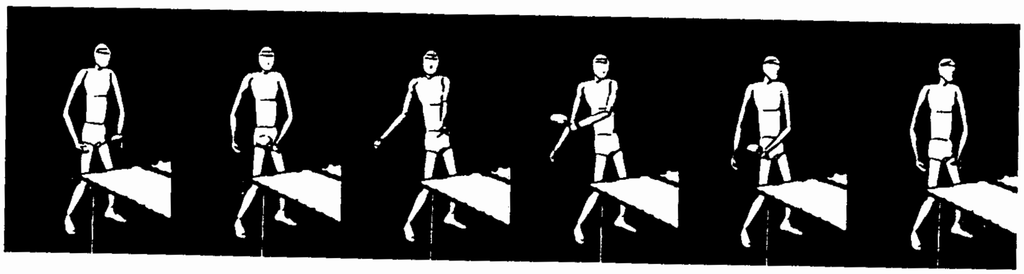
Moving Image Theory



Ecological Considerations

Edited by
Joseph D. Anderson
and Barbara Fisher Anderson

With a Foreword by
David Bordwell

2005

Southern Illinois University Press
Carbondale

2 The Value of Oriented Geometry for Ecological Psychology and Moving Image Art

Robert E. Shaw and William M. Mace

SCIENTISTS AND ARTISTS share the same environmental habitat (roughly, where they live) but occupy distinct, somewhat intersecting econiches (roughly, how they live). Although evolving within the same natural frame, their arenas of life are so dramatically different—the former tending toward the rational and the latter toward the expressive—that no easy comparison can be made of their methods or content. Yet, they have much in common. For instance, they have both made major contributions to the broadening of our culture of shared experiences. Such experiences are of two kinds: first, those that arise from *direct* perception of the environment, something all animals have in common; and, second, those that arise vicariously, as second-hand experiences, through *indirect* perception, or the use of substitutes for the real thing.

Historically, humankind has distinguished itself from other species by its attempt to produce a vision of nature—to produce records of that vision, with various degrees of fidelity and stylistic expression, to be shared and appreciated by others. Where art has pioneered our expressive side through poetry, dramaturgy, painting, sculpture, and music, among other things, science has advanced our rational side through basic research, theory, and technology. Milestones for both science and art were the discovery of various means for reproducing objects and events of general social interest vis à vis drawing, sculpting, painting, writing, printing, the telegraph, the telephone, photography, the phonograph, radio, movies, television, and computers. Drawings or paintings of people, landscapes, seascapes, or social events, such as sports, dance, travels, and trials, when framed and hung in a public place, become sources that capture some of the information contained in artists' once-personal experiences but which can now be shared publicly by many. Let's consider more carefully what this act of reproducing might entail.

We are so familiar with various forms of reproduction that we scarcely recognize what marvels they really are. Why do they work? There are two fundamental reasons: one having to do with intentionality, the other with causality. First, the very nature of one object, the *object of intention,* being in some way a reproduction of another object, the *object of reference,* is that the first refers beyond itself to the second. This is what is meant by the intention of the first being to refer to the second. *To refer* entails, at least, that when we perceive the first object, something about it formally resembles the reference object, and that thus in our experiencing the first, there is some part that would agree with our experiencing the second if such experiencing should occur. Second, there must be a causal basis for such intentional reference. But the nature of the referential relationship between the two objects is such that the absence of information about their causal connection does not mean absence of information about their intentional connection. Hence (and this is the main point) *the intentional entailment that exists between them is not solely dependent on the causal entailment.* Because this is so, a certain freedom for expressive variety exists for intentional entailment that is not allowed for rational entailment. The concept of ecological validity[1] is useful for comparing views of the world acquired through direct perception as opposed to indirect perception. To illustrate this concept, consider the following example.

A relief map of a landscape may show the lay of the land, the shape of forests, the meander of rivers or wiggle of streams and where their courses take them, relative to mountains and valleys; and, perhaps, it also shows virgin countryside where no houses or roads have yet been built. Later maps may show progressive variations in the topography after erosion has shifted the lay of the land, say, due to a forest being burned and tumultuous runoffs now allowed where before soaking action contained the water. Still later maps may show that houses and other buildings have cropped up since the addition of a major highway and its access roads have made commuting easier. Thus, the series of maps show what was, what is, and over successive differences, what transpired in the periods between cartographical perspectives. The series of maps offers graded records of a natural dynamical perspective unfolding over time, a historical event that can be causally explained by natural processes acting over the time samples.

Now imagine that someone accidentally shuffles the series of maps because their time-tags were lost. The differences between the maps would be out of causal order, with laws of nature appearing to be violated. Burned forests would sprout immediately full-replacement growth, and erosion creases would be inexplicably erased, as houses became dismantled and roads became covered by dirt, rocks, underbrush, and trees. No ecology could change in such a manner. Consequently, we would be justified in concluding that the properly ordered series was *ecologically valid* as a historical event because it conformed to the laws of nature. On the other hand, the improperly ordered one was *ecologically invalid* because the processes witnessed were unrealistic, being in violation of natural law. Direct perception of the landscape daily, say, by a forest ranger from a tower on a mountain peak or by a pilot whose plane flew daily over the landscape, could confirm the naturally ordered series but not the unnaturally ordered one. Direct perception has ecological validity because, in principle, it has direct access to confirmatory information while indirect perception may or may not. To ensure that indirect perception of the landscape over the series of maps was ecologically valid, that is, that it conformed to direct perception, one would have to supply the missing time-tags or some other means that would leave their temporal order inviolate.

One way is to label the maps with numerals and to give instructions that the maps are to be looked at in the order assigned, assuming that all who are allowed access to the maps understand the forward-counting convention. In this way we can see that a convention acts as a constraint that makes indirect perceptions conform to direct perceptions. Such conventions are required because perspectives taken on the world through indirect perception have more uncontrolled degrees of freedom than those taken through direct perception. That is to say, indirect perception allows for ecologically invalid information to be fashioned about an event even though the source-event is always ecologically valid, a fact that can always, in principle, under ordinary circumstances, be validated by direct perception.

There are, however, extraordinary circumstances where the conventional constraint is suppressed or unavailable. Under such circumstances, the indirect perceptual event can take on a life of its own. To the extent that such extraordinary circumstances defy rational (lawful) explanation, they serve to increase the mystery, metaphoric depth, and hence expressive power of the indirect perceptual event. This is one important way, perhaps the most important way, by which great arts attain expressive dimensions that surprise, challenge, and entice the viewer. Contrary to what some have argued (Goodman, 1968), art is not in this sense conventional like language but unconventional. Art is always lacking some degree of ecological validity because the expressive stylistics imposed by the individual artist are unique and defy rational conventions which would make the art object an easy read.

Of course, there is much that can be rational in art depending on how obvious its representational content; but there are also dimensions of expressive depth to be exploited by defying convention, as is found in the extreme in abstract expressionism—with impressionism falling between these poles of rational content versus expressive style.

The main goal of this paper is to show an example of extending the scope of lawfulness of projective geometry and thereby the basis for direct perception, but in doing so we also show a basis for controlled violations of ecological validity—available for use by the artist. To take liberties with the laws, one must know what the laws are and how to violate them skillfully so as to preserve some more-general constraint.

Contrasting Theories of Perspectives

Traditional theories of visual perspectives have been based on ordinary projective geometries. The technique of central projection is typically adopted without question as being the proper one for optics (e.g., photography), art (e.g., linear perspective), and psychology (e.g., retinal image theory). Here we wish to offer a glimpse of another theory of projective geometry that promises a simpler and more accurate description of visual perception and, perhaps, will have more potential usefulness for photography and art. By describing this alternative, *oriented* projective geometry, we mean to bring underlying geometry into focus as part of what can be tested and modified in the course of our science. Sometimes, it appears that researchers take projective geometry to be given and unmodifiable, leaving hypothesis formation and testing to be about tricks and assumptions for applying the geometry rather than revising the geometry itself. The emphasis of ecological psychology on lawfulness leads us to look to modifying theory as deeply as possible in order to minimize arbitrariness in the hypothesized system. It is not uncommon or unreasonable to regard limits on geometry to indicate limits on lawfulness. If we can extend the reach of geometry, we may justify a broader scope for lawfulness. We will caution that oriented geometry does not have all the properties we ultimately seek in a geometry but that it offers an advance worth making.

Projective theories have many practical uses in both art and science—a most important one being to model linear perspective in drawing and graphical computation. A second popular use has been to model the optical projection of objects and scenes observed in the world into the visual system. Traditional theorists treat the optical projection of the retinal image as a putative first stage (p_1) in visual processing. A second neurological projection (p_2) over the optic nerve tract and past the optic chiasm eventually reaches the visual cortex. And, finally, a third phenomenological projection (p_3) takes the cortical information into a visual experience in some way still not fully understood. Under this view, the retinal image is the first and most primitive site containing the visual information to be projected and, perhaps, cortically processed before being experienced.

The fact that, geometrically speaking, the retinal image is a two-dimensional object representing three-dimensional objects and scenes has posed a perplexing puzzle for the traditional perceptual theorist. How can we recover the third dimension from a two-dimensional image? This has been called the *tridimensionality problem*. If it were possible, however, to render the retinal image superfluous as a stage of processing, the main issue would then be how *information* gets into the visual system, without worrying about the specific properties of the retinal image. We could then move our theoretical concern to the second stage of projection described above without further ado.

In traditional psychological terms, we would say that the distal object was the referent rather than the proximal object (i.e., the retinal image). The issue would be *not* what image is projected but *how* the information about the object remains invariant under such projection. The optical physics connecting the distal referent to the eye dynamically influences the retinal firing pattern so that the visual pathways project the information experienced with high fidelity in a special sense. If perception is to be a direct (uncorrupted) specification of the world vis à vis information detected and directly experienced, then the medium of the central nervous system inside the body, like the medium of air outside the body, would have to be "transparent" and so pass the properties of the referent invariantly into experience. This transparent projection may be instantiated in many different energetic modes in between the reference object and the intentional object, but this is of no concern to the perceiver (unless the perceiver is a scientist); the perceiver merely sees *the world as it is* through his visual system, which has been carefully and relevantly tuned by evolution and learning to help him remain adapted to the environment. That is, it has been tuned to yield ecologically valid experiences.

Why should the retinal image be noticed in the course of perceiving the world? It is well known that our brains are insensate to being touched by probes; why then might not the retina, an extension of the brain, be insensate to the ephemeral touch of dancing photons and their rhythmic image? Is it required that the retina be treated as an image plane? Could this light-sensitive surface at the back of the eye be treated instead as a window? When one looks at the world through a window, there is a flat surface (the window glass) interposed between the observer and the world, but we do not say that the observer is looking at the window in order to look *out* the window. A window washer needs to look *at* the window, but ordinary observation *through* a window does not involve reading an image *off* the window. Consider looking outside a building through a window that is open versus one that is closed. Are these cases very different from one another? To the extent that window glass is transparent, we do not see it. We see the plane of the window only to the extent that it is not transparent, and what we see when we look at dirt on a window is something about the window itself, not the scene on the other side. A few people have argued that the retinal image, treated as a stage of analysis, is unnecessary or, even worse, a red herring which confuses rather than clarifies our understanding of perception (Gibson, 1966, 1979/1986; Haber, 1983). Let's consider two of these arguments, followed by a third, which is the primary focus of this chapter.

First argument: Retinal image is a red herring. The retinal image is an inverted, smaller-scale image of objects in the world, being projected upside down on the back of the eyeball. Yet, we do not experience the world itself as being inverted or small enough to fit into the eyeball. Furthermore, we have learned that by the wearing of inverting-prism goggles, before adaptation occurs, a person's reaching behavior is disturbed in predictable ways (Kohler, 1964; Dolezal, 1982). Thus, it seems that the image can be functionally inverted but not experienced as inverted under ordinary viewing conditions. The stage of interest therefore does not seem to be the first inversive projection (p_1) or even the second insensate cortical projection (p_2) but the *information resultant* of composing the first and second projection ($p_1 \times p_2$) into a third and final experiential projection ($p_3 = p_1 \times p_2$). Thus, the perceptual experience is not an event at the end of the train of three projections, not an effect that magically "pops out" at the end of a causal chain; rather, the experience longitudinally penetrates all three projections, with one foot in the environment and the other in the perceiver, and nothing but *transparent* physical and neurological media lying between. The resultant projection is over-mixed media (air and tissue) that are informationally transparent to the invariant properties of the environment—a distributed experience whose support is over the three projections. Knock out any of the distributed causal supports anywhere along the three projections, however, and the immediate consequence is some kind of blindness. The transparency would be destroyed.

Vision may fail because there is no light, or eyes are shut, or when the lenses are clouded by cataracts, or when the humors of the eye are too filled with debris (diabetic hemorrhages), or when fluid pressure compresses the ocular nerve (glaucoma), or when the retinae are detached, or when ocular tract has lesions or arterial occlusions, or when there is cortical damage, or when one is hit hard on the head, or when one is chronically inattentive or when temporally distracted. If causes of blindness can be distributed at different sites along the causal chain, *then so can causes of sightedness.* Why restrict experience arbitrarily to any specific location? Hence, the head is more likely *in the experience* than the experience is *in the head,* say, at the retinal image or some particular brain state. Because no one has solved the hard problem of where experience is located in the central nervous system (Chalmers, 1996), then we may locate it distributively over the field of concern. Our experiences join us with the objects experienced because our objects of intention directly specify our object of reference (Hintikka, 1975), so long as our history has appropriately attuned our perceptual systems to the relevant information that information is detected by us (Chan & Shaw, 1996).

This ecological view of direct perception differs somewhat from that of the Gestaltists' principle of psychoneural isomorphism, for it incorporates their brain field into an ecological field; their brain field is integrated into a more comprehensive *psycho-neuro-physical* field that interfaces a functionally defined environment (econiche) with its functionally defined organism (a perceiver-actor). (Note: Experiences that arise from dreaming, imagining, or hallucinating are allowed but simply do not have the referential transparency that direct perceiving and knowing do).

Second argument: Ganzfeld is experienced as three-dimensional. Assume that the total field of view is entirely filled by an illuminated, white, featureless surface. Such a homogeneous field of light with no visible boundaries is called a *Ganzfeld.* There are no focusable contrasts for binocular hunting to stereoscopically lock onto or to which the eyes' lenses can accommodate (i.e., change their shape). Nevertheless, shouldn't the experience be one of two-dimensionality? If the retina is an image plane, would we not still see the retina in the absence of a projection? A blank canvas is still an object to be seen. If the retina is an image plane to be seen, then shouldn't it show up as a flat surface if no perspective projections are given? But will a person really see a two-dimensional image of lightness, as if a white surface has been painted on the retina? Or will one see instead a white, featureless surface located at some determinate distance from the point of observation? If so, how far? Such Ganzfeld experiments have been done (Metzger, 1930; Gibson & Waddell, 1952; Cohen, 1957) but with an outcome that could not be predicted from retinal image theory.

Instead of experiencing a white surface at some indeterminate distance in the so-called frontal plane, people report experiencing a three-dimensional translucent volume of indeterminate depth. Statements are made like: "I am looking into a penetrable white fog that completely surrounds me!" Thus is our most primitive visual experience, as the Gestalt theorists argued, an autochthonous experience of three-dimensional, unbounded openness (Ganzfeld) that arises independently from nowhere. Ecological psychology would explain this experience otherwise than being a mysterious autochthonous "force." Like the rest of science, we would look for a sufficient reason for the phenomenon.

The basic premise of an ecological theory of perception is that *we see what we see because the information from the environmental situation is what it is.* In other words, we do not see what is simply in the light to the eye, as the physicist might construe it; rather we see what is functionally specified by the light to a highly evolved visual system—one that has been adaptively designed to fit its environment by evolution and further attuned by experience.

The information contained in a Ganzfeld specifies no surface because there are no focusable features on the surface creating the Ganzfeld. The eyes cannot accommodate to any given distance because no specific distance information is given. A situation in which there is light but no surface information is an insubstantial medium (like fog) quite capable of indeterminate penetration. This explains the first part of the argument needed: namely, how a 3-D object can be represented on a 2-D surface.

Third argument: Ordinary projective geometry does not preserve orientation information. Because traditional perceptual theory depends on the retinal image projection but such theory fails to explain the experience of tridimensionality, we must search for a different theory. No matter what projective theory is needed for describing the information input for visual experience, it must be based on a different kind of projection than that which describes the retinal image. Even if this were not sufficient to cast doubt on retinal image theory, there is a third even more telling argument having to do with the fact that ordinary projective geometry fails to preserve orientation information. Such a theory fails because the retinal image does not distinguish two distinct kinds of projections that need distinguishing if mischief is to be avoided. Consequently, this brings us to our third argument, which in many ways is the most important one.

Orientability, among other things, allows us to recognize counter-clockwise rotations from clockwise ones, left from right, top from bottom, and inside from outside. The argument we wish to present is based on whether the topological property of orientability is present or absent in the projective space of interest—whether this be the retinal image treated as a two-dimensional projective space or the dynamical retinal image treated as a three-dimensional projective space. We discuss next this important property of orientability, which is missing from all ordinary projective geometries regardless of their dimensionality. After that, we shall turn to framing a mathematical basis for an ecologically valid projective geometry.

Orientability and Sidedness

Take a few minutes to scrutinize carefully the following figures, and then we shall pose a few telling questions.

Look carefully at figure 2.1A. It depicts an ordinary two-sided carton with two cells. There is clearly an outside and an inside. The left-most arrow is inside the left cell; the middle arrow is outside the left cell and inside the right cell; and the right-most arrow

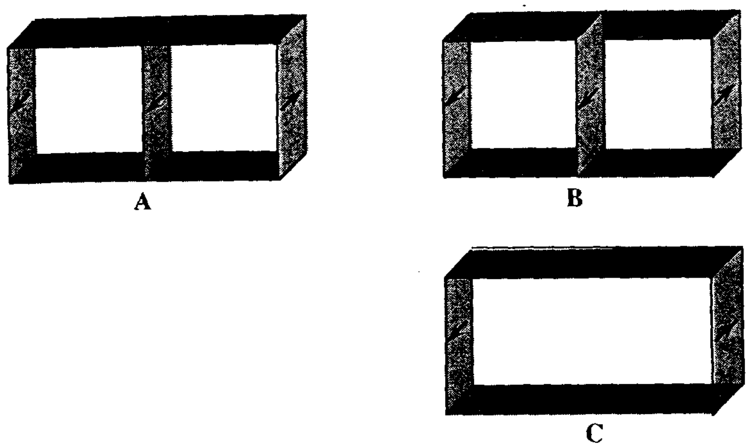


Fig. 2.1. Orientability and sidedness. *A* is a two-sided, surface with orientability while both *B* and *C* are one-sided surfaces without orientability. Can you see why?

is outside the right cell. Compare *A* with *B: B* is also a carton with two cells; but how many sides does it have? Notice that the three arrows in *B* have exactly the same placement as those in *A* relative to the flat surface of the page but not relative to the surfaces of the depicted cartons. The right-most arrows in *A* and *B* point in opposite directions on different sides of their respective cells: In *A*, the right-most arrow is outside and points backward while the corresponding arrow in *B* is inside and points forward. If this does not seem remarkable, then compare the middle arrows in each. In *A*, this arrow is inside the right cell and points outward but has no clear orientation in *B:* It seems to be outside of both the left and right cells and to point outward and inward, respectively, at the same time! *C* simplifies the picture so that it is easier to see that the surfaces of the *B* carton are based on the one-sided Möbius band. To clarify the relationship of orientability to sidedness, consider figure 2.2.

To understand this breakdown of the orientability property, we need to understand *sidedness*—an important topological property that projective geometries usually do not preserve. Later we shall see that *two-sidedness* is a necessary property of projective geometries to have distinguished, as pointed out earlier, front from back, inside from outside, left from right, top from bottom, and clockwise from counter-clockwise. To anticipate further: *Two-sidedness* is the minimal property any geometry must have, whether projective or not, if it is to have a way to handle occlusion information—a key informational invariant of a theory of the three-dimensional layout of surfaces in the environment and hence one of the most ubiquitous sources of information for perceiving three-dimensionality.

Try this demonstration: As shown in figure 2.2A, glue the corresponding ends of a paper strip together to make a cylindrical band (i.e., *a*⇔*a, b*⇔*b*). Notice that an ant crawling on the inside circumference of the band would stay on the inside or if crawling on the outside surface would remain on the outside surface. It would have to crawl over an edge to change sides. For this reason, the cylindrical band is called a *two-sided, bounded* surface.

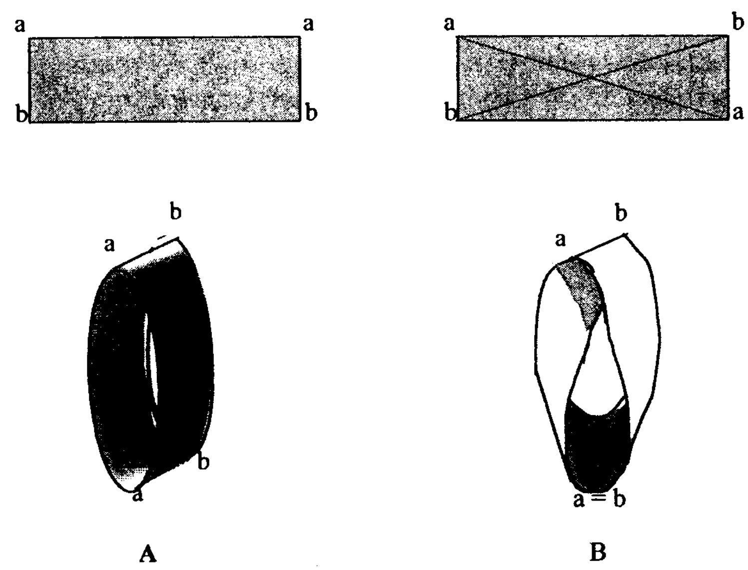


Fig. 2.2. The two-sided cylindrical band versus the one-sided Möbius band.

Now, as shown in figure 2.2B, take another paper strip and glue the noncorresponding ends together by giving the paper strip a half-twist (i.e., *a*⇔*b*, *b*⇔*a*). Notice that an ant crawling on this surface, even without crossing over an edge, will nevertheless cover what appears at one moment to be the inside but at another moment the outside. This is why the Möbius band is called a ***one-sided unbounded*** surface. If we draw a closed path on the circumference of the cylindrical band, an arrow transported around this path will retain its orientation (see fig. 2.3), but an arrow transported around the corresponding closed curve on the Möbius band will not retain its orientation. The property of orientability is a consequence of the surface being two-sided, while the loss of this property is a consequence of a surface being one-sided.

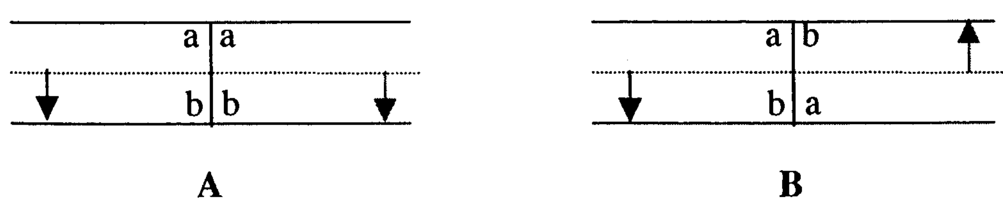


Fig. 2.3. Preservation or loss of orientability. *A*, parallel transport around a two-sided surface preserves orientability while *B*, parallel transport around a one-sided surface does not.

Orientable and Nonorientable Objects

To make clear the way in which projective transformations typically lose orientability information, consider the simple example of rotating a triangle in the plane.

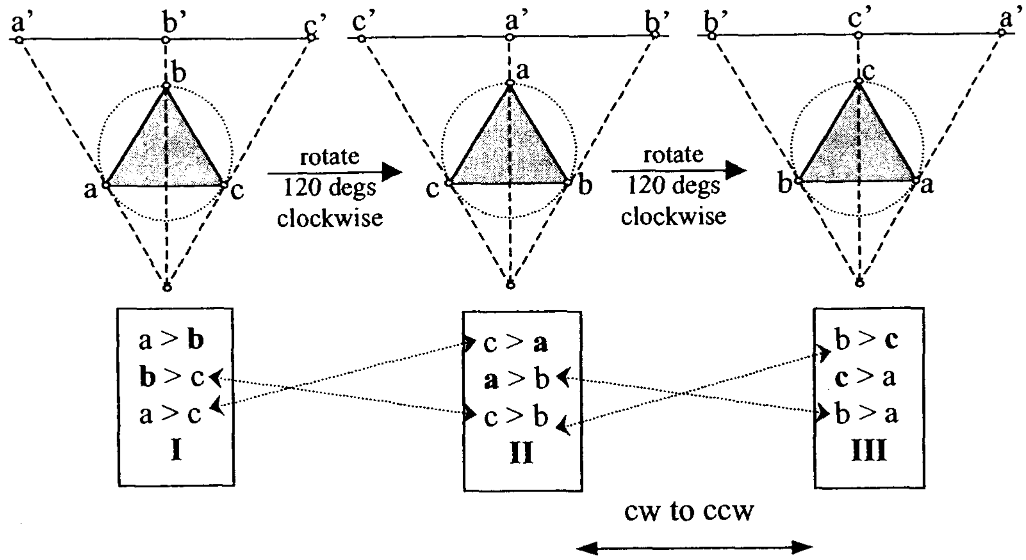


Fig. 2.4. Removing ambiguity from a projected rotation event. Here > specifies order of sequential occurrence (i.e., *to the left of* on the projective line) and bold letters denote the front range of the projective mapping.

The sequence *I, II, III* in figure 2.4 denotes a clockwise rotation, while the sequence *I, III, II* denotes a counter-clockwise rotation. Rotation direction reverses if the back range and front range are interchanged. Because occlusion information for orientability is suppressed, the projected dynamical shadow of any rotating object appears to reverse direction spontaneously. This is explained by the fact that some of the successive relationships on the projective line reverse order (indicated by the arrows). If occlusion or any other information is available to "mark" the front range (indicated by the bold letters), then there is no misidentification of what is in the front range and in the back ranges. Hence, orientability information is preserved under projected rotation. These arguments apply to any objects regardless of shape. Consider another case of loss of orientability: the so-called Necker cube.

Real Cubes, Necker Cubes, and Projection Topology

Assume that we have a cube in 3-D space (fig. 2.5, column I) that is projected onto a 2-D surface (column II-top) thus collapsing the cube's six faces into a complex with a maximum of seven co-planar, polygons (depending on the orientation of the cube in I, the number can be smaller); this 2-D polygonal complex is then topologically transformed into the unit circle (column III-top), preserving the seven regions but not their shapes. Alternatively, the ordinary cube can be projected in two other ways: either onto a one-sided (nonoriented) 2-D representation of a 3-D cube, called the *Necker cube* (II-middle), or onto a two-sided (oriented) 2-D representation of a 3-D cube (II-bottom). The number of each face is placed in the center of that face. For example, the numeral 5 refers only to the square that its face is centered in, and 6 refers to the square that its face is centered in. The one-sided figure (Necker cube representation) is then also topologically transformed into a unit circle in such a way as to preserve the ambiguous orientation of its faces, while the two-sided (real cube representation) is so transformed as to preserve the unambiguous orientation of its faces. The ambiguity of column II-middle is shown by arrows leading to both the middle and bottom topological maps in column III. Column II-bottom, on the other hand, is drawn with an arrow only to the column III-bottom topological map. The middle and bottom maps in column III indicate that two faces are projected to each of the seven regions. In column III-middle, the bold numbers—1, 4, and 6—depict the faces that are seen in front. The numbers not in bold—2, 3, and 5—indicate the faces seen *through* the faces that are in front. The bottom map in column III represents the alternative orientation of the cube with fronts and backs interchanged.

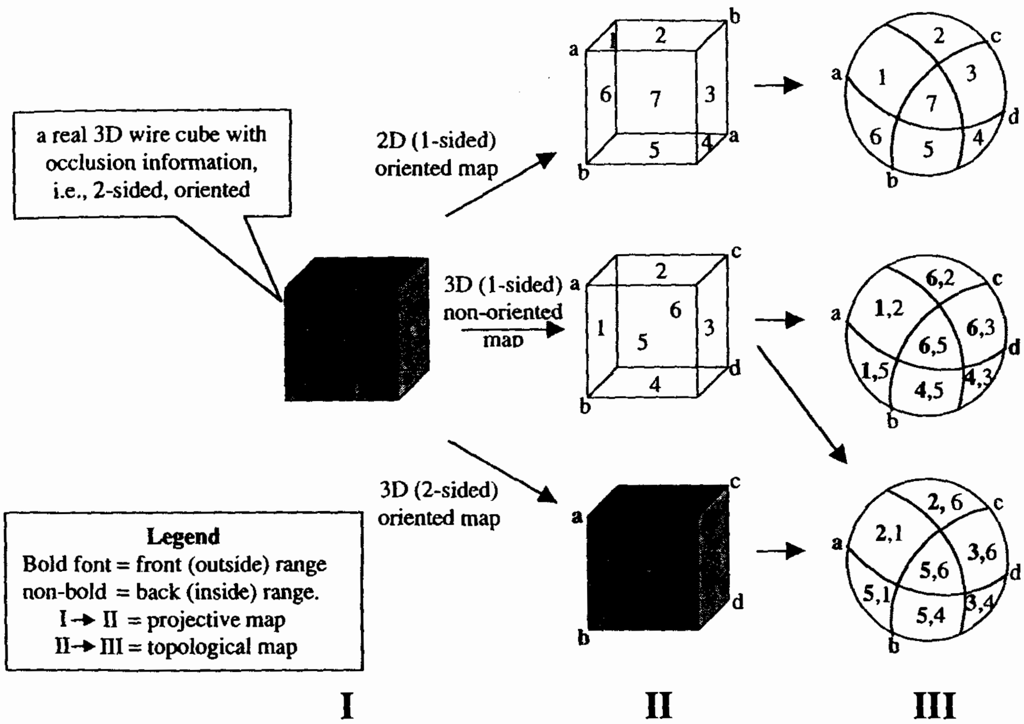


Fig. 2.5. Contrasting perspective theories: nonoriented and oriented projective geometries.

Hence we have a reason for the three different experiences. Namely, they correspond to the three projections:

- I => II-top => III-top: no occlusion information
- I => II-middle => III-middle *and* III-bottom: occlusion information is ambiguously specified
- I => II-bottom => III-bottom: occlusion information is unambiguously specified

If information is defined as specification of reference-object properties under a projective mapping and if we have three projective mappings that convey information from the environment to the visual system in three different ways, then we should have three different experiences (intentional objects); and of course we do!

Just as in the argument that the Ganzfeld is to be explained by the projective mapping being a faithful specification of three-dimensionality, we now argue analogously for the Necker cube. It is an ambiguous figure, not because some creative cognitive magic takes place but because its projective mapping, like the Mobius band, is a forgetful specification, leaving behind occlusion information. This contrasts sharply with the faithful projection of the occlusion information in the case of the unambiguous representation of the cube. Because the latter projective mapping is the most faithful in specifying invariantly the properties of the reference object, then it provides the most ecologically valid experience although it should be understood that all three experiences are accounted for by direct perception. Namely, *you see what you see because the information is what it is.* The specification is of just those properties experienced. Nothing need be cognitively constructed, remembered, or inferred—that is, no autochthonous "forces" need be postulated to account for the alternative experiences. More would need to be said to explain the selection of one of these alternatives at a given time, but this would be a *selection* among justified alternatives, not a creation or a construction.

Next, we must discover why the orientation-specific information is lost in our experience of the Necker cube.

Oriented Projective Geometry

Ordinary projective space, such as the Mobius band and the Necker cube, is one-sided as shown in figure 2.6. The spherical model of this geometry represents the fact that the projections of a point on the back of the sphere and of a point on its front both have the same image in the Euclidean (projective) plane, represented here as an infinite disk. (Note: The circumference of the disk actually lies at infinity where the angle of projection reaches 180 degrees, i.e., lying in the *xy*-plane, and completely covers the plane with images of points from the sphere.) All of the projected points, regardless of the hemisphere to which they belong, cover the projective plane in the usual way without any designation of where they originated. The loss of orientability is due to this failure of the projective mapping to preserve the distinction between the front and back range, collapsing both into positive values of the dimension of depth *w*. This loss of orientability is represented by the fact that relationships (e.g., the arrows) invert when the projective angle passes through the points at infinity.

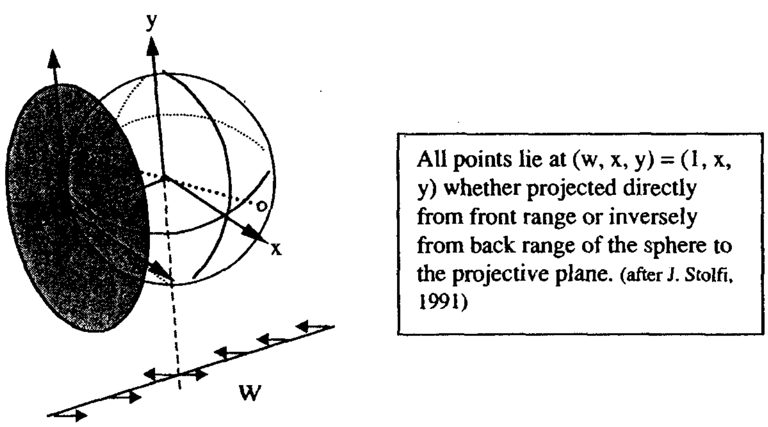


Fig. 2.6. The spherical model for ordinary projective geometry.

To keep the front and back ranges distinguished, traditional computational geometries use the line at infinity as a reference. This means we would have to exclude certain "degenerate" cases, such as line segments with one end on that reference line. But this move is not a real solution to the orientability problem in ordinary projective geometry because it is tantamount to a return to Euclidean geometry and hence to a geometry without a natural theory of perspective. Graphics programmers use many tricks to distinguish the front and back ranges: among them normalized signing, ray tracing, and a negative weight-clipping rule. These are ad hoc provisos rather than a systematic change in the basis of the geometry itself. For this reason, in traditional perspective theories, occlusion information is not principled. There is a better way of keeping the orientability information intact. *Oriented* projective geometry introduces a principled way to distinguish the front and back ranges. (We follow, in part, Jorge Stolfi's 1991 book in this presentation, which we highly recommend for those who are mathematically inclined.)

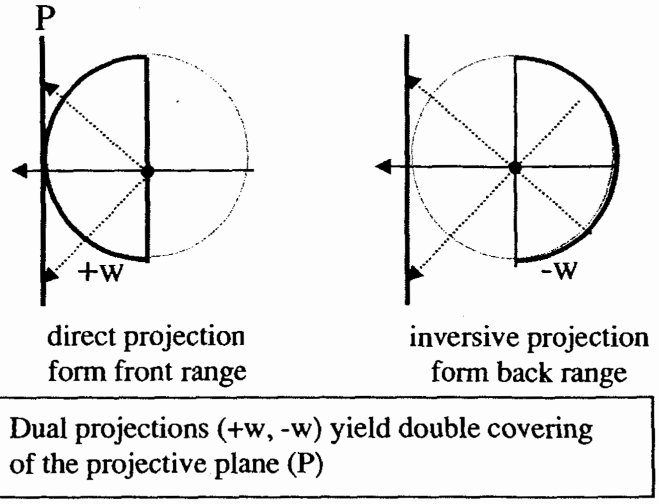
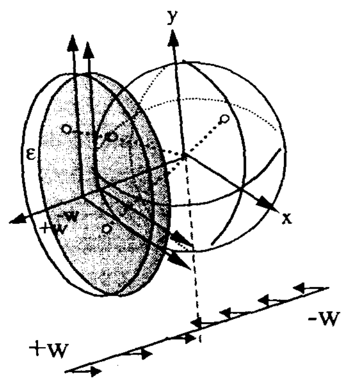


Fig. 2.7. Oriented projections with duomorphic projections.

In figure 2.7, we assign a dual range, +w and –w, to represent the front and back ranges of the spherical model, respectively (the front range and the back range are shown with the opposite dimension suppressed).

In figure 2.8, the projective plane is no longer without thickness but is a manifold (surface) of *infinitesimal* (*e*) thickness. Hence, every point on the "thick" plane is a double point, with each member of the pair being marked by either +w or –w, depending on whether it occupies the front or back range. Also, the line at infinity is no longer needed as a reference line.

Let's take note of a few of the technical concepts needed to describe the new projective geometry. Here we get a duomorphic or double covering of the projective manifold, that is, a covering by double points. A *double point* is not just two coincidental points but is also a neighborhood defined by a duomorphism. A *duomorphism* comprises two distinguishable functions, such as, a pair of dual projective transformations, with distinct ranges lying within the same topos—a concept from category theory referring



Points lie at (+w, x ,y) = (1+ ε, x, y) if they project directly from front range and (-w, x ,y) = (1 - ε, x, y) if the project inversely from back range of the sphere to the projective plane.

Fig. 2.8. The spherical model for oriented projective geometry.

to an infinitesimal region containing both points and a rule for distinguishing them in terms of their origin rather than their destinations. Here what looks initially like a *two-to-one* mapping from domain to range is actually a pair of dual mappings, or duomorphisms. Recall that an *isomorphism* is a mapping that is *one-to-one* and *onto,* while a *duomorphism* is a kind of isomorphism that is only *reflexive* and *symmetrical* but *not transitive* as are other isomorphisms, such as equivalence and identity.

Finally, we spoke of a "thick" two-sided surface, or manifold. But how thick is infinitesimal thickness? An *infinitesimal* number is a number that is greater than zero but smaller than any real number and belongs to the *hyperreal* domain consisting of both the real points and the infinitesimals nested among them. (See J. L. Bell, 1998, for a lucid introduction.)

In the next section, we apply this model to discuss some key issues of perceptual theory.

A Plethora of Double-points

The "depth" seen at an occluding edge of a surface (fig. 2.11) involves a *scission effect,* just as does a surface seen though a semi-transparent surface (fig. 2.9A). A scission effect is where a single projection carries information for more than one surface. Clearly, all the points along the line defined by an occluding edge qualify, as do all the points seen through a semi-transparent surface. In each case, a point *c* on the projection surface involves at least a pair of other points, *a* and *b:* One point *a* is seen to lie either in front of or behind another point *b* (see fig. 2.10). Although the separation *and* order of the surfaces in depth is nearly always clearly specified in occlusion, only the separation of the surfaces in depth, not the order, is clearly specified in transparency. In transparency their order usually appears indeterminate. The indeterminate order of separation can be understood as the failure to break parity.

Convexity: The Missing Ingredient?

Two unanswered questions deserve attention: If projective mapping can convey surface separation information but not determinate order, then by what information is determinate order specified? Of course, the answer is whatever information conveys two-sidedness ipso facto specifies ordered separation in depth. But this does not say precisely what such information is. To appreciate the nature of scission effects, it will be instructive to pay careful attention to another property that often accompanies orientability and is only defined if it is, namely, convexity. After defining this new concept, we will try to show how it may be the missing piece of the puzzle of depth perception and that it introduces order into the scission effects, regardless of how they are achieved.

Please study the displays in figure 2.9 for a moment.

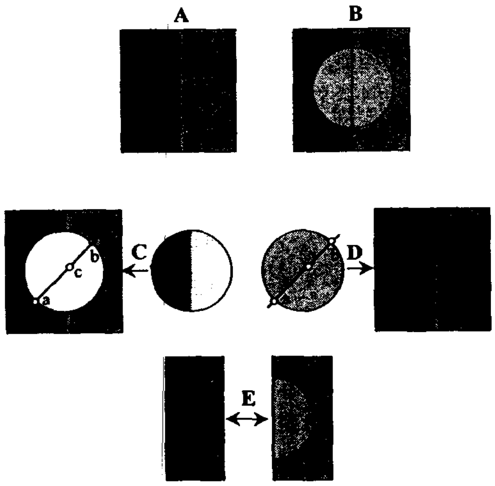
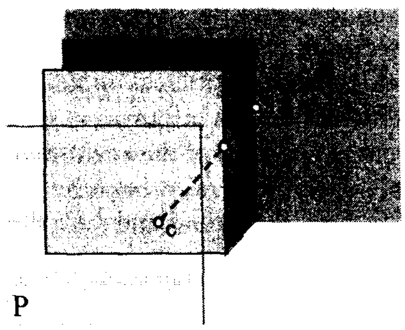


Fig. 2.9. Transparency and convex sets.



Assume the line of projection a, b, c takes point a on the back surface along with point b on the front surface onto c on the projective plane P. An ordinary projection would collapse the two points a and b into a point of coincidence at c, while an oriented projection would specify a double-point [a, b] at c.

Fig. 2.10. Occlusion information specifies double-points.

One experience typically specified by figure 2.9A is of a semi-transparent disk covering a surface with two contrastive regions. A certain law of psychophysics (Talbot's law) has been shown to account for a broad class of transparent depth phenomena so long as certain initial conditions are satisfied (Metelli, 1974; Anderson, 1997). One condition is that light contrast values must be present; and the second is that the light contrast values be in a certain order. In figure 2.9A, both conditions are met, and a transparent depth is experienced, while in B, they are not, and no such experience arises. This constitutes a general law of ecological physics in that it systematically links information conditions with a specific experience.

Note the chord across the open disk at figure 2.9C and also across the closed disk at figure 2.9D. Think of the circular area in *C* as indicating empty space. This makes *C* a nonconvex figure and *D* a convex one for the following reason. A *convex set* is one such that if the end-points, *a* and *b* of the chords lie in the set, then so must any point, *c*, lying between the endpoints. Clearly, then, the open disk of *C* is nonconvex while the closed disk of *D* is convex. We can use the convex set property to clarify the figural condition so as to distinguish the two spurious cases from the ecologically valid one, as illustrated.

An ecologically valid display for transparent depth must satisfy the following geometric and optical information conditions:

First, the light contrast conditions must have the values dictated by Talbot's law;

Second, the light contrast values must be arranged properly, such that

- they have values in the back range that do not belong to the front range, and
- they specify a convex set, which has values in the front range that do not belong to the back range.

If these conditions are met, then there is information for a scission effect that could only originate from a source with an ordered separation of surfaces. Hence, the information would have the fidelity required to qualify as a direct specification of an ecologically valid experience. Circularity is avoided in defining these transparent depth conditions in that the scission effect needed for surface separation and order is assimilated to the new oriented projective geometry as a consequence of two-sidedness—something that would not be possible in the old projective geometry.

Dynamical Occlusion as Displaced Accretion-Deletion Fronts

If the visual system is to distinguish between the front and the back range of an environmental projection, then there must be information for the *order of separation of the surfaces* specified through the optical projection. Ecological optics, as opposed to traditional optical physics, has accepted the task of discovering such information sources that account for our most important and most salient experiences of the environment—an environment shared by all life forms and within which they must organize and direct their behaviors in adaptive fashions. Some of the most ubiquitous and most important information is that for specifying the occlusion of one surface by another surface as seen by a perceiver.

Occlusion information through interpositioning, however, is not the only means for specifying the order of surface separation. Nonoccluding surfaces may be ordered in depth if they occupy different positions of an optical texture gradient or if they move at different rates toward or away from the perceiver. Because we cannot survey all such cases here, let's consider dynamical occlusion as our last example.

It has been well established that ordered depth effects are specified by accretion and deletion of texture, even in the absence of occluding real surfaces (Kaplan, 1969). Imagine that on a computer screen or in a movie, one sees a randomly textured pattern that completely covers the screen (fig. 2.11). Then suddenly a small rectangular section of the random texture is seen to emerge (at t_1) from the background camouflage, moves in a straight line for one-third of the screen width (t_2-t_3), and finally stops, merging back into the camouflage of the background (t_4). This merging into the background shows that the edges over the accretion and deletion change and do not exist in the static image.

In the real-world case in which one surface dynamically occludes another, the leading and the trailing edges of the surface in front will define moving fronts of accretion and deletion. The moving accretion and deletion front is the information that optically

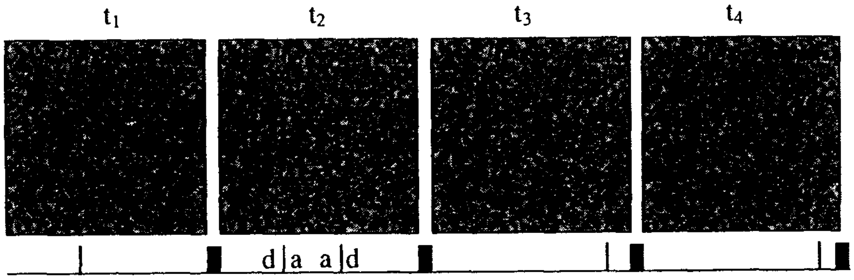


Fig. 2.11. Accretion and deletion of texture specifies object motion.

specifies an edge with depth. However, for this to be unambiguous, the texture between the accretion and deletion edges must be preserved.

A case of more-pure accretion and deletion without preservation of internal texture shows us that accretion-deletion alone is not sufficient to unambiguously specify ordering. Suppose that one creates a case in which background texture is deleted while foreground texture is accreted—and that's all. This would define the ***leading edge*** of a possible occluding surface, albeit an odd one. Deleting foreground texture and accreting background texture define its ***trailing edge*** (see fig. 2.12). For the case depicted, the texture trapped between these nonadjacent fronts defines the occluding surface as a convex set with values in the front range.

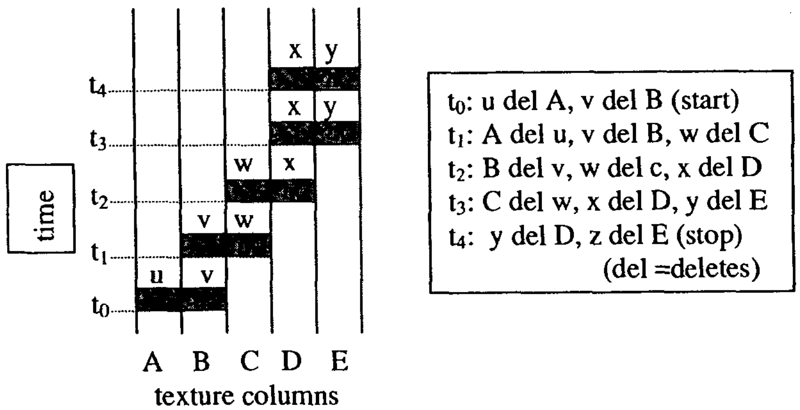


Fig. 2.12. Illustrating accretion and deletion fronts.

In figure 2.12, regions of background texture *(A, B, C, D, E)* get replaced by regions of foreground texture *(u, v, w, x, y)*. The sub-regions *A, . . . , E* and the sub-regions *u, . . . , y* are static regions on the surface of the screen. Texture replacement is defined by accretion of new texture in the place of existing texture, which is correspondingly deleted. Where the accreting and deleting take place, changes occur—optical disturbances but no texture is actually transported over locations.

In this case, however, the display may reverse so that what was the occluding convex set in the front range becomes an occluded surface in the back range, and the occluding surface is now not a convex set. To make the possibilities clear, we depict three cases that might be experienced from the same display (fig. 2.13A, B, C; these cases are meant to be read as fig. 2.12, but in fig. 2.13, the small rectangular regions are offset for illustrative purposes.) Case *A* is the accretion-deletion fronts just as before; but these may be experienced as either case *B* or case *C*. In case *B*, the occluding surface is seen as convex while in case *C*, the occluding surface is seen as nonconvex, that is, as a small moving aperture through which is seen the background texture.

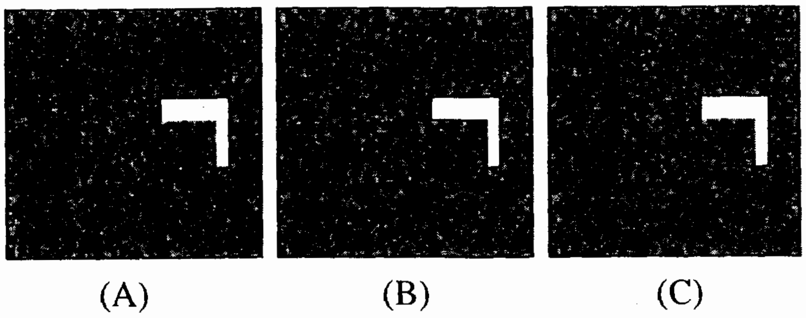


Fig. 2.13. Ambiguous occlusion.

The dynamical display is therefore reversible, like the Necker cube. In figure 2.13C, Rule 1, if the leading edge accretes behind itself, and the trailing edge deletes in front of itself, then the occluding set is convex; while in figure 2.13B, Rule 2, if the leading edge accretes in front of itself, and the trailing edge deletes behind itself, then the occluding set is nonconvex. These relative directions of accretion and deletion, therefore, partition the total display into those regions belonging to the front range and are thus occluding; and these relatives directions also partition the total display into those regions belonging to the back range and are thus occluded. Rules 1 and 2 define a duomorphism and are dual theorems.

Perceptual Fidelity Is a Function of Ecological Validity

In this final section, we return to some of the earlier issues regarding ecological validity. With the differences between ordinary and oriented projective geometry firmly in mind, we now have a principled basis for our argument.

From the discussion of transparency, we saw that information may be from a source (e.g., display) that need not have the property experienced. Notice that *A* in figure 2.9 is a source of information that specifies a transparent depth experience. In constructing the display, however, one may have followed any of the three procedures, although only *D* represents an actual case of a transparent surface being placed over the contrastive background surface. Hence, the source of the information might not possess the property that the information from that source specifies.

There are, of course, other procedures one might have followed, such as painting a picture (intentional object) in the proper way to specify a transparent reference object (source) or writing a graphics program to create such a display. The question that is nagging is the following: If a display (optic array structure) can be contrived to create information experienced as it having some property X but fails to possess that property, then might we not be fooled about our natural environment? Might perception have low fidelity?

The short rebuttal to this question is that our perceptions have fidelity to the degree that the actions we take in accordance with experiences succeed in achieving the intended goals. There can be no better yardstick for perceptual fidelity than the degree to which information is lawful in helping organisms as perceivers achieve positive outcomes as actors. The ecological fidelity for information detected by successful actors must be high, whether they be simple or complex organisms. From a pragmatic point of view, then, truth in perceiving the world is determined by the value gained or lost. Ecologically valid experiences are simply those that are most lawful in preserving the "right" values. Truth is what truth does! Hence, the issue of information fidelity is neither more nor no less than the issue of ecological validity.

Ecological Realism as a Critical Perspectival Realism

Occlusion is a perspectival relationship that only makes sense when the point-of-view is such as to place one surface between the perceiver and another surface. For this reason, what distinguishes the front range from the back range is the place of the perceiver, as the projective surface, within the layout of surfaces in the environment.

This makes the problem of occlusion perspective-dependent. Its laws, as formulated in Rules 1 and 2, belong to ecological physics rather than ordinary physics, where laws are intended to be universal and perspective-free rather than general (socially invariant) and perspective-dependent.

Perhaps, the notion of ecological realism that perception is direct but does not entail naïve realism is the most difficult principle of ecological psychology to grasp; namely, for example, that the visual world consists of a 3-D optical array of texture where all forms of change are merely optical disturbances in various regions of the array with its nested n-tuple ranges. We have seen cases where critical realism has come into conflict with common-sense (naïve) realism often in science: most dramatically, perhaps, when Copernicus refused to accept the naïve realism that claimed that the sun orbited the Earth because we "see" the sun rise in the east and set in the west; or when we reject the flat-Earth hypothesis, even though under our limited viewing conditions, the Earth does indeed look flat.

These naïve claims are not based on ecologically valid experiences, because we have drawn inferences that go beyond the limited information available. We may only conclude legitimately that the optic-array information samples that we have under our restricted circumstances do not themselves rule out either the geocentric theory or the flat-Earth hypothesis.

The information available surely does not affirm them but leaves room for two alternative hypotheses—the heliocentric theory and the round-world hypothesis. This wiggle room for critical realism to assert itself is justified because it allows a change in possible viewing circumstances, say, by taking a ride into outer space. From this broader, unrestricted perspective, we see directly how the local flatness of the Earth gives way to global curvature, and how the Earth moves among its sister planets to circle the sun.

As scientists, we must be conservative with our guesswork. For we do not have the ontological luxury of assuming the character of surfaces, objects, and events are as commonsense experience tells us. Rather, we must discover the information that specifies their particular perceptual character by broadening our perspectives. In this way, we accept no cheap ontological conclusions about environmental sources (reference objects) but must work to justify all such claims through sound experimental epistemology to identify the information and restrictions on our circumstances responsible for our ordinary experiences (intentional objects).

We then must work toward a systematic relaxation of those restrictions to reveal the larger truth of ecologically valid experiences. Art helps us do this, especially the special effects created by visionary artists. Artists create such circumstances (displays) that inform us of ways to transcend our world of ordinary experiences and in this sense provide intuitive bases for ecological physicists to study nature more directly.

Both the motion and the rectangle are specified by the relative accretion and deletion functions. Thus, there is no denying that what we see in movies, videos, or computer graphics is quite different from what actually happens. To understand the optical information for motion, we must distinguish between what is specified by the optical information, the intentional object of our experience, and what happens at the display that is the source of the information, the reference object of our experience. But one should not think of information as mere appearance and its source as the true reality for they are both equally real.

Of course, the information may be presented in such a way that it may specify something other than its source. This follows naturally from the condition that more than one source may display the same information and hence give the same perceptual experience. The basic postulate of ecological psychology, you will recall, is that the source is always specified as well. Just because we do not recognize the source for what it is simply means we may have to sample the information more extensively over many more perspectives before getting it right. "Getting it right," so to speak, is to elaborate our perspective sampling of the optic array until we have a valid ecological experience—an experience that stands up to all lawful scrutiny.

Obviously, both scientists and artists learn to contrive situations that look one way under a given set of perspectives and another way under another set of perspectives. Presentation constraints are very important. Special effects technicians, like magicians, know this only too well. The other side of this issue is that special effects that dissimulate their true sources follow from lawful practices that can be understood and reliably reproduced. By a careful study of alternative ways to present the same information, we eventually discover the ecological laws upon which to base our theories of perceptual experience.

Note

1. Many writers have used the phrase *ecological validity* in an intuitive way that is not especially technical. They are aware, however, that Egon Brunswik (1956) was well known for the concept of ecological validity and frequently cite him. In most cases, it appears that Brunswik is credited as a scholarly courtesy. Some writers apparently did not extend Brunswik the added courtesy of reading him. For examples and a scolding, see Hammond, 1998. For extensive material on Brunswik, see Hammond, 1966. Brunswik's use of ecological validity was a very specific one, the correlation between a cue (say, retinal size in vision) and an environmental property (real size). Ecological validities could have any value on the 0 to 1 range of correlations. Our usage begins closer to the intuitive usage and then is developed more technically within our version of the ecological program (Shaw, Turvey, & Mace, 1982, p. 209). We respectfully notify our readers that we are *not* using Brunswik's concept of ecological validity.

References

Anderson, B. (1997). A theory of illusory lightness and transparency in monocular and binocular images: The role of contour junctions. *Perception, 26,* 419–54.

Bell, J. L. (1998). *A primer of infinitesimal analysis.* Cambridge: Cambridge University Press.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory.* New York: Oxford University Press.

Chan, T. C., & Shaw, R. E. (1996). What is ecological psychology? *Psychologia: An International Journal of Psychology in the Orient, 39,* 1–16.

Cohen, W. (1957). Spatial and textural characteristics of the Ganzfeld. *American Journal of Psychology, 70,* 403–10.

Dolezal, H. (1982). *Living in a world transformed.* New York: Academic Press.

Gibson, J. J. (1986). *The ecological approach to visual perception.* Hillsdale, NJ: Lawrence Erlbaum. (Original work published 1979).

———. (1966). *The senses considered as perceptual systems.* Boston: Houghton Mifflin.

Gibson, J. J., & Waddell, D. (1952). Homogeneous retinal stimulation and visual perception. *American Journal of Psychology, 65,* 263–70.

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

Haber, R. N. (1983). The impending demise of the icon: A critique of the concept of iconic storage in visual information processing. *Behavioral and Brain Sciences, 6,* 1–54.

Hammond, K. (1998). Ecological validity: Then and now. Available: http://www.brunswik.org/notes/essay2.html.

———. (Ed.) (1966). *The psychology of Egon Brunswik.* New York: Holt, Rinehart, & Winston.

Hintikka, J. (1975). *The intentions of intentionality and other new models for modalities.* Dordrecht: D. Reidel.

Kaplan, G. (1969). Kinetic disruption of optical texture. *Perception & Psychophysics, 6,* 193–98.

Koffka, K. (1935). *The principles of Gestalt psychology.* New York: Harcourt, Brace.

Kohler, I. (1964). The formation and transformation of the perceptual world. *Psychological Issues, 3* (monograph 12).

Metelli, F. (1974). The perception of transparency. *Scientific American, 230,* 90–98.

Metzger, W. (1930). Optische Untersuchungen im Ganzfeld II. *Psychologische Forschung, 13,* 6–29.

Shaw, R. E., Turvey, M. T., & Mace, W. (1982). Ecological psychology: The consequence of a commitment to realism. In W. Weimer & D. Palermo (Eds.), *Cognition and the symbolic processes* (Vol. 2, pp. 159–226). Hillsdale, NJ: Lawrence Erlbaum.

Stolfi, J. (1991). *Oriented projective geometry.* Boston, MA: Academic Press.