *The story of art.* London: Phaidon, 1971.

Goodman, N. *Languages of art: An approach to a theory of symbols.* Indianapolis: Bobbs-Merrill, 1968.

Gregory, R. L. *The intelligent eye.* New York: McGraw-Hill, 1970.

Hay, J. C. Optical motions and space perception: An extension of Gibson's analysis. *Psychological Review,* 1966, **73**, 550–565.

Hochberg, J. In the mind's eye. In R. N. Haber (Ed.), *Contemporary theory and research in visual perception*. New York: Holt, Rinehart & Winston, 1968.

Hochberg, J. E., & Brooks, V. Pictorial recognition as an unlearned ability: A study of one child's performance. *American Journal of Psychology,* 1962, **75**, 624–628.

Hudson, W. Pictorial depth perception in subcultural groups in Africa. *Journal of Social Psychology,* 1960, **52**, 183–208.

Hudson, W. Cultural problems in pictorial perception. *South African Journal of Science,* 1962, **58**, 189–195. (a)

Hudson, W. Pictorial perception and educational adaptation in Africa. *Psychologia Africana,* 1962, *9*, *226–239*. (b)

Hudson, W. The study of the problem of pictorial perception among unacculturated groups. *International Journal of Psychology,* 1967, **2**, 89–107.

Kennedy, J. M. Line representation and pictorial perception. Unpublished doctoral dissertation, Cornell University, 1971.

Kilbride, P. L., & Robbins, M. C. Linear perspective, pictorial depth perception and education among the Baganda. *Perceptual and Motor Skills,* 1968, **27**, 601–602.

Mooney, C. M. Age in the development of closure ability in children. *Canadian Journal of Psychology,* 1957, **11**, 219–227.

Mundy-Castle, A. C. Pictorial depth perception in Ghanaian children. *International Journal of Psychology,* 1966, **1**, 288–300.

Nadel, S. F. A field experiment in racial psychology. *British Journal of Psychology,* 1937, **28**, 195–211.

Omari, I. M., & Cook, H. Differential cognitive cues in pictorial depth perception. *Journal of Cross-Cultural Psychology,* 1972, **3**, 321–325.

Pirenne, M. *Optics, painting and photography.* Cambridge, England: Cambridge University Press, 1970.

Polanyi, M. Introduction. In M. Pirenne, *Optics, painting and photography.* Cambridge, England: Cambridge University Press, 1970.

Purdy, W. C. *The hypothesis of psychophysical correspondence in space perception.* (Doctoral dissertation, Cornell University) Ann Arbor, Mich.: University Microfilms, 1958, No. 58-5594).

Ryan, T. A., & Schwartz, C. Speed of perception as a function of mode of representation. *American Journal of Psychology,* 1956, **69**, 60–69.

Segall, M. H., Campbell, D. T., & Herskovits, M. J. *The influence of culture on visual perception.* Indianapolis: Bobbs-Merrill, 1966.

Shaw, R., McIntyre, M., & Mace, W. M. The role of symmetry in event perception. In R. B. MacLeod & H. Pick (Eds.), *Studies in perception: Essays in honor of J. J. Gibson.* New York: Cornell University Press, in press.

Smith, O. W. Comparison of apparent depth in a photograph viewed from two distances. *Perceptual and Motor Skills,* 1958, **8**, 79–81.

Smith, O. W., & Gruber, H. Perception of depth in photographs. *Perceptual and Motor Skills,* 1958, **8**, 307–313.

Stern, W. *Psychology of early childhood.* New York: Holt, 1924.

Wilcox, B. L., & Teghtsoonian, M. The control of relative size by pictorial depth cues in children and adults. *Journal of Experimental Child Psychology,* 1971, **11**, 413–429.

Wohlwill, J. F. The perspective illusion: Perceived size and distance in fields varying suggested depth, in children and adults. *Journal of Experimental Psychology,* 1962, **64**, 300–310.

Wolfflin, H. *Principles of art history: The problem of the development of style in later art.* (Trans. by M. D. Hottinger) New York: Dover, 1950.

Yonas, A., & Hagen, M. A. Effects of static and kinetic depth information on the perception of size in children and adults. *Journal of Experimental Child Psychology,* 1973, **15**, 254–265.

(Received August 10, 1973)