

Part I
Linking Psychophysics
and Qualities

COPYRIGHTED MATERIAL

I

Inferential and Ecological Theories of Visual Perception

Joseph S. Lappin

Visual Phenomena

A basic principle of phenomenology: phenomena vary
with the observer's perspective

Visual phenomena also vary with one's theoretical perspective. Viewer-dependence plays different roles in inferential and ecological approaches to perceptual theory. Inferences about the environmental causes of sensory data are complicated by viewer-dependent variations; but viewer-dependence has a central and explanatory role in the ecological approach.

Objective and subjective

Intuitively, the world we experience usually seems an objective reality—shaped by what is rather than by the instruments of our senses and technology. We know, of course, that what we see depends on our vantage point, on our eyes being open, whether the TV is turned on, and so forth. But when eyes and doors are opened and when a video receiver is switched on, then we generally regard a revealed scene as having been there all along, independent of our eyes and technology. A belief in the objectivity of observation has seemed essential to the professional practice of many scientists and engineers, as well as to the tacit knowledge of most of us. To be sure, the world we experience is not a fiction of our imagination.

Nevertheless, for painters, poets, musicians, and photographers, perceptual experience is neither deterministic nor the product of an objective world. Artists design objects for purposes of “orchestrating experience”¹—to give meaning and emotional significance to both the objects and processes of observation. Making art is obviously

creative, but observing art is also creative. Observing and making art both involve active choices of attention to form, context, and meaning. And observing art is strongly influenced by one's vantage point and knowledge. In the everyday world as well as in museums, what we observe is selected from what might be seen. Attention is guided by context, learning, memory, meaning, and emotional significance.

Our choices of attention and action are also constrained by what our perceived surroundings afford—by walls and hallways, forests and trails, and traffic on the roads we travel. Our lives depend on the compatibility of our choices with changing environmental conditions. Our senses may sample limited patterns in our surroundings, but these perceived patterns must not conflict too often with the available constraints and opportunities.

Thus, visual phenomena are multifaceted. Different perspectives afford different descriptions and different explanations. Scientific experimenters and the observers who serve as subjects have importantly different perspectives. One's subjective, personal experience looking, as it were, from the inside out is obviously very different from that of a scientist studying vision by looking from the outside at another person's behavior. Scientific observations about other persons' visual experiences are obviously limited. If visual experience is not objectively observable by another person, does it belong to science?

In fact, logically rigorous psychophysical methods have been developed to characterize other persons' perceptual discriminations (e.g., Garner, Hake, & Eriksen, 1956; Green & Swets, 1966). Effective psychophysical methods usually concern subjects' discriminations among physical objects rather than the subjective experience per se. Does subjective experience belong at all within the domain of science?

The method of introspection, developed in the late 19th century, was designed to observe the characteristics of other persons' subjective experience. Subjects in an introspective experiment provided verbal descriptions of their phenomenological experiences, thereby offering to the experimental scientist indirect evidence about that experience. In the words of E. B. Titchener:

The first object of the psychologist . . . is to ascertain the nature and number of the mental elements. He takes up mental experience, bit by bit, dividing and subdividing, until the division can go no further. When that point is reached, he has found a conscious element. (1896/1899, p. 16)

Titchener regarded the introspective method as a psychological analog to chemical or anatomical analysis, supposedly revealing the structure of perceptual experience. "Structuralism" and the introspective method both failed to achieve their goals, however.

Structuralism and introspection depended on several important assumptions, including the following two:

1. Experience was assumed to be composed of sensations—products of the senses rather than properties of environmental objects.
2. Sensory experience was thought to be composed of discrete elements defined independently of their context. Thus, perceived objects, events, scenes, and pat-

terns were regarded as compositions of elementary sensations—analogue to molecular structures of chemical elements, or to anatomical structures of cells, organs, and so forth.

From the personal perspective of an observer, visual experiences usually seem to be composed mainly of environmental objects and events. The method of introspection failed partly because subjects found it difficult to describe sensations rather than stimulus objects; they too often made “stimulus errors” by describing stimulus objects rather than the sensations *per se*. Vision research has progressed more rapidly by focusing on the objects of perception rather than sensory experience as such. Perhaps the objective and subjective aspects of perception cannot even be clearly distinguished.

Psychological structuralism largely disappeared after Titchener’s death. Nevertheless, relatives of the two ideas above have survived, clothed in modern concepts of sensory, perceptual, and cognitive processes. Persisting ideas about the physiological components of perception derive from implicit intuitions about the material and causal bases of visual phenomena. Empirical support for these two ideas is actually very limited. The supporting rationale is mainly just implicit in the conceptual background of many scientific perspectives.

Material objects, immaterial relations, and “the really hard problem”

From the perspectives of most scientists, visual phenomena have properties quite different from those experienced by observers. Vision occurs through the actions of material mechanisms that transfer energy by optical, physiological, chemical, and neural processes. If visual phenomena have meanings and qualities, then these properties must, in the standard scientific view, be immaterial additions produced by inference, memory, cognition, and emotion.

The problem of understanding how material processes of the eye and brain produce meaningful experience, with properties of meaning, quality, and value, is an abiding and fundamental problem in science and philosophy. Flanagan (2007) identifies this as “the really hard problem.”

Properties of meaning, sensory quality, and affective value are seemingly unobservable—to the scientist on the outside at least—and vision scientists typically ignore them for that reason. But what, exactly, is observable? Observables are often thought to be objects and events with spatial and temporal dimensions. Thus, vision scientists manipulate and measure “stimuli” (environmental objects and events or optical patterns on the eyes) and record “responses” (discriminations of stimuli or physiological responses in nerve cells and brain areas).²

Individual stimuli and responses do not have directly observable properties of meaning, quality, or value. Relations among stimuli and among responses, however, certainly can permit inferences about such immaterial properties. Physiological responses in certain brain areas are also found to correlate with certain stimuli that elicit emotional behaviors or judgments. A contemporary example: Mormann et al. (2011) found that neurons in the human amygdala responded selectively and

with shorter latency to a stimulus category consisting of (pictures of) animals (both aversive and cute) but not to other categories of persons, landmarks, or inanimate objects; and similarly selective responses were not found in other areas of the brain. Converging evidence from clinical, behavioral, and neurophysiological studies supports the role of the amygdala in emotional responses. We can infer that the human subjects probably perceived affective properties of the animal pictures. Did the experimenters observe such affective properties? Or are affective phenomena necessarily only subjective, and not directly observable?

A broader question is whether immaterial properties are observable. Are observable objects and properties only those things that are measurable on well-defined physical variables such as length, duration, wavelength, mass, and energy? Implicitly if not explicitly, scientists have often represented perceived patterns as composed of sensory elements, specified by individual receptors at given spatial and temporal locations. Patterns as such are sometimes treated as not directly observable. Optical patterns, for example, can be represented as arrays of intensity values at discrete spatial and temporal positions, as in photos and movies recorded by cameras. Much of vision science has proceeded from just such representations directly analogous to the image arrays in cameras.

What, then, is the status of *motion* as a visual phenomenon? Efforts to answer this question have significantly influenced vision science. Motion is, after all, a *relationship* among material “stimuli” at particular spatial and temporal positions. Can the change itself be considered a fundamental visual property? Psychologists and physiologists have not always embraced this idea. Historically, many scientists have intuitively preferred to think of perceived motion as an inference from a sequence of stimuli at discrete spatial and temporal positions. Spatial and temporal positions have sometimes, in both past and present, been regarded as physically more fundamental than relationships in space-time. Accordingly, the phenomena of perceived motion have had a pivotal place in the history of vision science.

Many converging lines of psychophysical and physiological evidence show convincingly that motion constitutes a fundamental visual phenomenon, not derived from more elementary sensations at well-defined spatial and temporal positions. A review of the extensive literature is beyond the scope of this chapter, but many helpful collections and reviews are available, including Jansson, Bergström, and Epstein (1994), Epstein and Rogers (1995), Sekuler (1996), Mather, Verstraten, and Anstis (1998), Wade (1998), Westheimer (1999), Lappin and van de Grind (2002), Simoncelli (2004), and Warren (2004). The fundamental role of motion in vision is no longer in doubt, but the transformation from optical patterns in the eye to coherent perceptions of moving objects involves unknown steps.

Motion involves a change in spatial position. How, then, are spatial positions defined? Is the visual frame of reference for motion given by anatomical coordinates of the eye or by features of the surrounding optical pattern? Different frames of reference have different implications for the visual mechanisms that convert optics to perception. Different frames of reference for spatial structure and motion may emerge at different “stages” of visual processing—for example, from 2D to a “2½D sketch” and then a 3D framework (e.g., Marr, 1982).

Analogous but seldom articulated issues have influenced the history of research on virtually all aspects of perception, including space, form, and environmental objects and events. Marr offered a clear hypothesis about the frame of reference for vision:

. . . a process may be thought of as a mapping from one representation to another, and in the case of human vision, the initial representation is in no doubt—it consists of arrays of image intensity values as detected by the photoreceptors in the retina. (Marr, 1982, p. 31)

This statement describes a common belief among vision scientists, but it is an assumption. The topology of the surrounding optical pattern affords other frames of reference.

Do visual phenomena begin as 2D images spatially organized by the eye rather than by environmental objects and events? If so, then the perceived coherent organization, meaning, qualities, and values of our surroundings are necessarily products of our eyes, brains, memories, and imaginations. If visual phenomena begin this way, then phenomenology seems only an entertaining diversion from the sciences of neurophysiology and cognitive science. If visual phenomena are products of the physical, chemical, physiological, and neural mechanisms of the eye and brain, then understanding how organization, meaning, quality, and value arise from these material mechanisms really is a really hard problem.

Perhaps, however, visual phenomena do not arise from such impoverished beginnings. If optical patterns are structured by the environment rather than by the eye and brain, then perhaps the subjective impression that we observe environmental objects and events “directly” is plausible after all.

Is *information* a material thing—a “stimulus” or signal or symbol? Do relationships constitute information? If so, what relationships? How, exactly, do material mechanisms of the eye and brain, at discrete locations in space and time, carry “information” about the organization of the visual world?

What, indeed, is the relationship between the material world and the mental world? How can the material processes of our eyes and brains support the mental world of our knowledge and experience? Do perceived objects and events have observable properties of meaning, quality, and value? Are meaning, quality, and value fundamentally immaterial ideas created by the mind? Does our experience of a meaningful world with qualities and values belong at all within the realm of science? Does science include only material objects and events? What is the place of visual phenomena within the realm of science?

Questions about the nature and content of visual phenomena entail basic questions about the nature of both *information* and *observation*. Concepts of information and observation are fundamental to the science of visual perception. The present chapter focuses on contrasting paradigms of vision research known as “inferential” and “ecological” approaches. These contrasting approaches diverge at differing conceptions of information and observation.

Inferential Theories

Logical responses to proximal stimulation

The inferential approach to the problem of vision encompasses a large collection of visual phenomena, theoretical concepts, research methods, results, and theoretical explanations. These phenomena and ideas usually entail a conception of the visual process as the product of material mechanisms. The visual process is seen to begin with an objective physical “stimulus” of the eye’s photoreceptors. The scientific problem, then, is to discover mechanisms by which stimuli produce coherent experience of environmental objects and events.

The mainstream scientific approach to the study of vision has been, and continues to be, such a materialist conception. This approach has developed over a long intellectual history that gained strength especially with the development of 19th century science. Almost any current textbook on perception begins with a description of vision that exemplifies this approach. The philosophical and scientific history of the inferential approach is beyond the scope of the present chapter, but we can identify a few logical landmarks in its development.

Photoreceptors in the eye convert optical stimulation into physiological responses. Both stimuli and responses are regarded as objective packets of energy at specific locations in space and time. Sensory evidence about “distal” objects and events in the environment is, therefore, indirect and incomplete. Thus, perceived environmental objects and events must be inferred from the limited sensory evidence.

Coherent 3D organization of perceived environmental scenes, experienced “qualia,” and meanings must be created by the brain and mind. The objective matter and energy of physical stimuli and physiological responses have none of the properties of conscious experience. Accordingly, some version of dualism seems inevitable. The early empiricist philosophers—for example, Locke, Berkeley, Hume, Mill—all accepted that the material world and the perceptual experience of that world are incommensurate, neither reducible to the other. Lawful forces of nature, involving time-dependent material interactions, could be seen to govern causal events in the physical world. But these physical processes were evidently insufficient to explain the perception of organized scenes of solid objects and motions in a 3D world with meanings and qualities. Additional processes were needed, involving learned associations and rules of inference. Even the rationalist philosophers—for example, Leibniz and Kant—who disputed the empirical origins of perception accepted the premise that perceptual experience could not be reduced to the physical processes of the natural world.

The development of sensory physiology in the 19th and 20th centuries significantly strengthened the conception of perceptual experience and knowledge as constructed by inference from limited sensory evidence. As observations and understanding of the biophysics, anatomy, and physiology of the visual system have developed exponentially over the past 150 years, the resulting picture of visual mechanisms has become much clearer. Contemporary vision sciences are highly interdisciplinary—involving sensory physiology, neuroscience, psychophysics, cognitive science, computer science, optometry, and ophthalmology. The expanded and

clearer modern picture of the material aspects of vision has not revealed the origins of perceptual experience and knowledge of the world, however. In effect, the need for supporting roles of learning and inference has grown with increased knowledge of the material processes of vision. Computational and physiological details of such inferential processes remain largely unspecified, however.

Helmholtz (1910/1925) provided a memorably clear statement of the inferential conception of visual perception:

The sensations aroused by light in the nervous mechanism of vision enable us to form conceptions as to the existence, form and position of external objects. These ideas are called visual perceptions. (p. 1)

Perceptions of external objects being therefore of the nature of ideas, and ideas themselves being invariably activities of our psychic energy, perceptions also can only be the result of psychic energy. Accordingly, strictly speaking, the theory of perceptions belongs properly in the domain of psychology. (p. 1)

The general rule determining the ideas of vision that are formed whenever an impression is made on the eye . . . is that such objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism, . . . (p. 2)

Thus far the sensations have been described as being simply symbols for the relations in the external world. They have been denied every kind of similarity or equivalence to the things they denote. (p. 18)

These inductive conclusions leading to the formation of our sense-perceptions certainly do lack the purifying and scrutinizing work of conscious thinking. Nevertheless, in my opinion, by their peculiar nature they may be classed as conclusions, inductive conclusions unconsciously formed. (p. 27)

From Helmholtz's perspective as a physicist, physiologist, and mathematician, perceptual phenomena obviously required explanatory principles that were essentially cognitive—involving learning, symbolic representations, and rules of reason. Helmholtz explicitly rejected the idea that environmental objects and properties were somehow “directly” observed. The visual nervous system was believed to provide a symbolic representation of the perceived world. The natural laws governing material interactions in space and time were (and are) insufficient to account for logical operations on symbols, but principles from the logical or mental realms seemed necessary to explain the “ideas” of perception. Similar beliefs prevail today.

Physiological mechanisms

The concepts of contemporary vision science are supported by vastly expanded knowledge of both neural mechanisms and the feasibility and power of symbolic computations in physical systems. Nevertheless, the rationale for modern versions of the inferential approach is similar to that of Helmholtz. Key concepts in modern versions of inferential theories have included (a) specially tuned *receptive fields* of individual neurons which may encode specific stimulus features (e.g., Barlow, 1972); (b) multiple *cortical areas and visual pathways* with specialized functions involved

in perceiving objects, colors, motions, space, and controlling motor actions (e.g., Chalupa & Werner, 2004; Hubel & Wiesel, 2005; Livingston, 2002; Zeki, 1999); (c) linear systems and filters for abstracting the spatiotemporal organization of complex optical patterns (e.g., Cornsweet, 1970); (d) Bayesian statistical methods for integrating current sensory data with statistical evidence from past experience and other sensory cues to identify a likely interpretation of the environmental cause of the current sensory data (e.g., Purves & Lotto, 2003; Trommershauser, Kording, & Landy, 2011); and (e) computational theories (e.g., Churchland & Sejnowski, 1992; Marr, 1982). Wandell's (1995) book, *The Foundations of Vision*, develops many of these themes clearly. His concluding chapter on "Seeing" begins with the statement that "Seeing is a collection of inferences about the world" (p. 387).

The conversion from the material processes of the brain to the supposed symbolic processes of vision remains murky, however. A common idea has been that categories of stimulation are "made explicit" by the responses of neurons with receptive fields specially tuned to particular "trigger features" (Barlow, 1972). The currently known encoding of optical information by receptive field properties of single neurons is insufficient, however, to specify environmental objects, events, and scenes. Such a neural representation would not be invariant under changes in observational conditions associated with the vantage point and environmental conditions such as illumination and context. Probably few contemporary vision scientists believe in the sufficiency of the "neuron doctrine" as articulated by Barlow in 1972, but this idea has not yet been replaced by a clear and specific alternative.

Contemporary scientists all recognize that cortical areas with specialized functions must play a critical role in perception, and most also recognize that currently available knowledge about the brain is insufficient to account for the perception of environmental scenes. Many vision scientists seem to regard the brain mechanisms of vision as performing analog-to-symbolic transformations. The need for symbolic representations of visual stimulation is implicit in many current ideas about the mechanisms of visual perception.

A related recent development is the research program by a prestigious group of neuroscientists specifically aimed at identifying "neural correlates of consciousness" (NCC)—the minimal neural events or structures necessary and sufficient to produce a conscious percept (e.g., Crick & Koch, 1995; Dehaene, Sergent, & Changeux, 2003; Kim & Blake, 2005; Koch & Crick, 2001). The research program is essentially empirical, not driven by an explicit theory, but the motivating hypothesis is that conscious experience must have discoverable material bases. Neural correlates of conscious awareness are not necessarily symbolic, though that possibility is encompassed by the NCC effort.

A major research aim in the vision sciences is to elucidate the underlying neurophysiological processes. One need not believe that the material brain creates immaterial experience to see that optical information about the world must be communicated by the physiological mechanisms of the eye and brain. And one need not believe that explanation requires reduction of macroscopic visual phenomena to microscopic neurophysiological mechanisms to see that the correlation between these two levels of analysis constitutes a major scientific frontier. The inferential conception of perception has encouraged research on the neurophysiology of vision; and this line

of investigation is fruitful whether or not perceptual and brain processes are regarded as inferential.

Thirty years ago, the brain was commonly regarded as a collection of special-purpose local mechanisms for (symbolically) encoding local sensory data, recognizing familiar data patterns, inferring the meaning of these patterns, and choosing appropriate responses. The functions of most nerves and cortical areas were regarded as largely fixed by genetics and early experience. Now, we are coming to understand the brain as a vast and interconnected array of networks dynamically organized according to the particular task—analogue to a symphony orchestra, with activity patterns that change depending on the music and skills of the players. We are coming to understand that the functions of component parts at all levels of the brain, from molecules to networks, are flexible and can vary with the context in which they are used—analogue to the dependence of musical sound from a given instrument on the style with which it is played, on the sounds from surrounding instruments, and on the acoustics of the room.³ The brain seems now less like a symbolic logic machine than like an adaptive system of networks for recognizing, reproducing, and organizing patterns.

The interdisciplinary blending of neuroscience and psychology—neuropsychology—has many important applications. A recent book by Oliver Sacks (2010), *The Mind's Eye*, offers many compelling illustrations of the scientific, clinical, philosophical, and personal implications of clinical phenomena such as visual agnosia and alexia. Localized brain damage from a stroke, tumor, or injury may cause the sudden disappearance of what had seemed an automatic ability to recognize familiar objects such as faces, letters, words, or musical symbols. Visual functions are certainly tied to particular brain regions, but we have also discovered that brain mechanisms are plastic, that a given brain region can acquire a new function, and that the brain can accomplish old skills with new mechanisms.

Computational theory

Another theoretical strategy in vision research is to bypass the neural processes for converting physiological to symbolic representations, and simply treat all visual processes as symbolic operations. This strategy is used in much of the research in computer vision. Despite the intuitive simplicity and immediacy of everyday visual perception, almost a half-century of intensive research has failed to develop reliable and general computations by which machines can perceive environmental objects, events, and spaces. Significant gaps persist in our understanding of the logic and mathematics of vision for mapping optical input to perceptual output. A recent *New York Times* article offers a nontechnical review of the limitations of current machine vision and robotics (Markoff, 2011). Helmholtz's (1910, p. 2) principle—that “such objects are always imagined as being present in the field of vision as would have to be there in order to produce the same impression on the nervous mechanism”—now seems too vague to count as an explanation. No one has yet shown how to do this in natural and general environmental conditions.

Research on machine vision has usually focused on computational processes rather than the information they use. The optical input has often been regarded as physically

given, often represented as a planar array of intensities spatially structured by 2D Cartesian coordinates. The key computational problem, however, may be to find a suitable representation of the optical input, where the image structure reflects the environmental structure. Recent research in ecological optics shows how this can be done (Lappin, Norman, & Phillips, 2011).

Selective attention, information processing, and the demise of behaviorism

Behaviorism was the dominant force in American psychology in the first half of the 20th century. The theoretical strategy was to characterize all psychological phenomena in terms of associative relations among stimuli and responses. Behaviorists' emphasis on observable stimuli and responses was opposed to the subjectivity of introspection and phenomenology. Perception, attention, thought, and language were usually seen as outside the scientific domain. The Gestalt approach maintained some interest in perception, but this approach had limited influence. The Gestalt focus on self-organizing sensory patterns, where "the whole is more than the sum of the parts," seemed both immaterial and unhelpful in the *stimulus-response* (S-R) analysis of behavior; and the Gestalt laws had a limited range of applications and limited power for explaining learning and behavior. By the 1950s and '60s, however, empirical and theoretical insufficiencies of behaviorism had become evident to growing numbers of psychologists. Nevertheless, concepts of "stimuli" and "responses" remain common in contemporary perceptual theory.

Scientific developments in the last half of the 20th century yielded both the demise of behaviorism and renewed interest in perception. One such development involved experimental demonstrations of the role of selective attention in perception, learning, and memory. In effect, the causal sequence from stimulus to response was reversed; in effect, a "stimulus" depends on an attentional "response." The S-R conception of perception was undermined by the phenomena of attention. Before behaviorism, selective attention had been recognized as critically important (James, 1890/1981), but experimental research on attention was dormant in the first half of the 20th century. Now, it is a principal area of perceptual research.

Another major influence on both perception research and behaviorism came from cybernetics, information theory, and computer technology. This profoundly important intellectual development offered new ideas about both experiments and theory that departed sharply from the behaviorists' concepts of materially determined causes and effects. Perception, cognition, and decision making were now recast as phenomena of *information processing*. "Information" does not imply symbolic representation, but that was a common understanding in the "information-processing" approach to perception and cognition.⁴ By the 1970s, the information-processing approach dominated research on perception and cognition. The new analyses of computational processes were a major break from the deterministic constraints of behaviorism, opening the door to new ideas about perceptual phenomena and processes.

The information-processing ideas significantly strengthened and expanded inferential approaches. Interpretations of symbolic representations require logical rules and heuristics. Symbolically represented environmental objects and events and their

qualities and meanings were necessarily perceived “indirectly,” by inference and interpretation.

Theoretical and experimental efforts in the information-processing approach were focused on processes rather than information as such. Spatial and temporal characteristics of symbolic information were largely irrelevant. The behaviorists’ concepts of “stimuli” and “responses”—material objects and events at specific spatial and temporal locations—remained useful in the new paradigm, even though their deterministic connotations were abandoned. In practice, “information” was usually synonymous with a “stimulus.”

The information-processing approach has facilitated research on the limited but selective “capacity” of perception. Vast experimental evidence clearly demonstrates that the capacity of visual perception is quite limited—far more than subjective experience suggests. For example, an array of alphanumeric or geometric forms can be displayed for a short duration (e.g., 50–1,000 ms), with the observer’s task being to identify one or more target forms designated about 200 ms or more after the display. Such tasks are usually easy if the display contains only three or four items, but errors increase rapidly as the number of initial display items increases beyond four. A limited span of “apprehension” or “visual working memory” estimated by this method is typically about three or four items (e.g., Fougner, Asplund, & Marois, 2010; Luck & Vogel, 1997; Miller, 1956; Woodman & Vogel, 2008). Comparable results are obtained with many variations in specific stimuli and responses.

The restricted scope of perception is also well illustrated in experiments on “change blindness” (see Simons, 2000)—where observers consistently fail to detect optically large changes in photos or movies that do not alter the meaning of the scene. Mack and Rock (1998) conducted experiments on “inattention blindness”—where observers failed to detect features of images that are optically quite visible but seemingly irrelevant—and concluded that perception requires attention. Simons and Chabris (1999) reported a dramatic example: A gorilla strolls through a scene of humans bouncing balls to one another, stops in the middle of the scene, looks at the camera, beats his chest, and strolls away. When an audience is asked to attend closely to one group of the humans, about half the audience fails to notice the gorilla. Our subjective phenomenology is misleadingly incomplete: We don’t know what we don’t perceive (Levin, Momen, Drivdahl, & Simons, 2000).

Our perceived worlds are limited by our attention, but our selective attention is flexible. We recognize coherent and meaningful patterns that are organized by many converging factors—spatiotemporal patterns, prior knowledge and familiarity (e.g., Staller & Lappin, 1981), our interests and purposes, and competing patterns of organization. Accordingly, we cannot yet quantify the capacity of attention and perception. Various feature-based, object-based, and space-based models of visual attention have been proposed, but these are not independent of either the optical organization or the observer’s prior knowledge. The persisting difficulties in specifying visual capacity limits are reminiscent of the difficulties of structuralism in developing a general theory for analyzing and describing the structure of visual experience.

A shortcoming of the information-processing approach is that symbols are poorly suited for representing spatial and temporal patterns. For this reason, the

information-processing paradigm has not encouraged or assisted research on many basic aspects of perception—for example, spatial vision, shape perception, motion perception, natural scene perception, visual-motor coordination, esthetics, meaning, or affective properties.

Inferential concepts from phenomenology

In addition to the preceding influences, ideas about perception as inference have also developed from phenomenological perspectives. Two important developers of inferential concepts have been Richard Gregory (1998, 2009; Gregory, Harris, Heard, & Rose, 1995) and Irvin Rock (1983, 1997a, 1997b).

This line of research has been especially interested in phenomena in which perceived spatial organization and forms differ from what might be expected based on the 2-dimensional images. These apparent discrepancies between the perceived spatial structure and what is assumed to be the true image structure are attributed to post-visual cognitive interpretations. The present chapter cannot do justice to the large volume of evidence and writing developed by these two investigators and their students, but a few illustrations are provided in Figure 1.1.

Perceived spatial relations and forms in images such as those in Figure 1.1 are often taken as evidence that perception requires inference. The ecological approach, however, has paid much less attention to such phenomena. By studying how the structure of images constitutes information about the structures of surrounding scenes, the ecological approach has come to different conclusions about the role of inference.

Ecological Theories

A contrasting conception of perception has developed by examining how the spatio-temporal structure of optical images reflects the structure of the environment. The “ecological” and “inferential” approaches stem from distinctly different descriptions of (a) optical information and (b) the roles of the environment and the observer.

Like the inferential approaches of Gregory and Rock, the ecological approach also adopts the phenomenological strategy of investigating “why things look as they do.” A more important emphasis, however, is on the question “how can animals act as they do” in using optical information to interact with a changing environment.

The following ideas in the ecological approach diverge from the inferential approach:

1. *Optical information is given by spatiotemporal structure:* The optical input to vision consists of spatiotemporal patterns, rather than energies at spatial and temporal locations. Change and motion are fundamental, not derived from local energy measures.
2. *Optical images constitute information about both environmental structure and the observer’s position and motion within the environment:* Optical images, especially as produced by moving objects and moving observers, are mutually determined

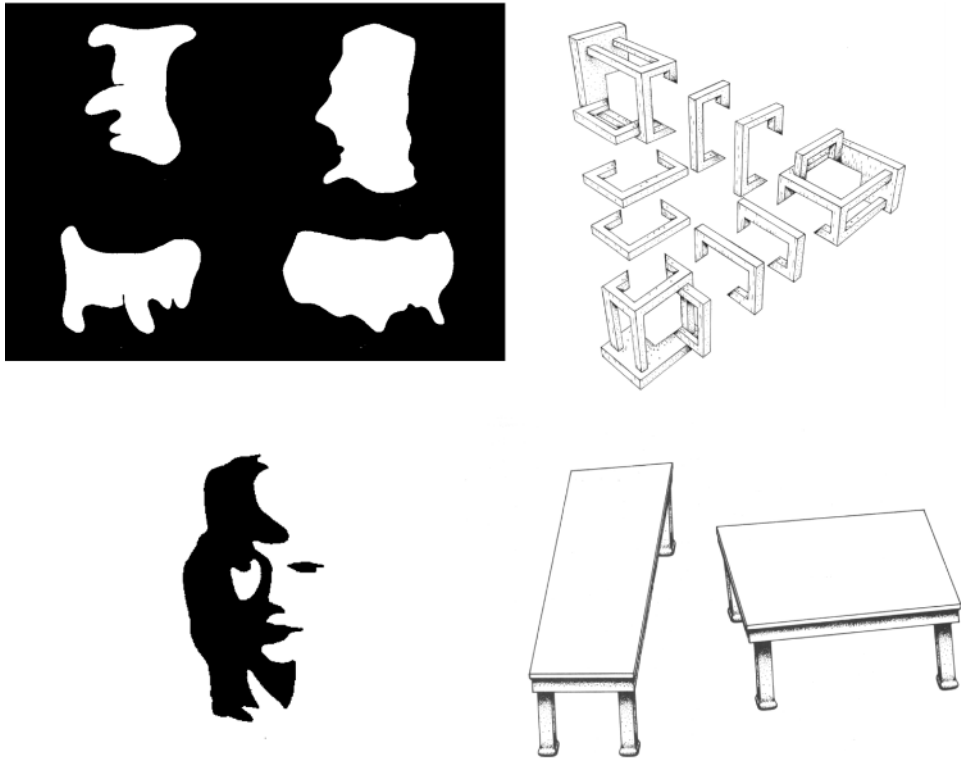


Figure 1.1 Perceived spatial forms often differ from the 2D Euclidean image structure. *Upper left:* Perceived shapes depend on orientation (from Rock, 1997b, p. 140). *Upper right:* Occlusion boundaries may be sufficient for perceiving solid objects, even impossible 3-dimensional objects (from Penrose, 1995, p. 333). *Lower left:* Boundaries between figure and ground permit perceptions of multiple alternative spatial forms (from Ferrante, Gerbino, & Rock, 1997, p. 167). *Lower right:* Perception of 3-dimensional form may preclude perception of even simple 2-dimensional forms. The parallelograms formed by the two table-tops are identical in the image plane (except for planar rotation) (from Shepard, 1990, p. 48).

by environmental structure and by the observer's positions within the environment.

3. *Perceived environmental properties are specified by optical variables at the retina:* Gibson hypothesized that "there is always some variable in stimulation (however difficult it may be to discover and isolate) which corresponds to a [perceived] property of the spatial world" (1950, p. 8).⁵
4. *Perception of invariants:* Spatial forms may be visually defined by the transformations under which they remain invariant. Optical transformations produced by moving objects and observers specify spatial forms that are invariant under motion. Vision is directly sensitive to spatial structure defined by invariance under motion.
5. *Direct perception of environmental objects and events:* Visual phenomena are composed of environmental objects and events rather than physiological events in

eyes or brains. The surrounding environment is perceived “directly” rather than “indirectly” through symbolic representations in the visual brain.

6. *Perceiving and acting are interdependent*: “We perceive in order to act, and we act in order to perceive” (Herb Pick, personal communication). The ecological approach has motivated research on visually guided locomotion and a search for common organizing principles for both visual perception and motor control.
7. *Meanings, qualities, values, and affordances for action are directly perceived*: Optical information about environmental objects and events is contingent on the observer’s aims, actions, and attentions. The meaningfulness of visual phenomena for the observer is, therefore, inherent in the optical information at the retina.

The ecological approach is more closely associated with the research and writing of James Gibson than anyone else. All the preceding ideas (and others) were clearly articulated and developed in his three books—*The Perception of the Visual World* (1950), *The Senses Considered as Perceptual Systems* (1966), and *The Ecological Approach to Perception* (1979). Reed and Jones (1982) provide a collection of his essays on key ideas, and Reed (1988) describes the historical context and development of Gibson’s research and thinking. Naturally, the ecological approach has intellectual origins before Gibson, including both Gestalt theory and American functionalism.

The ecological approach has also been significantly enriched by other lines of research, including research on (a) motion perception, for example by Gunnar Johansson (1950/1994a, 1973) and Hans Wallach (Wallach & O’Connell, 1953); (b) ecological optics, especially as developed by physicists Jan Koenderink and Andrea van Doorn; (c) contemporary psychophysical research by numerous researchers associated with the University of Connecticut (e.g., Geoff Bingham, Claudia Carello, Claire Michaels, Robert Shaw, James Todd, Michael Turvey, William Warren, and others); and (d) the International Society for Ecological Psychology (ISEP). ISEP was founded in 1981, supports a quarterly journal, *Ecological Psychology*, hosts international meetings every two years, and has spawned related organizations and meetings in several countries. Gibson stimulated, directly or indirectly, nearly all of this continuing line of research. The basic theory and evidence now stand on their own, however, independently of Gibson.

An ecological concept of sensory information:

(1) spatiotemporal structure

All theories of visual perception begin with a representation of the input optical information. Ideas 1–4 in the list above describe an ecological conception of sensory information that differs in important ways from the representation implicit in inferential theories.

Light may be regarded as information as well as energy. A basic premise of ecological theory is that the optical information consists of spatial and temporal variations in light, rather than the light energy as such. From the inferential perspective, information is often regarded as a “thing” of matter and energy, located in space and time—a stimulus, signal, symbol, or data point. From the ecological perspective,

however, information is given by the spatiotemporal patterns of light energy. Vision is certainly sensitive to this spatiotemporal organization.

Accordingly, the ecological description of spatial vision is importantly different from descriptions common to the inferential perspective. Inferential theories often assume that optical information may be represented as a 2-dimensional array of intensities at retinally defined positions (Marr, 1982). In contrast, the ecological representation is based on the topology of the optical patterns, where the position of any given point is described in relation to the surrounding optical pattern.

Information about surface shape, for example, is given by second-order spatial derivatives, on both environmental surfaces and their optical images. This second-order structure involves the 2-dimensional neighborhood around each point (Koenderink & van Doorn, 1992a; Lappin & Craft, 2000; Lappin et al., 2011). As Koenderink and van Doorn (1992b; Koenderink, 1990) have pointed out, higher-order spatial derivatives can be measured directly. Estimating the higher-order relations indirectly by comparing lower-order measures is impractical due to rapid increases in measurement errors. Psychophysical results on visual acuities, shape perception, and other spatial discriminations show that human vision is directly sensitive to second-order differential structure associated with local surface shape (Lappin et al., 2011).

Information has been sometimes misunderstood as inherently symbolic—where the physical form of a symbol is irrelevant and serves merely to distinguish between symbols. Wiener's (1954, 1961) conception of cybernetics, which preceded and guided Shannon's (1948) theory, involved a broader, non-symbolic conception of communication, control, and computation. Applying theories of information to the study of sensory systems certainly requires analyses of non-symbolic physical variations, as recognized by Wiener, Ashby (1963), Attneave (1954), Garner (1962), Meyer-Eppler (1969), and others, including Gibson.

An ecological concept of information: (2) optical images constitute
information about both environmental structure
and the observer's vantage point

The optic information available at a given vantage point within an environmental scene is contained in the *optic array*. The optic array may be described as a very broad hourglass-shaped bundle of light rays in two conical sections converging at their vertices, with the convergence point corresponding to the observer's vantage point.⁶ Each light ray projects in opposite directions from the vantage point, in one direction to a point on an environmental surface, and in the other direction to a point on an image of the surface. Each light ray corresponds to a visual direction. Importantly, the optic array is a function of both the structure of the surrounding scene and the location of the vantage point within the scene; it contains information about both. When the observer's vantage point changes, the optic array also changes.

The optic array is useful in conceptualizing the projective geometry of images because it is based on the location of the vantage point rather than the image orientation, viewing direction, focal length, or retinal eccentricity. As a first approximation, the eye rotates around its optical center (the nodal point), changing the direction of

view without changing its central position.⁷ The optic array should be distinguished from planar images like those in photographs. In planar images, the mapping of visual directions to image positions varies over the image, with the same angular change in visual direction projected to increasingly larger image separations toward the outside edges of a planar image. The spherical surface of the eye, however, preserves correspondence between shifts in visual direction and shifts in retinal position. Images formed with lenses, in eyes as well as cameras, depend on focal length, but the optic array involves neither lenses nor focal lengths.

The optical information available in all directions from a given location can be represented as an *optic sphere*. Suppose that this sphere is reduced in size, approaching a point; then this point-like sphere contains all the optical information available at a given position in a given scene, independent of the viewing direction. Much of this information is not visible at a given moment, of course, as much is behind the head and not visible in the momentary visual field. Gunnar Johansson (Johansson, 1994b; Johansson & Börjesson, 1989) discussed the usefulness of the “optic sphere” concept with a particular emphasis on the visual periphery.

The two eyes are located at slightly different vantage points, and differences between these two optic arrays and spheres constitute information available for stereopsis. At large viewing distances, one can consider a single “cyclopean” optic sphere centered at the observer’s head.

The projective structure of the optic array involves only variations in the relative directions of environmental points, not their distances. Projective information about the relative distances of stationary environmental objects is obtained only by varying the vantage point, through binocular vision or motion parallax. Nevertheless, the optic array preserves information about important spatial relations including connectedness, order, collinearity, parallelism, co-planarity, and relative heights of the vantage point and other objects in the scene. Figure 1.2 illustrates some of these relationships.

Specifically:

1. Points that are collinear in the environment are also collinear in a stationary optic array.⁸ Collinear points in the optic array are almost always collinear in the environment, although deviations from collinearity can be invisible from an “accidental” view, varying only in distance but not viewing direction.
2. Information about parallelism of environmental lines is also preserved in the optic array: Lines that are parallel in the environment converge toward a common point in the optic array. Lines that are parallel with the ground plane, either on the ground plane or at any height above it—for example, train rails, edges of floors and ceilings in a rectilinear building—converge at a common vanishing point on the horizon. The horizon line marks the observer’s eye height. Lines that converge on the horizon from above correspond to directions above the observer’s vantage point; and those converging on the same horizon line from below are below the vantage point. The horizon line corresponds to a great circle in the optical sphere, separating objects above and below the observer’s eye. Thus, for an observer standing in a building corridor, the four edges at the sides between the walls and the ceiling and floor all converge at a common image



Figure 1.2 Elementary characteristics of perspective and the optic array. The optic array preserves information about the observer's eye height and relative position within the scene: All lines parallel to the ground plane project toward the horizon line, which designates the observer's eye height (marked by the red horizontal line in the lower photo). (More generally, all lines parallel to any plane project to a common great circle in the optic sphere.) All lines parallel to any given direction (e.g., marked by the green lines in the lower photo) converge toward a common point. The green lines in this photo are approximately parallel with the camera's viewing direction, but their locus of convergence is independent of the observer's viewing direction. The horizon line and convergence points of parallel lines are independent of the focal length of the lens and the resolution of the image recording system, as well as the particular objects in the scene.

location at the observer's eye height, regardless of the viewing direction or position in the corridor. Tall and short observers see different optic arrays that specify their relative heights.

The optic array does not directly specify the slant of a plane, and does not distinguish the ground plane from other planes. Parallel lines in any direction converge (when extended) at a common vanishing point; and all lines parallel with any given plane vanish at a common great circle in the optic sphere. Nevertheless, the ground plane is identifiable in most human-made and natural scenes through the effects of gravity on orientations of walls, trees, and other objects that are usually perpendicular to the ground plane. Appearances can be made to deceive, but relative heights of the observer and surrounding objects usually are visible.

An ecological concept of information: (3) perceived environmental properties are specified by retinal variables

A guiding principle in James Gibson's development of the ecological approach was that the perceived environment is fully specified by optical information at the retina. Gibson emphasized the importance of this hypothesis in his book *The Perception of the Visual World* (1950). Unfortunately, this idea still seems implausible to most scientists outside the ecological community.

The question of whether sensory information is sufficient or insufficient marks a key distinction between ecological and inferential theories. An important fact about contemporary vision science is that this issue has remained effectively unresolved for at least 60 years. The hypothesis has motivated psychophysical research to identify retinal variables that may account for specific perceptual properties, but psychophysical research and theory that directly addresses this hypothesis has so far been limited. In principle, the issue is empirical, but the debate has been more philosophical than empirical.

The ecological and inferential approaches reach different conclusions about the sufficiency of retinal information partly because they derive from different conceptions of retinal information. Gibson endeavored to show that structural details of the retinal patterns are determined by, and must specify, the spatiotemporal structure of the surrounding scene. The speed and reliability of visually guided performance in piloting planes and in athletics, for example, support his intuitions about the sufficiency of the optical information. Gibson concluded that the information must be contained in "higher-order variables." Identifying these higher-order variables has proven difficult, however, but progress has occurred recently.

Two parts of this research problem are to show (a) a specific correspondence between environmental structure and retinal image structure and (b) that human observers can reliably and precisely discriminate this optical structure. Both parts of this problem require a demonstration that the supposedly informative structure is invariant with changes in other variables that might also account for the perception. Lappin et al. (2011) recently described significant research progress toward both of these subproblems associated with shape perception. Specifically, visible information about local surface shape is associated with the second-order spatial differential struc-

ture of optical patterns of motion, binocular disparity, contours, texture, and shading. This is a promising beginning, but it is only a first step toward showing how optical patterns at the retina determine perceived environmental scenes.

The hypothesis that retinal information must be sufficient is also supported by logical problems with the alternative hypothesis:

1. Visually guided actions: The speed, reliability, precision, and robust variety of perception-action coordination in athletics and in animals as simple as houseflies seem incompatible with the idea that such phenomena require cognitive interpretations.
2. Physical and computational implausibility: A premise of inferential theory is that the correlation between the environment and perception is greater than the correlations between the environment and its images or between the retinal images and perception. This idea seems both physically and computationally implausible. Visual processes can only detect coherent organization of the retinal variables. Correlations between past and present retinal variables may sometimes improve discriminations between alternative objects (via Bayes's theorem), but the bandwidth (resolution and speed) of real-time correspondence between current environmental events and actions cannot be increased by information from past events. Such increases would seem to violate both Shannon's (1948) fundamental theorem about the bandwidth of a communications channel and the second law of thermodynamics.
3. Lack of explanatory value: Generally speaking, appeals to logical inference, intelligent interpretation, previous learning, evolution, heuristics, and other inferential processes have lacked explanatory detail. From ecological perspectives, inferential explanations often seem appeals to magic. The precision, speed, reliability, and robustness of visually guided actions have exceeded the explanatory capabilities of inferential theories.

Accordingly, the ecological approach has studied retinal image information as an explanation for visual phenomena, in contrast to inferential explanations based on the processing of insufficient information. The ecological approach regards retinal information as identifiable only by investigating correlations between environments, images, and perceptions.

An ecological concept of sensory information: (4) perception of invariants

The concept of "invariance" was important in Gibson's (1950, pp. 153–154) early development of the ecological approach, and was often mentioned in subsequent publications. The concept of invariance is fundamental to the definition of "information," and is a basic criterion for identifying spatiotemporal variables that carry information in natural systems (Lappin et al., 2011). Nevertheless, "invariance" has only recently been used to identify visual information.

"Invariance" refers to the permissible transformations of a structure that do not alter its correspondence with the structure in another system. Counterintuitively, the best way to define a "structure" is by the transformation groups that leave it

unchanged. This seemingly indirect definition is rigorous because it avoids arbitrary choices of component elements of a pattern. This method is logically and experimentally powerful because it is deductive rather than inductive: One can begin with a group of transformations under which invariance is required, and then identify the structure that satisfies the requirement.

In theoretical physics and mathematics, invariance is called “symmetry.” All physical laws may be expressed as symmetries—as structural relations that are conserved under specific groups of transformations of observational parameters (Lederman & Hill, 2004). The conservation of energy, for example, is equivalent to the invariance of physical interactions under shifts in time of occurrence.

Thus, in vision science, we may seek to identify structures of environmental objects and their images that remain invariant under transformations of observational conditions. Relevant transformations involve motions of the observer or object in 3D space; eye movements, which change the location of an object’s image in the eye and change visual resolution; changes in intensity and spectrum of ambient illumination; and changes in scene context.

For the problem of shape perception, Koenderink and van Doorn (1997) and Lappin et al. (2011) used this approach to identify information about local surface shape. Specifically, (a) the second-order differential structure of environmental surfaces corresponds to the differential structure of the retinal images of surfaces. And (b) psychophysical experiments have found that human discriminations of local surface shape remain precise under image transformations produced by movements in 3D space. Simpler properties such as depth and surface slant do not satisfy the required invariance and are poorly discriminated by human observers.

Ecological theory of observation: direct perception of environmental scenes

A contentious debate between the inferential and ecological approaches concerns the ecological hypothesis that environmental objects and events are perceived “directly”—specified by optical patterns on the eyes, without “indirect” inferences from ambiguous image cues (Gibson, 1979; Reed & Jones, 1982; Rock, 1997a; Ullman, 1980; Warren, 2005). Whether perception is “direct” or “indirect,” however, depends on descriptions of the optical input and perceptual output.

Considered as physical transformations of matter, energy, and spatial structure, vision seems impossible, even miraculous. The environmental input consists of complex 3D scenes of moving solid objects, which stimulate the eyes in continually changing 2D patterns of light; and these optical images are transformed into patterns of electrochemical events in an almost infinitely vast network of nerve cells, synapses, and brain areas, which then produces subjectively compelling real-time experience of the environmental scenes, discriminations among subtly different objects (e.g., human faces), and coordinations of bodily movements with rapidly moving objects. As a physical process, vision is certainly not “direct.”

As a logical or computational process, however—involving transfers of information—can vision possibly be considered “direct”? The answer depends entirely on how the input and output information is described. Ecological and inferential descriptions of

the input information are sharply different, and their studies of perceptual output usually differ as well.

Visual depth illusions are often cited in support of the inferential approach. However, relative depths, either within or between objects, are indeterminate in optical images, despite subjectively compelling appearances of 3D Euclidean relations among objects. Accordingly, human observers are usually both inconsistent and inaccurate in judging absolute depths, distances between objects, or surface slants (e.g., Koenderink et al., 2001; Lappin et al., 2011; Norman & Todd, 1998).

In contrast, the ecological strategy has focused on phenomena in which perceived spatial relations derive from identifiable optical image information. This strategy has been successful in accounting for shape perception, where a reliable relationship between environmental surface shape and retinal images can be identified (Lappin et al., 2011). The ecological strategy has also been fruitful in research on visually guided movements, as described in the next section.

Choices among alternative descriptions of the input and output information are ultimately empirical. If visual processes transform optical input into perceptual output, then research can identify the input and output information that permits such “direct” transformations.

Ecological theory of observation: interdependence of perception and action

The ecological approach is motivated by the everyday performance of animals in coordinating movements with environmental events and spatial layout. Consider, for example, the optical information that permits piloting a plane (an early interest of James Gibson), driving a car, walking through a thick forest, running to catch a baseball, or a housefly avoiding a flyswatter. Prey animals dart in changing directions to avoid obstacles and escape predators; and predators require complementary information to capture moving prey. The sensory information that enables real-time coordination of perception and action is obviously spatiotemporal and obviously jointly structured by the environment and by the animal’s movements.

Observers’ eyes, heads, and bodies move—to explore a scene, to better see an object, to move toward, around, or away from an object. The observer’s movements immediately change the optical images. Even without active movements of head and body, as in reading or viewing a video screen, vision involves active shifts of visual fixation and attention to sample information from spatially distributed locations. The ecology of observer–environment systems demands coordinated control of attention and action.

Three generic problems illustrate the dynamic optical information for controlling movements relative to moving objects: (a) 1 dimension—anticipating time-to-contact; (b) 2 dimensions—pedestrian and driver navigation, anticipating collisions of planar trajectories; and (c) 3 dimensions—the “outfielder problem,” intercepting a curvilinear trajectory in a different plane.

The changing location and size of an object’s image in the optic sphere provide both spatial and temporal information for guiding bodily movements relative to the object. The azimuth position of the image can be specified by reference to the observer’s locomotion direction; and the image elevation can be represented

in relation to the ground plane. Images of other environmental objects, both moving and stationary, offer additional information that may be used for visual navigation (see Land & Tatler, 2009; Warren, 2004). The following analysis is based on simple aspects of the optic sphere.

1-dimensional trajectories—time-to-contact. Suppose an observer is moving in a straight line toward a target, which may be either stationary or moving. At a constant relative velocity, the time-to-contact is proportional to the distance between observer and target. And the target's image size is also (approximately) inversely proportional to the distance. Thus, if a_1 and a_2 are the angular image sizes of a target object at two successive moments, and if t_1 and t_2 are the times-to-contact with the target at these two moments, then one can easily show that

$$t_1/t_2 \approx a_2/a_1 \quad (1.1)$$

The terms on both sides of Equation 1.1 are scale-free ratios. The rate of decreasing relative time-to-contact equals (approximately) the rate of increasing relative image size.⁹

Lee (1976) showed that this relationship offers visual information for controlling the rate of approach to a target object. Lee and Reddish (1981) pointed out that such information must be used by plummeting gannets, birds that dive ballistically into the ocean from variable heights up to 30 m, often reaching speeds over 50 mph. To avoid injury, the birds must use optical information to fold their wings before hitting the water. Yilmaz and Warren (1995) showed that similar optical information controls human drivers' braking at stoplights.

Colliding trajectories in 2-dimensional space—for pedestrians, drivers, and terrestrial predators and prey. Suppose that the observer and target travel in different directions and velocities in a plane, and that the problem is to anticipate whether the two will collide. Optical information is given by variations in the target's image location. If the image location of the target is constant, and if this image is expanding, then collision will occur. The angle of impact is given by the azimuth and elevation of the target's image. Collision can be avoided if the observer or target changes either direction or speed, thereby causing the target's image to drift.

Suppose, for example, that the observer's direction is 0° azimuth (in the optic sphere), and that the expanding image of a moving target object appears at 135° azimuth, with 0° elevation (at the horizon line). If the target image continues to expand at the same optical position, then collision will occur at a time predicted by the rate of image expansion, as given by Equation 1.1.¹⁰ If the target's image position drifts continually, or if the size of the target image is decreasing, then collision will be avoided. One can easily verify this simple relationship by working backward in space and time from the point at which a collision occurs, increasing the spatial separation between observer and target at constant (but different and arbitrary) velocities. Changes in relative trajectory of observer or target alter the target's optical image location.

Converging trajectories in 3-dimensional space—"the outfielder problem": When the observer and target move relative to one another in a 3-dimensional space, the geometry for predicting their intersection is obviously more complicated. Such visual

control problems are both common and important: A baseball outfielder runs to catch a fly ball headed in a different direction; a hawk dives in pursuit of a rabbit on the ground below; a hiker on a twisting hilly forest trail adjusts her stride to avoid obstacles and maintain balance. The apparently routine ease with which such problems are solved by a wide variety of animals, including houseflies and fish, suggests that the controlling optical information may be simple. For concreteness, we will consider the case of an outfielder catching a baseball.

If the motions of ball and fielder are described in a 3-dimensional reference frame from the perspective of a stationary spectator, then the fielder's visual skills are amazing and difficult to explain. The baseball rises with unpredictable speed to a variable height and distance, then curves and falls with increasing speed. Usually, the falling ball is caught by an outfielder running with a direction and speed adapted to meet the falling ball. Ignoring aerodynamic perturbations, the ball's trajectory is parabolic in a plane perpendicular to the ground plane. Constancy of the physical forces is not critical to the fielder's performance, however.

From the fielder's perspective, the task is simpler. The image of the ball rises from the horizon at an angle that indicates, roughly, the direction the fielder must run to catch the ball (McBeath, Shaffer, & Kaiser, 1995). To intercept the ball as it falls again toward eye height, the fielder must move in a direction and speed to minimize variations in the azimuth and elevation of the ball's image in the optic sphere (see Fink, Foo, & Warren, 2009; McLeod, Reed, & Dienes 2006). One can imagine the fielder as traveling along the ground in a vehicle with a spherical windscreen. Thus, the fielder is free to move his or her eyes and head without changing the image positions of the ball and surrounding scene on the (imaginary) spherical windscreen. The image positions of surrounding objects move as the fielder moves; but image positions of the moving ball can be compensated and stabilized by appropriate movements of the fielder.

The azimuth of the ball's image location is measured most conveniently relative to the fielder's direction of travel. If the plane of the ball's trajectory is approximately perpendicular to the ground, then the fielder's direction and speed must simply maintain a constant azimuth position of the ball's image—just as in the 2-dimensional problem (Fink et al., 2009; McLeod et al., 2006). If the ball's ground speed were constant, then the fielder could maintain a constant optic azimuth of the ball's image by running in the correct constant direction and speed. The ball's ground speed varies, however, with its changing image elevation. Therefore, the fielder's speed and/or direction must also vary to maintain a constant azimuth of the image. In fact, fielders' paths are often curved (Fajen & Warren, 2007; Fink et al., 2009; McBeath et al., 1995; Shaffer & McBeath, 2002).

A successful interception path is also controlled by the elevation of the ball's image above the fielder's optical horizon. Unlike the ball trajectory described by a spectator or camera, the optic elevation of the ball's image for the fielder increases monotonically but at a decreasing rate (see McLeod et al., 2006). As the fielder moves beneath the approaching ball, its increasing optic elevation angle approaches a constant equal to the angle of its descent toward the fielder. If the elevation angle accelerates, then the ball is headed over the fielder's head; and if the elevation angle decreases, then the ball is headed toward the ground in front of the fielder. Chapman (1968)

first pointed out that for a fielder already in the plane of the ball's trajectory, the ball can be caught by running forward or backward so as to maintain a constant rate of increase of $\tan \alpha$, where α is the optic elevation angle—i.e., so that $d^2(\tan \alpha)/dt^2 = 0$ —hence “optic acceleration cancellation” as a name for this strategy. This elevation control strategy generalizes to non-parabolic trajectories (Fink et al., 2009; McLeod & Dienes, 1993), and to trajectories angled away from the fielder (McLeod, Reed, & Dienes, 2001, 2006). As McLeod et al. (2001, 2006) pointed out, however, the optic information is given more directly as a decelerating increase in the elevation angle α . The image elevation approaches a constant value $< 90^\circ$, equal to the angle of the ball's descent toward the ground.

Reference frames and visual mechanisms. Difficulties in understanding visual navigation phenomena such as the outfielder problem depend on the reference frame used to describe the phenomena. The preceding analysis shows that the critical information is given by the azimuth and elevation of the target object's images in the optic sphere at the observer's location. The observer's gaze direction and, therefore, the position of the image on the observer's eye were not considered. What, then, may be the underlying mechanisms?

Evidently, optical information for the outfielder problem is not based on the retinal position of the target image. The “linear optical trajectory” (LOT) theory of ball-catching (McBeath et al., 1995) entails measuring the azimuth relative to the home plate, an angle that changes continuously with movement of both the ball and the fielder. These and other optical relationships between the ball and other environmental objects seem unrealistically complicated. Moreover, Oudejans, Michaels, Bakker, and Davids (1999) found that fielders could reliably catch luminous balls in the dark. Optical information about the surrounding environment is evidently unnecessary.

Retinal coordinate frames also cannot account for the outfielder phenomena. Subjective experience in catching balls and measures of fielders' eye movements and fixations both show that successful catches generally involve visually tracking the ball—see Land and Tatler (2009) for extensive evidence about the role of looking in acting. Thus, eye movements tend to minimize but not eliminate retinal image motions of balls-to-be-caught and other environmental targets to be intercepted and avoided. The specific mechanisms that represent the relative motions of a target object and the body probably do not begin with precise measures of retinal positions.

The “generalized optic acceleration cancellation” (GOAC) theory of McLeod et al. (2006) proposes that fielders maintain fixation on the ball as they run to catch it. Accordingly, visual information about the azimuth and elevation of the ball relative to the fielder would be given by proprioceptive feedback from head and eye movements of the running fielder—involving the vestibular and motor systems and multiple parts of the visual system. The required precision of fixation is not known, however. For pilots of planes and drivers of cars, variations in gaze direction and fixation distance seem allowable and even beneficial. In moving vehicles, the windscreen provides a stable reference frame for the relative azimuth and elevation of a target object; and perhaps this reference frame is critical. Fixation stability might be important when a windscreen is not available. The underlying visual mechanisms are not yet known.

Ecological theory of observation: direct perception of meanings and affordances

The least intuitive idea in ecological theory is that meanings and affordances can be and often are perceived “directly.” From an inferential perspective, based on material mechanisms of the eye and brain, this ecological idea seems incomprehensible. The physical, chemical, and physiological signals involved in vision have no inherent meaning, quality, value, affective significance, or affordance (say “meanings” as a simplified blanket term). Meanings are not material.

The ecological approach, however, leads naturally to ideas that perception is meaningful and meanings are perceivable. These counterintuitive ideas are important for understanding the ecological rationale.

Consider perceptions by persons engaged in conversation, or tennis, or chess, or reading, or exploring, or searching for a familiar face, or engaged in almost any activity, energetic or even sedentary. Performance of such activities rests on selective pattern recognition guided by past experience and comprehension of the task environment. Observing is active, and activities are purposeful. Observed objects and events are inherently meaningful in relation to the observer’s aims, actions, and attentions.

Thus, perceptual processes begin at different points in the ecological and inferential perspectives. The ecological description begins with an observer’s aims and actions in an environmental setting. Inferential descriptions of perception, however, typically begin with the proximal stimulation of an observer’s sensory receptors. Optical stimulation is seen in the inferential view as objectively definable and independent of the environmental context and the observer’s activities. In the ecological view, information is a basic commodity in the animal’s commerce with its environment. Visual information necessarily involves both the environmental context and functional relevance for the active observer.

The ecological concept of information involves both the observer and the environment. Correlated variations at both source and destination are also implicit and basic in the theories of information and communication developed by Wiener (1954, 1961) and Shannon (1949). Based on a relation between receiver and sender, between the observer and the observed, information is inherently subjective. This subjective aspect of information is often overlooked, but it is fundamental. In Shannon’s model, the potential signals and messages have been identified beforehand, and are known by both sender and receiver prior to any given signal transmission. For many pattern recognition problems, however, and in vision science in particular, a key problem is to identify the specific image variables that correlate with variations between environmental objects (see Lappin et al., 2011).

Sensory information is based on variations, not individual stimuli, signals, or symbols as such—on probabilistic variations among events that *might* occur. Individual stimuli carry sensory information only by virtue of discriminating among potential alternatives. The material properties of a stimulus are irrelevant to its information except as they distinguish among alternatives. Thus, the observer’s acquisition of information involves comprehension of potential variations, among sensory signals and among environmental objects.

In short, optical and physiological information is inherently meaningful because it involves the observer's knowledge of the environmental context as well as his or her aims, actions, and attentions in that setting.

Conclusions

Visual phenomena vary with one's perspective, and they are described and explained in different ways by inferential and ecological theories. These contrasting approaches also focus on different phenomena. The inferential approach has often been interested in neurophysiological mechanisms and in the influence of attention on perception. The ecological approach has paid more attention to the environmental setting and to the coordination of perception and action. None of these interests is incompatible with either approach; they are complementary.

The ecological approach typically looks first at the environmental organization of perception, and is sometimes seen as incompatible with physiological research. A priority in the inferential approach is to analyze the material physiological mechanisms; so the ecological approach seems misdirected. From a broader perspective, however, both macroscopic analyses of ecological optics and information and microscopic analyses of neurophysiology are necessary for a full understanding of vision—involving both structure and function.

Similar comments apply to the study of attention. Attention exerts a decisive influence on visual phenomena, but the attention may be described in different ways. Regarded as a particular mental mechanism, visual attention resembles a homunculus that interferes with the reality of an observer's contact with the world. But attention can also be recognized as a necessary and vital aspect of vision, involving active information-acquisition guided by the perceiver's knowledge and purposes in the world. Indeed, environmental information has no objective definition independent of the observer's knowledge and purposes. The ecology of attention is a vital aspect of visual phenomenology, and it merits better scientific understanding.

Nevertheless, the inferential and ecological approaches also differ more fundamentally. Inferential theories usually derive from materialist conceptions of both visual mechanisms and visual information. From an inferential perspective, information is usually composed of individual stimuli, signals, and symbols. The ecological concept of information, however, is immaterial—based on variations and on coordinated structures of variation in the environment, optic images, physiological patterns, and perceptual discriminations and actions. Such an ecological understanding of information and perception seems both theoretically sound and necessary for explaining many visual phenomena.

Notes

- 1 The painter Chuck Close used this descriptive phrase in a PBS Television program devoted to the neuroscience of creativity, hosted by Charlie Rose and organized by Eric Kandel, October 28, 2010.

- 2 The concepts of “stimuli” and “responses” require scrutiny, which we postpone for the moment. For now, we may understand these terms simply to mean material and observable objects and events.
- 3 The analogy between the brain and a symphony orchestra was suggested recently during a panel discussion of “Neurological, psychiatric, and addictive disorders” in the Charlie Rose Brain Series 2, November 3, 2011, PBS Television.
- 4 By definition, symbolic representation is not based on any physical correspondence between a symbol and its referent. Symbol-processing operations are governed by rules of logic rather than laws of nature. These rules involve neither mass, time, space, nor energy.
- 5 Jim Todd (personal communication) brought this statement to my attention. He also pointed out that a valid version should refer to a correspondence between stimulation and *perceived* properties of the environment.
- 6 This point may be the aperture in a pinhole camera, or the nodal point of an eye or camera. When images are focused by lenses, as in eyes and cameras, the directions of the light rays are not straight lines, and the light projecting to a given image point arrives from multiple directions. Such complexities associated with lenses may be temporarily ignored in considering the projective geometry of the optic array.
- 7 The eye’s center of rotation is near but not precisely at the nodal point. Rotations of the eye within a stationary head can provide a small amount of depth information from motion parallax, but this small effect can be ignored for present purposes.
- 8 When environmentally collinear points are projected onto a spherical surface such as the back of the eye, they form a geodesic, but are not actually collinear in Euclidean space. If the points are projected onto a planar surface, as in a camera, then the image of the points will be collinear in that plane.
- 9 The angular image size, a , is given more precisely by $a = 2 \arctan(S/2D)$, where S is the linear object size perpendicular to the visual direction, and D is the distance of the target object from the observer. When $S \ll D$, then the angle a is closely approximated by $a \approx (S/D)$, and $a1/a2 \approx D2/D1$. This relationship is often written as a differential equation, $\tau = a/(da/dt)$, where τ is the time-to-contact at a given velocity, a is the visual angle between any two points on the object, and da/dt is the derivative of the angle with respect to time (Lee, 1976).
- 10 Geometry of the 2-dimensional “pursuit curve” has been well described in the mathematical literature (e.g., <http://mathworld.wolfram.com/PursuitCurve.html>). The classical pursuit curve assumes that the predator’s pursuit is always directed at the prey. Thus, this pursuit strategy maintains a constant optical location of the prey’s image at 0° azimuth relative to the predator’s direction. Pursuit also succeeds if the azimuth of the prey’s image is a different constant.

References

- Ashby, W. R. (1963). *An introduction to cybernetics*. New York: Wiley.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*, 183–193. doi:10.1037/h0054663
- Barlow, H. B. (1972). Single units and sensation: A neuron doctrine for perceptual psychology? *Perception*, *1*, 371–394. doi:10.1068/p010371
- Chalupa, L. M., & Werner, J. S. (Eds.). (2004). *The visual neurosciences*. Cambridge, MA: MIT Press.
- Chapman, S. (1968). Catching a baseball. *American Journal of Physics*, *36*(10), 868–870.

- Churchland, P. S., & Sejnowski, T. J. (1992). *The computational brain*. Cambridge, MA: MIT Press.
- Cornsweet, T. N. (1970). *Visual perception*. New York: Academic Press.
- Crick, F., & Koch, C. (1995). Are we aware of neural activity in primary visual cortex? *Nature*, 375(6527), 121–123. doi:10.1038/375121a0
- Dehaene, S., Sergent, C., & Changeux, J.-P. (2003). A neuronal network linking subjective reports and objective physiological data during conscious perception. *Proceedings of The National Academy of Sciences of the United States of America*, 100(14), 8520–8525. doi:10.1073/pnas.1332574100
- Epstein, W., & Rogers, S. (Eds.). (1995). *Perception of space and motion*. New York: Academic Press.
- Fajen, B. R., & Warren, W. H. (2007). Behavioral dynamics of intercepting a moving target. *Experimental Brain Research*, 180(2), 303–319. doi:10.1007/s00221-007-0859-6
- Ferrante, D., Gerbino, W., & Rock, I. (1997). The right angle. In I. Rock (Ed.), *Indirect perception* (pp. 163–170). Cambridge, MA: MIT Press. (Original work published 1995: Retinal vs. environmental orientation in the perception of the right angle. *Acta Psychologica*, 88, 25–32)
- Fink, P. W., Foo, P. S., & Warren, W. H. (2009). Catching fly balls in virtual reality: A critical test of the outfielder problem. *Journal of Vision*, 9(13), 14, 1–8. doi:10.1167/9.13.14
- Flanagan, O. (2007). *The really hard problem: Meaning in a material world*. Cambridge, MA: MIT Press.
- Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of visual working memory? *Journal of Vision*, 10(12), 27, 1–11. doi:10.1167/10.12.27
- Garner, W. R. (1962). *Uncertainty and structure as psychological concepts*. New York: Wiley.
- Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *The Psychological Review*, 63(3), 149–159. doi:10.1037/h0042992
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton-Mifflin.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston: Houghton-Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton-Mifflin.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley.
- Gregory, R., Harris, J., Heard, P., & Rose, D. (Eds.). (1995). *The artful eye*. New York: Oxford University Press.
- Gregory, R. L. (1998). *Eye and brain: The psychology of seeing*, 5th ed. New York: Oxford University Press.
- Gregory, R. L. (2009). *Seeing through illusions*. New York: Oxford University Press.
- Helmholtz, H. von. (1910/2005). *Treatise on physiological optics: Vol. 3* (J. P. Southall, Trans., 1925). Mineola, NY: Dover.
- James, W. (1981). *The principles of psychology*. Cambridge, MA: Harvard University Press. (Original work published 1890)
- Jansson, G., Bergström, S. S., & Epstein, W. (Eds.). (1994). *Perceiving events and objects*. Hillsdale, NJ: Lawrence Erlbaum.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201–211.
- Johansson, G. (1994a). Configurations in event perception. In G. Jansson, S. S. Bergström, & W. Epstein (Eds.), *Perceiving events and objects* (pp. 29–122). Hillsdale, NJ: Lawrence Erlbaum. (Original work published 1950, Doctoral dissertation, Uppsala University. Uppsala: Almqvist & Wiksell)

- Johansson, G. (1994b). Visual vector analysis and the optic sphere theory. In G. Jansson, S. S. Bergström, & W. Epstein (Eds.), *Perceiving events and objects* (pp. 270–294). Hillsdale, NJ: Lawrence Erlbaum.
- Johansson, G., & Börjesson, E. (1989). Toward a new theory of vision: Studies in wide-angle space perception. *Ecological Psychology*, *1*, 301–331.
- Kim, C.-Y., & Blake, R. (2005). Psychophysical magic: Rendering the visible “invisible”. *Trends in Cognitive Sciences*, *9*(8), 381–388. doi:10.1016/j.tics.2005.06.012
- Koch, C., & Crick, F. (2001). On the zombie within. *Nature*, *411*(6840), 893. doi:10.1038/35082161
- Koenderink, J. J. (1990). *Solid shape*. Cambridge, MA: MIT Press.
- Koenderink, J. J., & van Doorn, A. J. (1992a). Second-order optic flow. *Journal of the Optical Society of America A*, *9*(4), 530–538.
- Koenderink, J. J., & van Doorn, A. J. (1992b). Generic neighborhood operators. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *14*(6), 597–605. doi:10.1109/34.141551
- Koenderink, J. J., & van Doorn, A. J. (1997). The generic bilinear calibration-estimation problem. *International Journal of Computer Vision*, *23*(3), 217–234.
- Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2001). Ambiguity and the “mental eye” in pictorial relief. *Perception*, *30*, 431–448. doi:10.1068/p3030
- Land, M. F., & Tatler, B. W. (2009). *Looking and acting: Vision and eye movements in natural behaviour*. Oxford: Oxford University Press.
- Lappin, J. S., & Craft, W. D. (2000). Foundations of spatial vision: From retinal images to perceived shapes. *Psychological Review*, *107*(1), 6–38. doi:10.1037//0033-295X.107.1.6
- Lappin, J. S., & van de Grind, W. A. (2002). Visual forms in space-time. In L. Albertazzi (Ed.), *Unfolding perceptual continua* (pp. 119–146). Amsterdam: John Benjamins.
- Lappin, J. S., Norman, J. F., & Phillips, F. (2011). Fechner, information, and shape perception. *Attention, Perception, & Psychophysics*, *73*(8), 2353–2378. doi:10.3758/s13414-011-0197-4
- Lederman, L. M., & Hill, C. T. (2004). *Symmetry, and the beautiful universe*. Amherst, NY: Prometheus Books.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, *5*(4), 437–459. doi:10.1068/p050437
- Lee, D. N., & Reddish, P. E. (1981). Plummeting gannets: A paradigm of ecological optics. *Nature*, *293*(5830), 293–294. doi:10.1038/293293a0
- Levin, D. T., Momen, N., Drivdahl, S. B., & Simons, D. J. (2000). Change blindness blindness: The metacognitive error of overestimating change-detection ability. *Visual Cognition*, *7*(1–3), 397–412. doi:10.1080/135062800394865
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*(6657), 578–581. doi:10.1038/36846
- Mack, A., & Rock, I. 1998. *Inattention blindness*. Cambridge, MA: MIT Press.
- Markoff, J. (2011). In search of a robot more like us. *New York Times*, July 11, 2011. Retrieved December 7, 2012, from [http://www.nytimes.com/2011/07/12/science/12robot.html?pagewanted=1&r=1&sq=in search of a robot more like us&st=cse&scp=1](http://www.nytimes.com/2011/07/12/science/12robot.html?pagewanted=1&r=1&sq=in%20search%20of%20a%20robot%20more%20like%20us&st=cse&scp=1)
- Marr, D. (1982). *Vision*. San Francisco: W. H. Freeman.
- Mather, G., Verstraten, F., & Anstis, S. (Eds.). (1998). *The motion aftereffect: A modern perspective* [Electronic version, ISBN: 0585003041]. Cambridge, MA: MIT Press.
- McBeath, M. K., Shaffer, D. M., & Kaiser, M. K. (1995). How baseball outfielders determine where to run to catch fly balls. *Science*, *268*(5210), 569–573. doi:10.1126/science.7725104

- McLeod, P., & Dienes, Z. (1993). Running to catch the ball. *Nature*, *362*(6415), 23. doi:10.1038/362023a0
- McLeod, P., Reed, N., & Dienes, Z. (2001). Toward a unified fielder theory: What we do not yet know about how people run to catch a ball. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(6), 1347–1355. doi:10.1037//0096-1523.27.6.1347
- McLeod, P., Reed, N., & Dienes, Z. (2006). The generalized optic cancellation theory of catching. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(1), 139–148. doi:10.1037/0096-1523.32.1.139
- Meyer-Eppler, W. (1969). *Grundlagen und Anwendungen der Informationstheorie*. Berlin: Springer.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review*, *63*(2), 81–97. doi:10.1037/h0043158
- Mormann, F., Dubois, J., Kornblith, S., Milosavljevic, M., Cerf, M., Ison, M., Tsuchiya, et al. (2011). A category-specific response to animals in the right human amygdala. *Nature Neuroscience*, *14*(10), 1247–1249. doi:10.1038/nn.2899
- Norman, J. F., & Todd, J. T. (1998). Stereoscopic discrimination of interval and ordinal depth relations on smooth surfaces and in empty space. *Perception*, *27*(3), 257–272. doi:10.1068/p270257
- Oudejans, R. R. D., Michaels, C. F., Bakker, F. C., & Davids, K. (1999). Shedding some light on catching in the dark: Perceptual mechanisms for catching fly balls. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(2), 532–542. doi:10.1037/0096-1523.25.2.531
- Penrose, R. (1995). Mathematics of the impossible. In R. Gregory, J. Harris, P. Heard, & D. Rose (Eds.), *The artful eye* (pp. 324–334). New York: Oxford University Press.
- Purves, D., & Lotto, R. B. (2003). *Why we see what we do: An empirical theory of vision*. Sunderland, MA: Sinauer.
- Reed, E., & Jones, R. (Eds.). (1982). *Reasons for realism, selected essays of James J. Gibson*. Hillsdale, NJ: Erlbaum.
- Reed, E. S. (1988). *James J. Gibson and the psychology of perception*. New Haven, CT: Yale University Press.
- Rock, I. (1983). *The logic of perception*. Cambridge, MA: MIT Press.
- Rock, I. (Ed.). (1997a). *Indirect perception*. Cambridge, MA: MIT Press.
- Rock, I. (1997b). Orientation and form. In I. Rock (Ed.), *Indirect perception* (pp. 133–150). Cambridge, MA: MIT Press. (Original work published 1974: The perception of disoriented figures. *Scientific American*, *230*(1), 78–85)
- Sacks, O. (2010). *The mind's eye*. New York: Alfred A. Knopf.
- Sekuler, R. (1996). Motion perception: A modern view of Wertheimer's 1912 monograph. *Perception*, *25*(10), 1243–1258. doi:10.1068/p251243
- Shaffer, D. M., & McBeath, M. K. (2002). Baseball outfielders maintain a linear optical trajectory when tracking uncatchable fly balls. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 335–348. doi:10.1037//0096-1523.28.2.335
- Shannon, C. E. (1949). The mathematical theory of communication. In C. E. Shannon & W. Weaver (Eds.), *The mathematical theory of communication* (pp. 3–90). Urbana, IL: University of Illinois Press. (Original work published 1948)
- Shepard, R. N. (1990). *Mind sights: Original visual illusions, ambiguities, and other anomalies, with a commentary on the play of mind in perception and art*. New York: W. H. Freeman.
- Simoncelli, E. P. (2004). Local analysis of visual motion. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1616–1623). Cambridge, MA: MIT Press.

- Simons, D. J. (Ed.). (2000). *Change blindness and visual memory*. East Sussex, UK: Psychology Press. Also published as a special issue of *Visual Cognition*, 7(1–3), 1–412, ISSN: 1464-0716 [Online].
- Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattention blindness for dynamic events. *Perception*, 28(9), 1059–1074. doi:10.1068/p2952
- Staller, J. D. & Lappin, J. S. (1981). Visual detection of multi-letter patterns. *Journal of Experimental Psychology: Human Perception and Performance*, 7(6), 1258–1272.
- Titchener, E. B. (1896/1899). *An outline of psychology*. New York: Macmillan.
- Trommershauser, J., Kording, K., & Landy, M. S. (Eds.). (2011). *Sensory cue integration*. New York: Oxford University Press.
- Ullman, S. (1980). Against direct perception. *The Behavioral and Brain Sciences*, 3(3), 373–415.
- Wade, N. J. (1998). *A natural history of vision*. Cambridge, MA: MIT Press.
- Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, 45(4), 205–217. doi:10.1037/h0056880
- Wandell, B. A. (1995). *Foundations of vision*. Sunderland, MA: Sinauer.
- Warren, W. H. (2004). Optic flow. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1247–1259). Cambridge, MA: MIT Press.
- Warren, W. H. (2005). Direct perception: The view from here. *Philosophical Topics*, 33(1), 335–361.
- Westheimer, G. (1999). Gestalt theory reconfigured: Max Wertheimer's anticipation of recent developments in visual neuroscience. *Perception*, 28(1), 5–15. doi:10.1068/p2883
- Wiener, N. (1954). *The human use of human beings: Cybernetics and society*. Boston: Houghton Mifflin.
- Wiener, N. (1961). *Cybernetics, or control and communication in the animal and the machine*, 2nd ed. Cambridge, MA: MIT Press.
- Woodman, G. F., & Vogel, E. K. (2008). Selective storage and maintenance of an object's features in visual working memory. *Psychonomic Bulletin & Review*, 15(1), 223–229. doi:10.3758/PBR.15.1.223
- Yilmaz, E. H., & Warren, W. H. (1995). Visual control of braking: A test of the τ -dot hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, 21(5), 996–1014. doi:10.1037/0096-1523.21.5.996

