

Visual Perception

With this summary of visual system structure and function in mind, the following sections consider the end product of visual processing—that is, what we actually see. The primary visual qualities that describe visual perception are *lightness*, *brightness*, *color*, *form*, *depth*, and *motion*. These qualities are the foundation that allows us to make the associations needed to recognize objects and conditions in the world.

Lightness and brightness

A good place to begin a consideration of visual percepts is with *lightness* and *brightness*, the terms used to describe our visual experience of light and dark elicited by different light intensities. **Lightness** refers to the appearance of a surface such as a piece of paper; **brightness** refers to the appearance of a light source, such as the sun, or a lightbulb. Vision is impossible without these perceptual qualities, whereas some other qualities, such as color, are expendable (some animals have good vision generally, but little or no color vision). Like all percepts, lightness and brightness are not subject to direct measurement, and can be evaluated indirectly only by asking observers to report the appearance of one object or surface relative to that of another (**Box 3B**). The physical correlate of brightness is **luminance**, a measure of light intensity made by a photometer and expressed in units such as candelas per square meter. As will be apparent, however, the relationship between luminance and lightness/brightness is deeply puzzling.

A logical assumption would be that luminance and lightness/brightness are directly proportional, since increasing the luminance of a stimulus increases the number of photons captured by photoreceptors. Another intuition is that two objects in a scene that return the same amount of light to the eye should appear equally light or bright. It has long been known, however, that perceptions of lightness/brightness fail to meet these expectations. For example, a patch on a background of relatively low luminance appears lighter or brighter than the same patch on a background of higher luminance—a phenomenon called **simultaneous lightness/brightness contrast** (**Figure 3.13A**). This effect becomes even more dramatic when the stimulus includes more detailed information (**Figure 3.13B**).

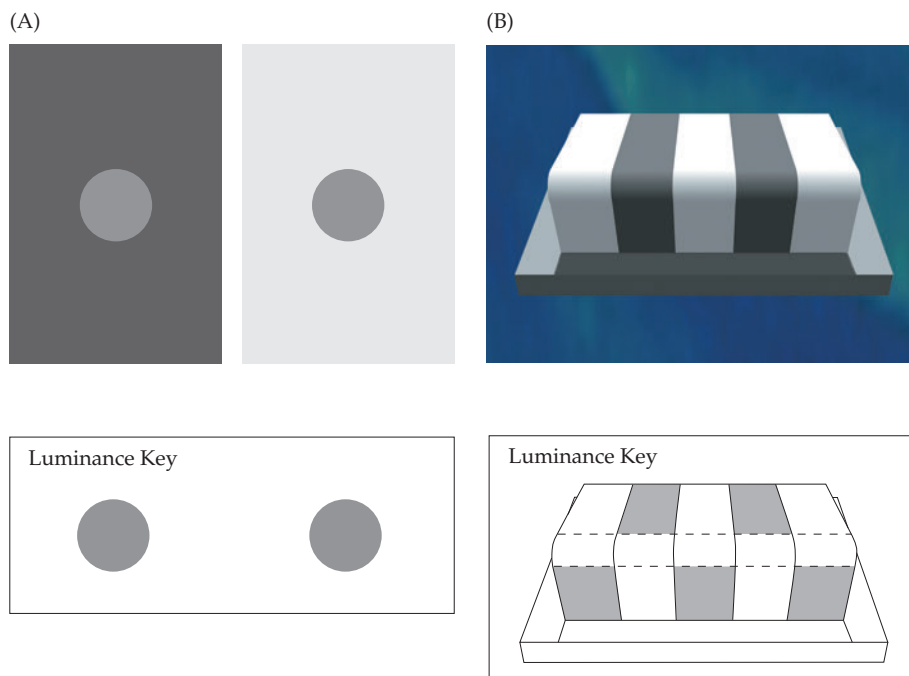
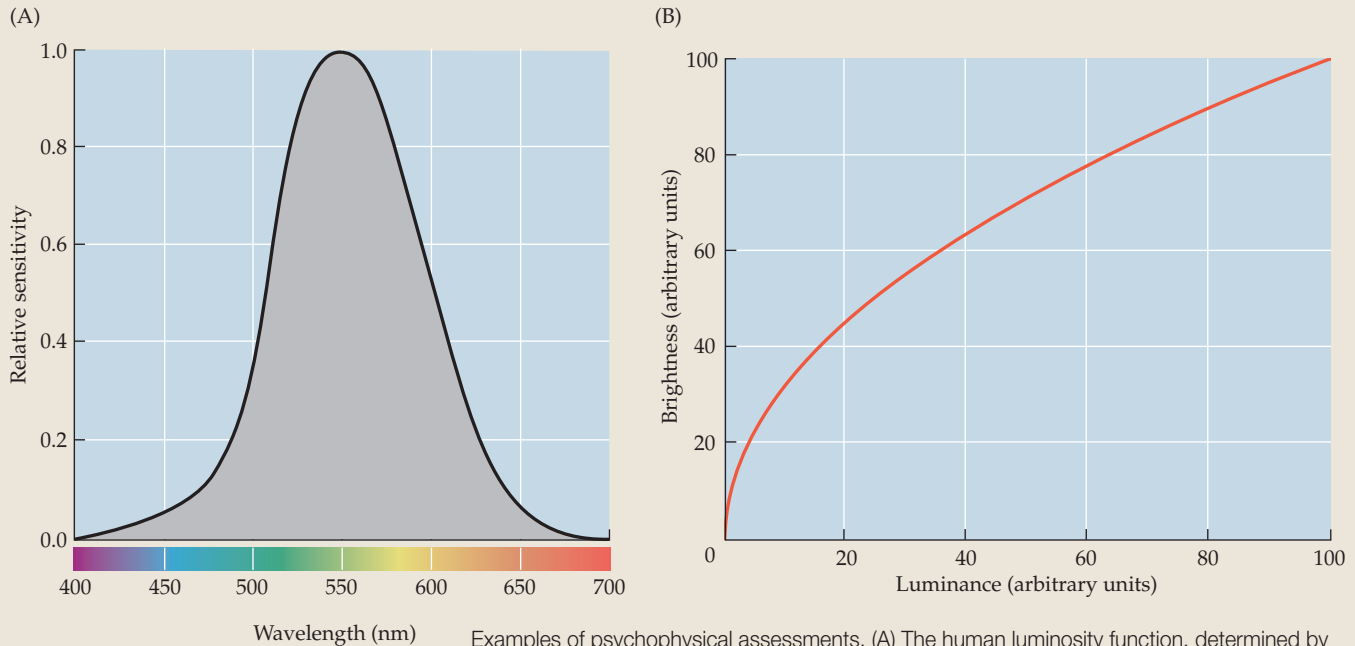


Figure 3.13 Simultaneous lightness/brightness contrast (A) In this standard presentation of the effect, the two circular patches have exactly the same luminance (see the key), but the one in the dark surround looks somewhat lighter/brighter. (B) Simultaneous lightness/brightness contrast effects can be much greater when the scene contains more information; as the key shows, the patches whose lightness/brightness appears very different again have the same luminance. Thus, the amount of light returned to the eye does not determine the lightness/brightness seen. (B from Purves and Lotto 2003.)

BOX 3B MEASURING PERCEPTION



The physical properties of stimuli can be measured with arbitrary precision. Measuring percepts, however, is quite another matter. The perceptual consequences of stimuli are subjective, and as such they can't be measured in any direct way. They can, however, be reported in terms of thresholds, least discernible differences, or other paradigms in which subjects state whether a percept is brighter or darker, larger or smaller, slower or faster than some standard of comparison. Such evaluations of perceptual responses are broadly referred to as *psychophysics*. The effort to make the analysis of percepts scientifically meaningful dates from 1860, when the German physicist and philosopher Gustav Fechner decided to pursue the connection between what he referred to as the “physical and psychological worlds” (thus the rather unfortunate word *psychophysics*).

In practice, there are only a limited number of ways to assess perception in relation to physical stimuli, although there are many permutations of the basic techniques. A conceptually straightforward but technically difficult measurement is to ascertain the least energetic stimulus that elicits a perceptual response in a particular sensory modality, such as the weakest retinal stimulation perceived as something seen by dark-adapted sub-

jects. By varying the amount of energy delivered, a *psychophysical function* can be obtained that defines the *stimulus threshold value* (Figure A). Since at threshold levels of stimulation subjects have difficulty saying whether they saw something or not, such tests are usually carried out using a *forced-choice paradigm*, in which the observer must respond on each trial. Typically, a series of trials is presented in which stimuli of different energetic levels are randomly interspersed with trials that do not present a stimulus. Because 50 percent correct responses (i.e., saying “Yes, I saw something” or “No, I saw nothing” when a stimulus was or was not present, respectively) would be the average result obtained if the subject merely guessed on each trial, a 75 percent correct-response rate is conventionally taken to be the criterion for establishing the threshold level of stimulus energy.

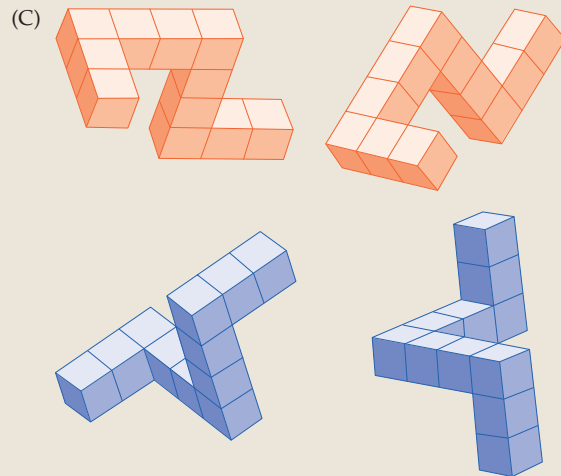
A technically easier and more generally applicable way of getting at the sensitivity of a sensory modality is to measure—at any level of stimulus intensity—how much physical change is needed to generate a perceptual change. The resulting functions, called *difference threshold functions*, have many practical implications. The *Weber-Fechner law* is a good example. The law states that the ability to notice a difference (called a test of *just-noticeable* or *equally noticeable differences*) is determined by a fixed proportion of the stimulus intensity, not an absolute difference. This proportion is referred to as the *Weber fraction*; if, for example, the Weber fraction is 1/10, then if a 1-gram increment to a 10-gram weight can just be detected, 10 grams will be the minimum detectable increment to a 100-gram weight.

What is now known about the physiology of sensory systems indicates that the proportional relationship between just-noticeable differences and stimulus magnitude expressed by the Weber-Fechner law makes good sense. Recall that because neurons can generate only

measure—at any level of stimulus intensity—how much physical change is needed to generate a perceptual change. The resulting functions, called *difference threshold functions*, have many practical implications. The *Weber-Fechner law* is a good example. The law states that the ability to notice a difference (called a test of *just-noticeable* or *equally noticeable differences*) is determined by a fixed proportion of the stimulus intensity, not an absolute difference. This proportion is referred to as the *Weber fraction*; if, for example, the Weber fraction is 1/10, then if a 1-gram increment to a 10-gram weight can just be detected, 10 grams will be the minimum detectable increment to a 100-gram weight.

What is now known about the physiology of sensory systems indicates that the proportional relationship between just-noticeable differences and stimulus magnitude expressed by the Weber-Fechner law makes good sense. Recall that because neurons can generate only

■ BOX 3B (continued)



(C) A reaction-time task. Reaching a judgment about whether the object on the left is the same as the object on the right is a function of “task difficulty”; people make this judgment more quickly for objects that are closer to the same orientation in space (top pair) than for objects that are differently oriented (bottom pair). Reaction time is used in many different paradigms, as will be apparent in later chapters.

a limited number of action potentials per second, sensory systems must continually adjust their overall range of operation to provide subjects with information about the energy levels of, say, light, where those energy levels pertinent to humans span many orders of magnitude (see Figure 3.3). The Weber fraction thus provides an approximate measure of the *gain* of a sensory system under specified conditions.

Another psychophysical approach, called *magnitude scaling*, entails order-

ing percepts along an ordinal scale that covers the full range of a perceptual quality (brightness, for instance; Figure B). The most extensive studies of this sort were carried out by psychologist Stanley Stevens, who worked on this issue from about 1950 until 1975. To take an example, Stevens asked whether a light stimulus that is made progressively more intense elicits perceptions of brightness that linearly track the physical intensity. In making such determinations, Stevens simply asked subjects to rate the relative intensities of a series of test stimuli on a number scale along which 0 represented the least intense stimulus and 100 the most intense (similar in principle to the common practice of rating pain on a scale of 1 to 10). In this manner, he determined that brightness scales as a *power function*

with an exponent of approximately 0.5 under the standard conditions he used (Figure B). The power functions exhibited in such magnitude-scaling experiments are sometimes referred to as reflecting *Stevens' law*. Rationalizing Stevens' results presents another challenge to theories seeking to explain the how visual and other sensory systems generate the percepts they do.

Finally, another staple of psychophysics entails measurements of *reaction time*. A logical assumption is that the more complex the neural processing entailed in performing a given task, the longer it will take to perform the task (Figure C). This simple paradigm is the basis of many studies in later chapters.

References

- LEZAK, M. (2004) *Neuropsychological Assessment*. Oxford: Oxford University Press.
- STEVENS, S. S. (1975) *Psychophysics*. New York: Wiley.

Many investigators have supposed that the patch on the dark background looks brighter than the patch on the light background because of a difference in the retinal output. The percepts elicited by other stimulus patterns, however, undermine the idea that simultaneous lightness/brightness contrast effects are an incidental consequence of dark versus light surrounds. In the pattern in Figure 3.14, for example, the target patches on the left are surrounded by a greater area of *higher* luminance (lighter territory) than lower luminance, and yet

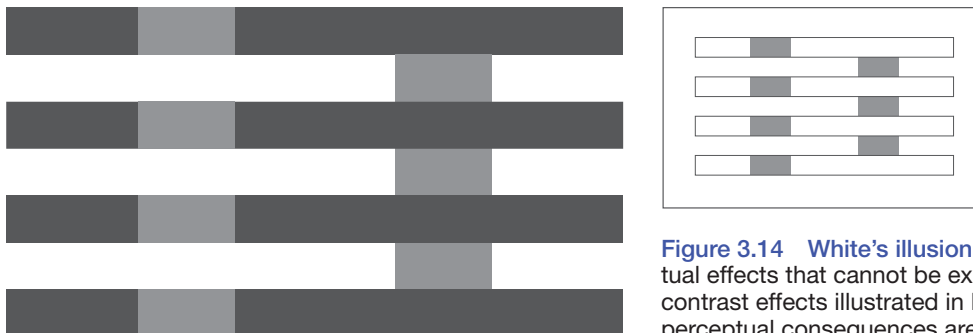
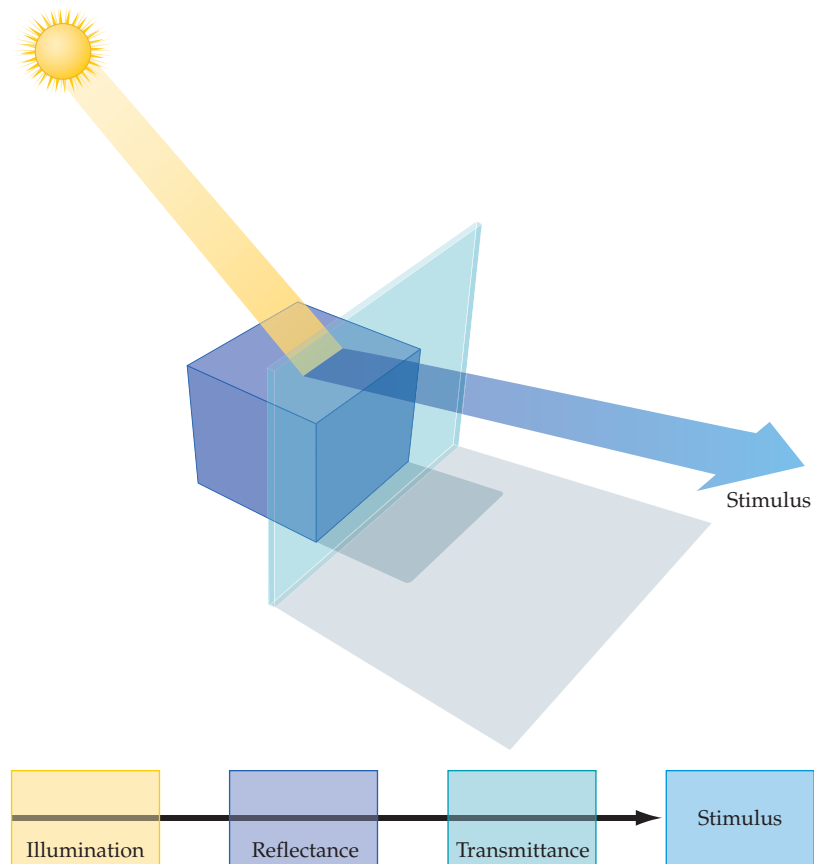


Figure 3.14 White's illusion This stimulus pattern elicits perceptual effects that cannot be explained in terms of the same local contrast effects illustrated in Figure 3.13A. The pattern and its perceptual consequences are called *White's illusion* after the psychologist who first described this stimulus more than 30 years ago. (After White 1979.)

Figure 3.15 Conflation of illumination, reflectance, and transmittance in a light stimulus An observer must parse these three factors—which the stimulus inevitably conflates—in order to respond appropriately to the pattern of luminance values in any visual stimulus.



appear brighter than the targets on the right, which are surrounded by a greater area of *lower* luminance (darker territory) than higher. Although the average luminance values of the surrounds in Figure 3.14 are effectively opposite those in the standard simultaneous lightness/brightness stimulus shown in Figure 3.13A, the brightness differences elicited are about the same in both direction and magnitude as in the standard presentation.

If the output of retinal neurons cannot account for the relative lightness/brightness values seen in response to such stimuli, what then is the explanation? A fact fundamental to answering this question is that the *sources* of luminance values are not specified in retinal images. The reason is not hard to understand. Retinal luminance is determined by three basic aspects of the physical world: the **illumination** of objects, the **reflectance** of object surfaces, and the **transmittance** of the space between the objects and the observer. As indicated in **Figure 3.15**, these factors are conflated in the retinal image; thus, many different combinations of illumination, reflectance, and transmittance can give rise to the same value of luminance. There is no logical or direct way in which the visual system can determine how these three factors are combined to generate a particular retinal luminance value. Since appropriate behavior requires responses that accord with the physical sources of a stimulus, the **inverse optics problem** presents a fundamental challenge in the evolution of animal vision (see Box 3C). For example, if we were unable to distinguish the same stimulus luminance arising from a high-reflectance surface in shadow and a low-reflectance surface in light, we would be unable to respond properly.

Evidently, our visual system has evolved to solve this problem by interpreting lightness/brightness according to past experience with the success or failure of our behavior in response to different combinations of illumination,

reflectance, and transmittance in the world. In this framework, lightness/brightness perceptions would presumably correspond to the relative frequency with which different possible combinations have proved to be the source of the same or similar stimuli in the enormous number of visual scenes witnessed during the course of evolution, as well as by individual observers during their lifetimes. This general idea harks back to the nineteenth-century vision scientist and polymath Hermann von Helmholtz, who suggested that empirical information is needed to augment what he took to be the veridical information supplied by sensory mechanisms. The more radical idea proposed by cognitive neuroscientists today is that vision and the neural connections that underlie it depend entirely on trial-and-error experience. In this conception, the lightness/brightness values seen by an observer accord with the behavioral significance of stimuli, rather than with the physical intensities of light falling on the retina. Many peculiarities of lightness/brightness, including the percepts elicited by the stimuli in Figures 3.13 and 3.14, can be explained in this way (see Box 3C).

Color

Lightness and brightness are perceptions elicited by the overall amount of light in a visual stimulus. **Color** is the perceptual category generated by the distribution of that amount of light across the visible spectrum—that is, the relative amounts of light energy at short, middle, and long wavelengths ([Figure 3.16A](#)). The experience of color comprises three perceptual qualities: (1) *hue*, the perception of the relative redness, blueness, greenness, or yellowness of a stimulus; (2) *saturation*, the degree to which the percept approaches a neutral gray (e.g., a highly unsaturated red will appear gray, although with an appreciable reddish tinge); and (3) *color brightness*, the perceptual category described in the previous section, but applied to a stimulus that elicits a discernible hue. Taken together, these three qualities describe a perceptual **color space** ([Figure 3.16B](#)). The ability to see colors evolved in humans and many other mammals because perceiving spectral differences allows an observer to distinguish object surfaces more effectively than by basing those distinctions on luminance alone.

In humans, seeing color is based on the different absorption properties of three different cone types with different photopigments (called *cone opsins*). Each cone type thus responds best to a different portion of the visible light spectrum (roughly speaking, to long, middle, and short wavelengths, respectively; see [Figure 3.16A](#)). The fact that human color vision is based on the different sensitivity of three cone types means that we humans are **trichromats**; color vision in most other mammals that have significant color vision is based on only two cone types, however, and they are thus referred to as **dichromats**.

A common disorder of human color perception arises from a genetic defect in one or more of the three cone types. The most common form of “color blindness” is deficiency of a single cone type and affects about 5 percent of males in the United States. The defective gene is located on the X chromosome, which explains the overwhelming preponderance of this problem in males. Although such individuals cannot distinguish between red and green hues (or, less commonly, between blue and yellow), most people with this problem have little practical difficulty in daily life, confirming that color perception is expendable, whereas lightness/brightness perception is not.

While successfully accounting for many aspects of color perception in the laboratory, explanations of color vision based on retinal output from the three human cone types have long been recognized to be inadequate, in much the same way that retinal output determined by luminance does not adequately explain the lightness or brightness that people see. The comparisons made by the three cone types provide only a partial account of the colors we end up seeing and therefore of how color sensations are generated. Like lightness/brightness

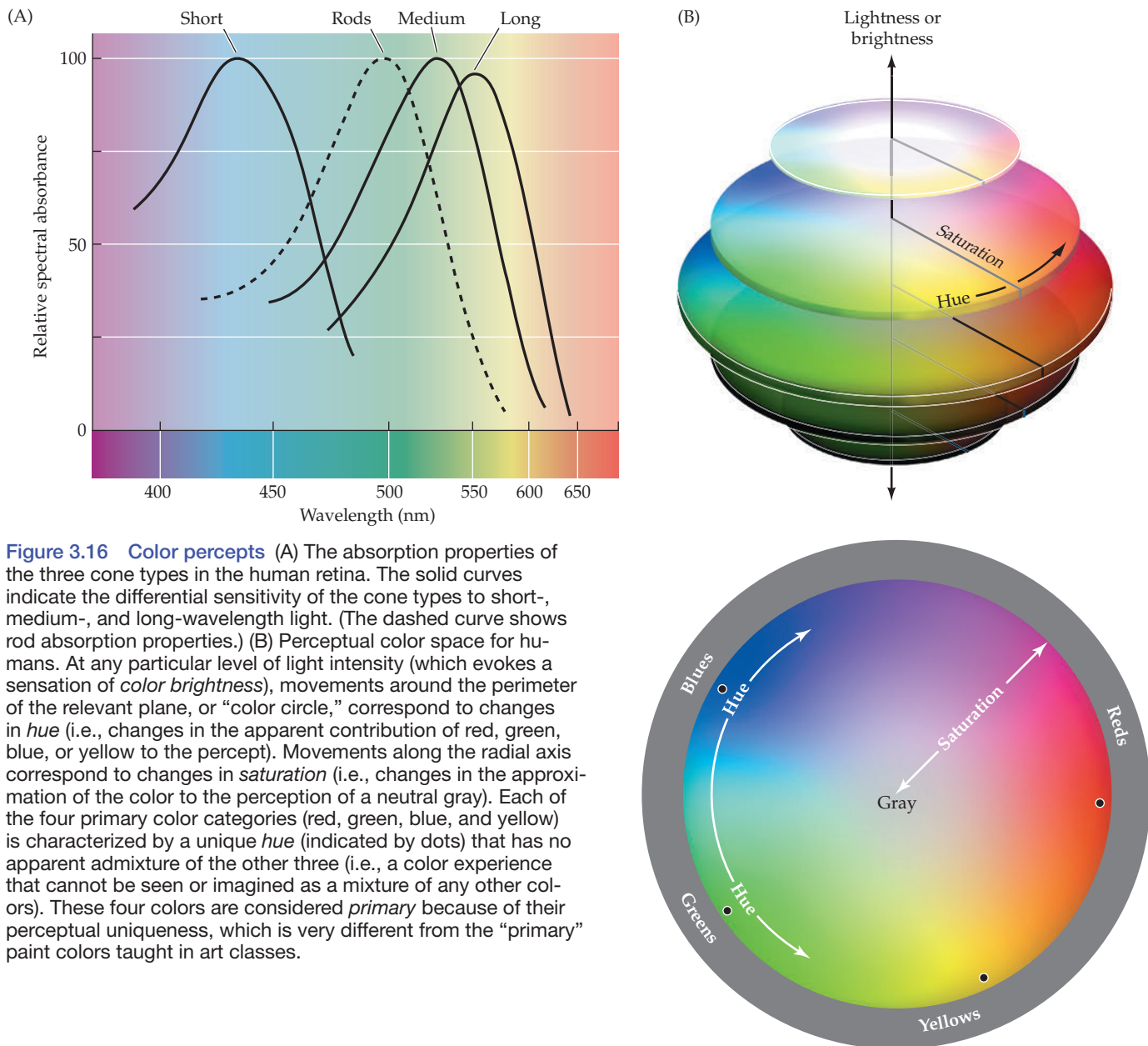


Figure 3.16 Color percepts (A) The absorption properties of the three cone types in the human retina. The solid curves indicate the differential sensitivity of the cone types to short-, medium-, and long-wavelength light. (The dashed curve shows rod absorption properties.) (B) Perceptual color space for humans. At any particular level of light intensity (which evokes a sensation of *color brightness*), movements around the perimeter of the relevant plane, or “color circle,” correspond to changes in *hue* (i.e., changes in the apparent contribution of red, green, blue, or yellow to the percept). Movements along the radial axis correspond to changes in *saturation* (i.e., changes in the approximation of the color to the perception of a neutral gray). Each of the four primary color categories (red, green, blue, and yellow) is characterized by a unique *hue* (indicated by dots) that has no apparent admixture of the other three (i.e., a color experience that cannot be seen or imagined as a mixture of any other colors). These four colors are considered *primary* because of their perceptual uniqueness, which is very different from the “primary” paint colors taught in art classes.

perceptions, the colors we see are strongly influenced by the rest of the scene. For example, a stimulus patch generating exactly the same distribution of light energy at various wavelengths can appear quite different in color depending on its surroundings—a phenomenon called **color contrast** (Figure 3.17A). Conversely, patches in a scene returning different spectra to the eye can appear to be much the same color—an effect called **color constancy** (Figure 3.17B).

Color contrast and color constancy effects present much the same problem for understanding color processing as do contextual lightness/brightness effects. Together, these phenomena have led to a debate about color percepts that has lasted more than a century. The key issue is how global information about the spectral context in scenes is integrated with local spectral information to produce color percepts. There is as yet no consensus about how central visual processing integrates local and global spectral information to produce the remarkable phenomena apparent in color perception. The answer may again

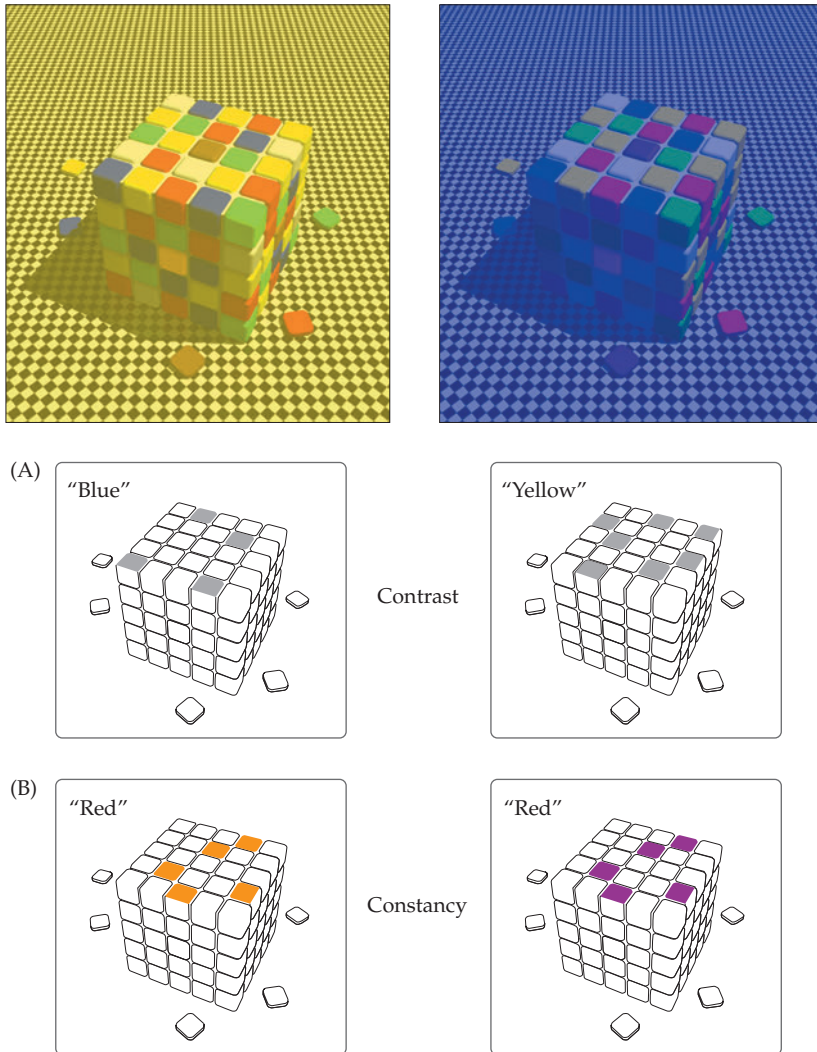


Figure 3.17 Color contrast and color constancy (A) The four blue patches on the top surface of the cube in the left panel and the seven yellow patches on the cube in the right panel are actually identical gray patches. In this demonstration of *color contrast*, these identical stimulus patches are made to appear either blue or yellow by changes in the spectral context in which they occur. (B) Patches that have very different spectra can be made to look more or less the same color (in this case, red) by contextual information—a phenomenon that demonstrates *color constancy*. (From Purves and Lotto 2011.)

be that the colors we see are determined empirically for the same reasons that lightness and brightness appear to be generated in this way—that is, to meet the challenge presented by the inherent ambiguity of light stimuli.

Information about central color processing has come from studies in non-human primates, clinical observations in patients with cortical lesions, and noninvasive brain imaging in normal subjects. This work suggests that extrastriate area V4 is especially important in color processing (see Figure 3.7). Particularly revealing have been neuropsychological and imaging studies of individuals suffering from a condition called **cerebral achromatopsia**. In effect, such patients lose the ability to see the world in color, although other aspects of vision, such as lightness, brightness, and form, remain intact. A good example described by the neurologist and essayist Oliver Sacks is a patient who, following lesions

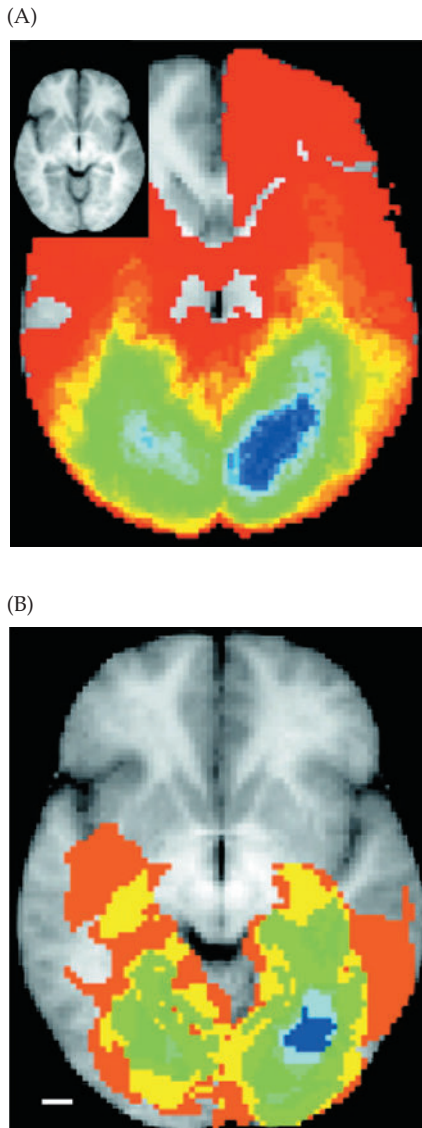


Figure 3.18 Damage to the ventral occipital cortex affects color vision A person with brain damage in this region (which includes visual cortical area V4) often suffers from an inability to perceive color (achromatopsia), despite being able to see lightness, brightness, and form more or less normally. (A) Degree of overlap in the location of lesions in a series of 46 patients with achromatopsia who also had other visual problems, such as difficulty in face recognition. Given the anatomy of the primary visual pathway (see Figure 3.5), such patients are often blind to stimuli of any sort in the contralateral visual field. (B) Degree of overlap in 11 patients in this series whose primary symptom was achromatopsia. The narrower overlap in these patients is consistent with the conclusion that the integrity of the cortex in the general vicinity of V4 is important for color vision. The inset in (A) shows the level of the horizontal sections shown. (From Bouvier and Engel 2006.)

in the region of V4, saw objects as all being “dirty” shades of gray. When asked to draw from memory, he had no difficulty reproducing relevant shapes or shading, but he was unable to appropriately color the objects he represented. A meta-analysis (see Box 1A) of a large number of cases of cerebral achromatopsia (Figure 3.18) shows that damage over an extensive region of the ventral occipital cortex that includes V4 can give rise to this condition, and that the region of injury typically affects other visual and cognitive functions. Thus, whereas V4 seems important to color vision, a number of related extrastriate areas probably participate as well in generating color percepts. Further support for the conclusion that V4 and surrounding regions of the extrastriate visual cortex are concerned with color processing comes from functional imaging studies in normal subjects, which show activation of these same regions when subjects undertake color-processing tasks.

Form

A third fundamental quality of visual perception is **form**. Perceptions of form entail simple geometrical characteristics such as the length of lines, their apparent orientation, and the angles they make as they intersect other lines, and understanding the responses to such stimuli is a first step toward understanding how complex object shapes are perceived.

A starting point in exploring how the visual system generates perceptions of form is examining how we perceive the distance between two points in a stimulus, as in the perceived length of a line, or the dimensions (size) of a simple geometrical shape. It is logical to suppose that the perception of a line drawn on a piece of paper or on a computer screen should correspond more or less directly to the length we see. But, as in the case of lightness, brightness, and color, perceptions of form do not correspond to physical reality. A well-studied example of this discrepancy is the variation in the perceived length of a line as a function of its orientation. As investigators have repeatedly shown over the last 150 years, a line oriented more or less vertically in the retinal image appears to be significantly longer than a horizontal line of exactly the same length; and the maximum length is perceived, oddly enough, when the stimulus is oriented about 30 degrees from vertical. This effect is evidently a particular manifestation of a general tendency to perceive the extent of any spatial interval differently as a function of its orientation in the retinal image.

There is a rich literature on other perceptual distortions (“geometrical illusions”) elicited by simple stimuli, showing in each case that measurements made with instruments like rulers or protractors are at odds with the corresponding percepts (Figure 3.19). These effects are similar to lightness, brightness, and

■ BOX 3C THE INVERSE PROBLEM

A fundamental problem in understanding perception in any sensory modality is the indirect relationship between human perceptions and the physical world. The challenge presented by this relationship is especially evident in vision, where it was recognized as early as the start of the eighteenth century. The problem is that images on the retina cannot, by their nature, uniquely specify the physical objects in a scene (see figure). Physical objects are of course three-dimensional, and what we see appears in three dimensions. But the retinal image is only two-dimensional. As a result of image projection, the actual size and arrangement of objects in real-world space cannot be specified by the retinal projection, and this information cannot be deduced by any logical operation on the image as such. The physical properties underlying the perceptual qualities of brightness, color, and form are, for much the same reasons, incapable of being specified by the information in the retinal image (see Figure 3.15). This quandary in vision is referred to as the *inverse optics problem*, and a key question in visual perception is how the visual system solves it.

Beginning with the work of Hermann von Helmholtz in the latter part of the nineteenth century, many vision scientists have supposed that the strategy for getting around the inverse problem is to make sensory systems depend to some degree on empirical information (i.e., information about the meaning of ambiguous stimuli based on past experience with what the same or similar stimuli

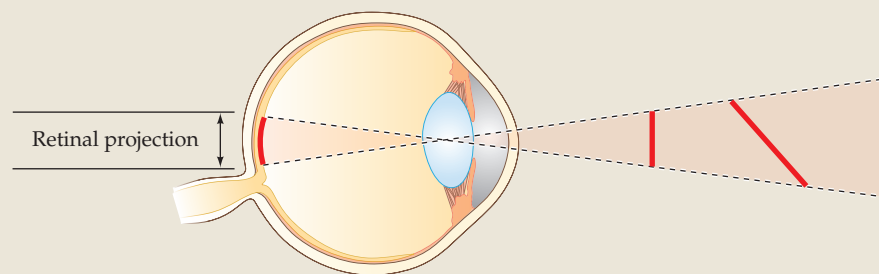
meant). In this way, Helmholtz proposed that the visual system might “disambiguate” images by an empirical boost that would help out their “interpretation.” Only recently have vision scientists and others considered the possibility that percepts might be generated *entirely* on the basis of the empirical success or failure of visual experience in the pasts of both the species and the individual. If this scenario were true, then the purpose of visual processing would essentially be to determine relative probabilities of inherently uncertain stimuli, rather than representing the “features” of the retinal projection.

As a result of these very different concepts of vision—seeing the “real world” by directly detecting its features versus seeing a brain-constructed world based on evidence determined by trial-and-error experience—there is considerable debate at present about what visual processing is actually doing and how best to interpret the wealth of information about the anatomy and physiology of visual neurons that modern neuroscience has provided over the last 50 years. One camp supposes that the way of resolving this debate in vision science is to continue obtaining more information about the properties of visual neurons and circuits, with the expectation that sooner or later it should be apparent how these properties give rise to visual perception. Another camp has tended to focus on visual percepts as such, asking whether what people actually see can tell us about the strategy of visual processing. In fact, both approaches

make sense and will eventually need to come together to complete our understanding of how percepts are generated. The central question in this prospective union is how the nuts and bolts of visual processing in the human brain instantiate the empirical evidence that is needed to solve the inverse problem.

References

- BERKELEY, G. (1709) *A New Theory of Vision*, Everyman's Library Edition. 1976.
- HELMHOLTZ, H. L. F VON (1866/1924–1925) *Helmholtz's Treatise on Physiological Optics*. (Third German Edition, Vols. I–III, 1909; translated by J. P. C. Southall) Rochester, NY: The Optical Society of America.
- HOWE, C. Q., R. B. LOTTO AND D. PURVES (2006) Empirical approaches to understanding visual perception. *J. Theor. Biol.* 341: 866–875.
- KERSTEN, D. (2000) High-level vision as statistical inference. In *The New Cognitive Neurosciences*, M. S. Gazzaniga (ed.). Cambridge, MA: MIT Press, pp. 353–363.
- MALONEY, L. T. (2003) Statistical decision theory and biological vision. In *Perception and the Physical World: Psychological and Philosophical Issues in Perception*, D. Heyer and R. Mausfeld (eds.). New York: Wiley, pp. 145–189.
- PURVES, D. AND R. B. LOTTO (2011) *Why We See What We Do Redux: A Wholly Empirical Theory of Vision*. Sunderland, MA: Sinauer.
- RAO, R. P. N., B. A. OLSHAUSEN AND M. S. LEWICKI, EDS. (2003) *Probabilistic Models of the Brain: Perception and Neural Function*. Cambridge, MA: MIT Press.



The inverse optics problem. The inherent ambiguity of retinal stimuli is illustrated here in terms of the perception of objects in space. The same linear projection on the retina can derive from an infinite number of linear objects at different distances, of different sizes, and in different spatial orientations. As a result, a retinal image cannot specify its physical source.

stimulus (Figure 3.20). The investigators took advantage of the retinotopic organization of the primary visual cortex and used fMRI to ask whether the area activated by a stimulus of a particular size corresponded to the actual size of the object in the retinal image or to its perceived size. Consistent with the idea that percepts do not derive from representations of image features,