

*Annual Review of Vision Science*

# The Certainty of Ambiguity in Visual Neural Representations

Jan W. Brascamp<sup>1</sup> and Steven K. Shevell<sup>2</sup>

<sup>1</sup>Department of Psychology and Neuroscience Program, Michigan State University, East Lansing, Michigan 48824, USA; email: brascamp@msu.edu

<sup>2</sup>Departments of Psychology and Ophthalmology & Visual Science and Institute for Mind & Biology, University of Chicago, Chicago, Illinois 60637, USA; email: shevell@uchicago.edu

Annu. Rev. Vis. Sci. 2021. 7:465–86

The *Annual Review of Vision Science* is online at [vision.annualreviews.org](http://vision.annualreviews.org)

<https://doi.org/10.1146/annurev-vision-100419-125929>

Copyright © 2021 by Annual Reviews.  
All rights reserved

## Keywords

ambiguity, visual perception, unconscious inference, bistable perception, contextual influences

## Abstract

Some images evoke bistable percepts: two different visual experiences seen in alternation while continuously viewing an unchanged stimulus. The Necker Cube and Rubin's Vase are classic examples, each of which gives alternating percepts of different shapes. Other bistable percepts are alternating colors or directions of motion. Although stimuli that result in salient bistability are rare and sometimes cleverly constructed to emphasize ambiguity, they have been influential for over 150 years, since the work of von Helmholtz, who considered them to be evidence for perceptual visual processes that interpret retinal stimuli. While bistability in natural viewing is uncommon, the main point of this review is that implicit ambiguity in visual neural representations is pervasive. Resolving ambiguity, therefore, is a fundamental and ubiquitous process of vision that routinely affects what we see, not an oddity arising from cleverly crafted images. This review focuses on the causes of widespread ambiguity, historical perspectives on it, and modern knowledge and theory about resolving it.

**ANNUAL  
REVIEWS CONNECT**

[www.annualreviews.org](http://www.annualreviews.org)

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

# 1. SEEING IS INTERPRETING: PERCEPTION REQUIRES RESOLVING AMBIGUITY

Perceptual resolution of ambiguity is fundamental for seeing. Von Helmholtz's (1867) well-known concept of unconscious inference refers to processes of interpretation, which act on initial sensory responses generated by an image on the retina and are necessary for perception. Gregory (1997, p. 1122) similarly reminds us that "retinal images are inherently ambiguous," so subsequent visual neural processes must resolve the ambiguity.

Ambiguity is not unique to seeing. Ambiguity in language was recognized by Aristotle and today remains an active research area. A modern conceptualization in that field identifies two poles of ambiguity: plurality and doubt (Ossa-Richardson 2019). Plurality refers to the multiple possible interpretations of a stimulus. Doubt is our perceived uncertainty of meaning caused by recognition of plurality. Both plurality and doubt can affect interpretation, and when discussing the ambiguity involved in vision, it is useful to distinguish these two aspects.

In the case of vision, ambiguity typically goes unnoticed because the perceptual interpretation of a scene is stable and functionally adequate; that is, the properties of the percept correspond well enough to those of the physical scene itself. The internal state of doubt, therefore, seldom arises when perceiving a natural retinal stimulus. This, however, belies implicit plurality because it suggests a direct correspondence between physical objects and what we perceive, an impression that conceals two disconnects. First, retinal images (proximal stimuli) do not convey complete, specific information about physical objects (distal stimuli) because a substantial amount of object information is either not encoded or confounded with other aspects of a scene. This creates ambiguity about the objects in view because different physical objects can generate an identical retinal image, and conversely, a single object can deliver an infinity of different images to the eye depending on, for example, viewing angle, distance, and occlusion. The second disconnect is between retinal images and the objects in our perception. The latter are richer than is determined by the objects' retinal image alone. Neural processes act on the retinal data in the process of creating perception. In other words, even if we end up with reasonably good agreement between physical objects and perception of them, this result follows from a not-so-straightforward explanation: Imperfect information at an initial encoding stage is followed by perceptual representations that are extrapolated from the incomplete retinal information.

Doubt does arise in vision in some cases, which are helpful for illustrating the above points. Many so-called optical illusions depend on creating doubt, for example, when a single stimulus evokes two different visual experiences either over time, as when viewing a Necker cube (Long & Toppino 2004, Necker 1832) (**Figure 1a, subpanel i**), or over space, as in Shepard's (1990)



**Figure 1**

(a) (i) A Necker cube. (ii, iii) The elements of the Necker cube separated into parts that include (ii) a truncated pyramid or (iii) seven planar shapes. (b) The blue edges of the left box appear longer than the magenta ones, while the blue edges of the right box appear shorter than the magenta ones. In fact, all blue edges and blue arrows are equal in length, and all magenta edges and magenta arrows are equal in length. The blue arrow at the far right reveals that magenta edges are actually longer than blue ones. Panel adapted with permission, © 2009 Lydia Maniatis; also appears in Maniatis (2017).

“Turning the Tables” and Maniatis’ (2017) variant of it (**Figure 1b**). In particular, a Necker cube stimulus gives the percept of a three-dimensional cube, and during steady viewing, one commonly sees a left-facing and then a right-facing cube in alternation, an example of bistability. These two percepts from a constant stimulus are interesting because they violate the maxim that seeing is believing (in other words, the alternating percepts create doubt). Further reflection reveals that many additional percepts are also consistent with the Necker cube’s retinal image, yet these percepts usually are not experienced. This shows that perceptual resolution of ambiguity can result in some alternating salient percepts but also, simultaneously, suppression of others. For example, the Necker cube stimulus could be interpreted—that is, seen—as having two parts, one of which is a truncated pyramid (**Figure 1a, subpanel ii**); alternatively, it could be interpreted as seven planar shapes (four trapezoids, two triangles, and a square) (**Figure 1a, subpanel iii**).

Note that using the term optical illusion for images such as these is misleading. As mentioned above, ambiguity in vision involves both a loss of information between distal and proximal stimuli (a loss that is only partly caused by optics) and subsequent visual processes (unrelated to optics) that determine one or more percepts evoked by the resulting retinal representation. Nonetheless, such illusions form excellent starting points for learning about perception and ambiguity. For instance, once one recognizes that perception of a Necker cube stimulus involves an alternation between two interpretations at the expense of countless conceivable alternatives, it is a small step to think of most everyday vision as involving a single, stable interpretation at the expense of all alternatives. Such strong suppression of alternatives eliminates doubt and thus cloaks the visual processes that resolve ambiguity.

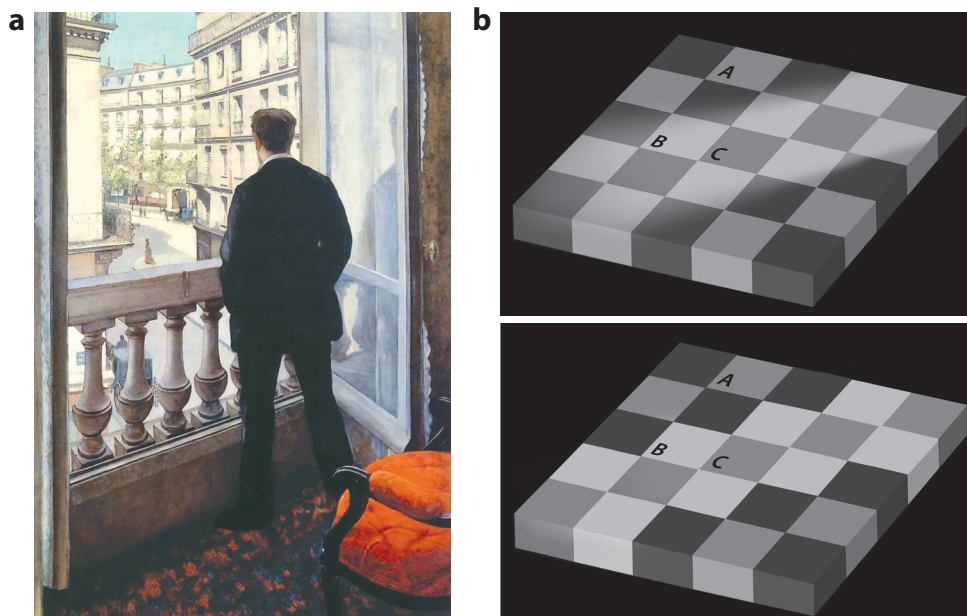
## **2. THE UNAVOIDABLE PERVASIVENESS OF AMBIGUITY: WHY EVERYDAY PERCEPTION REQUIRES AMBIGUITY RESOLUTION**

Perception of our three-dimensional world is based on two-dimensional optical projections that fall on the retina of each eye. A given retinal image can result from an infinite number of different three-dimensional objects; accordingly, retinal images long have been recognized as ambiguous with regard to three-dimensional space (discussed below in Section 3). This is not, however, the only source of ambiguity. This section reviews several causes of ambiguous neural representations.

### **2.1. The Inverse Optics Problem**

Projecting a three-dimensional object in a two-dimensional space is a straightforward application of linear algebra. Of course, some information about the three-dimensional object usually is lost in a two-dimensional projection. This causes the inverse optics problem faced by the visual system (Palmer 1999): construction of three-dimensional object percepts from two-dimensional retinal images formed by the eyes’ optics. Mathematically, achieving this inverse is underdetermined, which implies ambiguity about the objects in the physical world that generate any retinal image.

The primary information missing in two-dimensional retinal images is depth. The retinal image size of a projected object depends on its distance from the eye, so uncertainty about depth also causes ambiguity about size. While an observer’s stereoscopic pair of two-dimensional images can provide information about depth, perceived depth also depends on many additional factors, including oculomotor cues, blur, perspective, shading, and the known size of familiar objects (Banks et al. 2011, Howard 2012), none of which depend on corresponding binocular neural responses. In terms of coding, some cortical neurons respond similarly to either a stereoscopic-pair depth



**Figure 2**

(a) *Jeune homme à sa fenêtre* (*Young Man at His Window*) by Gustave Caillebotte (painted in 1875) (public domain). (b) (Top) The top-left and bottom-right corners appear to be under lower illumination because of the diagonal edges interpreted as shadow boundaries. This causes area A to appear lighter than area C although they in fact reflect the identical amount of light to the eye. (Bottom) The same image as above but with no shadow boundaries. Areas A and C appear to have the same lightness. Figure courtesy of P. Cavanagh, adapted with permission.

cue or a texture-gradient depth cue (Tsutsui et al. 2002). In general, depth can be inferred from cues in a monocular view, so depth is constructed in the process of perception, rather than being fixed by retinal disparity.

## 2.2. Objects Are Illuminated

Nearly every object that we see in the natural world is illuminated by an external light source, such as the sun (**Figure 2a**), a burning candle, or a light bulb. While some organisms generate light, for example, fireflies and some marine life (bioluminescence), most natural objects in our daily lives can only be seen with an external source of illumination. The illuminating light is reflected from the objects into the eyes. (While traffic signals and video screens are self-luminous, they are not natural objects.) Illumination causes ambiguity because the light from an object reaching the eyes depends on both the features of the object itself and the physical properties of the light source. Functional vision depends on perceiving stable object properties in spite of varying light-source properties.

An example is perceived lightness of achromatic surfaces (see Murray 2021), which varies along a continuum from black to gray to white and generally follows the relative proportion of illuminating light reflected from a surface to the eye. If the overall level of illumination falling on all surfaces is increased by a factor of 10, then retinal light stimulation from each surface in view also rises by a similar factor. The perceived lightness of the surfaces, however, is nearly unchanged.

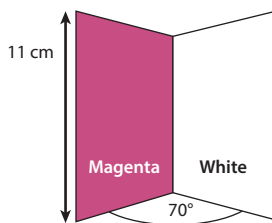
Importantly, lightness perception depends on the illumination as inferred by the observer (Adelson 1993), not necessarily on what is physically present. An observer's subjective interpretation of illuminating light affects visual mechanisms at several neural levels (even preretinally by altering pupil size; Castellotti et al. 2020). When the inferred level of illuminating light reaching a surface is reduced, but retinal illumination from that surface is fixed, this leads to greater perceived lightness (i.e., the inference that a greater proportion of light is reflected by the surface). An illustration comes from the perception of shadows. When regions of equal retinal illumination are separated by an edge with the properties of a cast shadow (areas A and C in **Figure 2b, top**), this results in the percept of two differentially illuminated regions (C being under higher illumination than A), each with a different lightness (A's lightness being greater than that of C). However, when no such edge is present, then illumination is inferred to be equal everywhere, so lightness is perceived as equal (areas A and C in **Figure 2b, bottom**). In general, retinal illumination from a surface is ambiguous in terms of surface lightness because lightness depends on both the retinal illumination and the inferred illumination of the surface.

Achromatic lightness varies along a single dimension, so resolution of lightness ambiguity is a one-dimensional problem. The range of ambiguity is much greater for the resolution of color, for which light reaching the eye depends on the full emission spectrum of the illuminant (i.e., the relative amount of energy from the light source at each wavelength in the visible spectrum), as well as the full reflectance spectrum of the object (the relative amount of energy at each wavelength reflected by the object). When viewing a given object against a dark background, the full domain of perceived colors that can be achieved by varying the object's emission spectrum is limited by the responses of the three cone photoreceptor types (trichromatic color matching) (Smith & Pokorny 2003). The domain of colors that one may see is even broader, however, when multiple objects are in view, so that contrast may evoke dark-color percepts (e.g., brown, maroon, navy blue) (Shevell 2003).

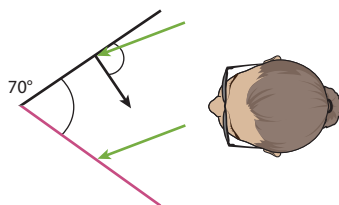
If objects took on colors according to only the light that reaches the eyes at their retinotopic location, then an object's color could shift dramatically with a change in the emission spectrum of the illuminant (Shevell & Kingdom 2008). Spectral illumination varies in natural environments (Webster 2020, Webster & Mollon 1997), yet the human visual system compensates for different illuminants reasonably well, although incompletely (Foster 2011). The result is that illuminated objects in many situations maintain a fairly constant perceived color due to neural processes that resolve the ambiguity. The ambiguity caused by the entanglement of inferred illumination and object reflectance is seldom noticed in color perception, although in certain situations, it can lead to stark differences in perceived color between individuals and, therefore, widespread doubt (for example, in the case of #thedress; Aston & Hurlbert 2017, Toscani et al. 2017).

**2.2.1. Interreflection.** Interreflection among two or more surfaces causes light reflected from one surface to illuminate another surface in the scene. This can add a further component to the factors that determine retinal illumination, as the light falling on a surface can be a combination of illumination directly from the light source and the interreflected light. In general, the visual system tends to discount illumination from interreflection, much as it does for overall illumination. Importantly, experiments show that inferred, not actual, interreflected light mediates resolution of ambiguity. When the left half of a V-shaped folded card is magenta and the right half white (Bloj et al. 1999) (**Figure 3a**), some light from the source falling on the magenta surface is reflected on to the white surface. This interreflected light is weighted toward wavelengths selectively reflected by the magenta side of the card, so the result is that light from the white surface reaching the eyes has a spectrum biased toward the longer wavelengths reflected from the magenta surface. Nonetheless, as long as the concave V shape is apparent to the observer, as in **Figure 3b**, the interreflected light

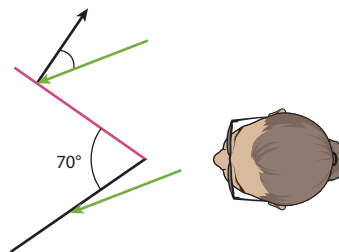
**a** Concave card



**b** Actual corner shape



**c** Apparent roof shape



**Figure 3**

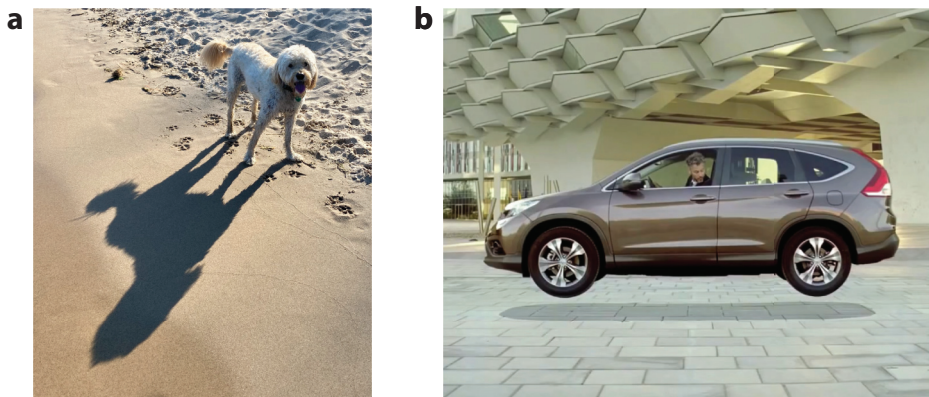
(a) A folded V-shaped concave card standing on end, with a magenta surface on the left and white surface on the right. (b) Card viewed directly. Illumination from the light source (green arrows) is reflected off each surface on to the other surface. Light reflected off the magenta surface has a spectrum biased toward longer wavelengths. (c) When the card is viewed through a pseudoscope, the physics of illumination and interreflection is still as in panel b, but the percept becomes a roof-shaped convex card, which is physically incompatible with interreflection between the two surfaces. Panel adapted with permission from Bloj et al. (1999).

that reaches the white side causes little change in the perceived color. A simple change in viewing optics, however, restores the substantial color shift predicted by the physical interreflection from the magenta surface. When the retinal image in each eye is left–right reversed using an optical device (a pseudoscope) but otherwise unchanged, the physics of light interreflection is the same, but the percept becomes a roof-shaped convex card instead of the concave V shape (**Figure 3c**). The perceived color of the white surface in this case is found to shift substantially toward magenta. The perceived convex shape is inconsistent with the physics of interreflection between the two surfaces, so the ambiguity implicit in the interreflected light is resolved as a property of surface spectral reflectance, not interreflected illumination.

**2.2.2. Specular reflection.** Depending on the smoothness of a reflecting surface, the reflected light that reaches the eye may come nearly uniformly from all parts of the surface or include a small area with substantially higher retinal illumination. The latter typically corresponds to mirror-like reflection from a smooth surface that reflects the illuminating light in a specific direction (angle of incidence from the illuminant equals angle of reflection toward the eye), unlike diffuse reflection from an unpolished surface that causes light to reflect off the surface in many different directions (as is the case for the interreflection from the V-shaped card in **Figure 3**). The ambiguity implicit in such local peaks in light reaching the eye from a particular region of a reflecting surface results from uncertainty about whether the peaks are indeed local mirror-like reflections of the illuminating light or, instead, due to locally different reflectance properties of the object's surface (D'Zmura & Lennie 1986, Yang & Maloney 2001). Disambiguating a fixed pattern of retinal illumination as a specular reflection or local change in surface reflectance can depend on the motion or perceived shape of a surface, demonstrating that the disambiguation process incorporates perceived object features beyond the retinal light pattern (Doerschner et al. 2011, Marlow & Anderson 2016).

**2.2.3. Shadows.** Shadows directly reduce the illumination reaching a surface or object, as discussed above, but also can introduce ambiguity beyond the percepts of lightness and color. In general, a shadow requires (a) a light source (illumination), (b) an object that casts a shadow, and (c) a surface on which the shadow falls. A shadow is directly affected by an object's size, position, shape, and motion and thus conveys information about properties of the object. However, because these properties are entangled with those of the illuminating light and the surface where





**Figure 4**

(a) A shadow depends on properties of the object, the illuminant, and the surface on which the shadow is cast. (b) The car appears to float above the floor because the darker floor area is interpreted as a shadow cast by the car. In fact, the darker area is a feature of the floor itself, as appreciated by viewing the figure upside down. Figure courtesy of P. Cavanagh; image credit Honda.

the shadow is cast (**Figure 4a**), inferring object properties from a shadow requires resolving ambiguity, for example, by invoking assumptions about objects or the surface (Casati & Cavanagh 2019). In addition, an object's position can be misperceived when the viewer erroneously interprets a region of reduced retinal illumination as a cast shadow instead of a local change in surface reflectance (**Figure 4b**). Finally, shadows share a property of silhouettes: Front versus back orientation is implicitly ambiguous (Troje 2017).

Many of the ambiguities caused by illumination are painstakingly depicted in art. Gustave Caillebotte's *Jeune homme à sa fenêtre* (*Young Man at His Window*), painted in 1875 (**Figure 2a**), captures several ambiguities resulting from daylight: shadows affecting surface illumination (on a Paris street, on the carpet) and suggesting a concave opening (a building's windows at top center), interreflection of sunlight reflected off the open window's glass, and specular reflection (the top of the chair's arm at its wooden tip). A painting, of course, portrays light reflected from objects and surfaces in each part of the scene and thus does not eliminate ambiguity. The artist depends on the observer to disentangle illumination and surface reflectance for each part.

### 2.3. Competing Neural Representations for the Percept at a Single Location in the Visual Field

When two very different images are presented in the same retinotopic location of each eye, for example, a face to the left eye and a house to the right (Tong et al. 1998), one sees the two images in alternation. The stimuli establish binocular rivalry, as described by Wheatstone (1838), and reveal another cause of perceptual bistability: incompatible proximal stimuli in the two eyes.

As described above, bistability does not require binocularly rivalrous stimuli. It can occur also with monocular viewing of two superimposed images (Breese 1899, O'Shea et al. 2009) or when both eyes view identical stimuli (e.g., **Figure 1a**). For present purposes, a useful distinction is between interocularly rivalrous (i.e., dichoptic) stimuli and bistable (or multistable) percepts. Perhaps surprisingly, dichoptic stimuli are common in natural viewing, even if they seldom cause a bistable percept. Consider an object that forms an image on the retina of the left eye but not the right because of an intervening occluder between the object and only the right eye (a

natural occluder can be the nose). In this situation, a particular retinotopic area has, in one eye, an image of the object but, in the other eye, an image of the occluder. Typically, the occluder is seldom, if ever, seen (Arnold 2011a,b; Shimojo & Nakayama 1990). Nonetheless, the near complete suppression of the occluder, resulting in the absence of doubt, does not diminish the basic problem created by unequal, competing neural representations for the same retinotopic area. The plurality introduced by naturally occurring stimuli is well documented, even if it is rarely experienced perceptually. In fact, “the ubiquity of diplopic [i.e., nonfusible] images away from fixation was discovered. . . in the eleventh century” (O’Shea 2011, p. 1).

### 3. ADVANCES IN UNDERSTANDING VISUAL AMBIGUITY: FROM ANTIQUITY TO NEUROSCIENCE

The sources of ambiguity discussed above have always been present in everyday life and have served as an indicator that perception amounts to more than can be explained by sense data alone. However, most people throughout the ages, including today, probably shared the intuition of a direct correspondence between perceptual objects and their counterparts in the physical world (Section 1). How have ideas developed about the distinctions between the external world, the proximal stimulus, and our perceptual world, and how have scholars viewed the associated ambiguity and resulting need to resolve it? With regard to the latter question, we focus in this section on the ambiguity that arises from experiencing a three-dimensional world through a two-dimensional projection (Section 2.1) because this type of ambiguity has attracted attention since ancient Greece.

Several schools of thought in antiquity formalized the intuitive notion that perception corresponds directly to the external world. The Greek atomists argued that objects shed three-dimensional simulacra of themselves that, upon entering the eye, communicate the properties of the original objects to the soul of the observer (Hatfield & Epstein 1979, Lindberg 1976). In Aristotelian thinking, which was influential well into the European Middle Ages, the sense organ quite literally takes on the qualities of the thing looked at (Meyering 1989). Accordingly, Aristotle wrote about the sentient object: “at the end of the process [of being acted upon by the external object] it has become like that object, and shares its quality” (quoted in Lindberg 1976, p. 58).

Even in antiquity, however, the difficulty of a three-dimensional percept rooted in two-dimensional impressions was recognized. Ptolemy (circa 100–170 AD) worked in the extramissionist tradition, which approached vision primarily from a geometrical perspective, conceiving of a cone of rays with its apex in the eye and its base on the surface of the viewed object. Each ray conveyed properties of part of the object’s surface, and the ordering of rays within the cone conveyed information about spatial layout. Ptolemy appreciated that this ordering communicates only shape in the plane orthogonal to the rays, not shape along the rays or distance from the eye. Thus, perception of those qualities would require some further process. Consistent with the extramissionist view that the rays emanate from the eye to the object, Ptolemy asserted that the eye can sense the length of the rays directly, providing an answer as to what that process could be (Hatfield 2002).

Several advances were made in approximately 1000 AD by the Arab scholar known in the West as Alhazen (circa 965–1040). He combined extramissionist geometry with an intromissionist physics in which rays enter the eye, a combination made possible by the still-accepted notion that each point on an object sends rays in all directions such that those rays that reach the pupil form a cone (Lindberg 1976, Meyering 1989, Sabra 1978). Alhazen recognized not only the ambiguity associated with three-dimensional vision but also the ambiguity involved in lightness and color perception (Section 2.2; Hatfield 2002, Howard 1996). Even more important in the present



context is that Alhazen's examination of the path between object and eye (what we would call optics) was supplemented by an equally important examination of unconscious processes that elaborate the eye's image to create perception (what we would call psychology or, more recently, neuroscience); thus, Alhazen's work represents a precursor to the theory of unconscious inference 800 years before von Helmholtz (Cavanagh 2011, Howard 1996, Sabra 1978). In Alhazen's writings, then, object, sense data, and percept are each separate, even though boundaries between them are sometimes blurry, as when Alhazen asserts that a veridical and noninverted image of the external object must be transmitted from the eye to the brain, or when he describes this transmission in terms that are quasi-optical (Hatfield & Epstein 1979, Lindberg 1976, Meyering 1989).

European scholars up to the sixteenth century followed Greek tradition in largely identifying the full process of vision with the path between object and eye (Meyering 1989). Even when translations of Alhazen's work reached the West, his followers primarily adopted his teachings on optics rather than those on postoptical processes (Cavanagh 2011, Howard 1996, Lindberg 1976, Meyering 1989). Accordingly, scholars viewed (parts of) the eye itself as the seat of vision, and although opinions differed on other matters, it was generally agreed that rays entering the eye could not intersect before causing sensation, as an inverted proximal stimulus is inconsistent with the notion that sense data and percept are effectively the same (Lindberg 1976, Meyering 1989). For example, Roger Bacon (circa 1214–1292), who was instrumental in spreading Alhazen's ideas in the West, included as a necessary condition for vision that the impression of the visible object be oriented properly (Matthews 1978). Similarly, when both Henry of Langenstein (circa 1325–1397) and Francesco Maurolico (1494–1575) independently criticized Bacon and his contemporaries, they did not dispute the necessity of a properly oriented impression. They did quite the opposite: Their criticism was that Bacon's ideas on the path of rays into the eye could not explain how this critical requirement was met (Lindberg 1976, Meyering 1989).

It was not until the seventeenth century that attention was redirected to postoptical processes and to the relationship between sense data and percept (Lindberg 1976, Meyering 1989). This was due in part to Kepler (1571–1630) and his essentially modern conception of the eye as a measuring device with a lens that projects a focused but inverted image onto the retina. Even Kepler himself was sufficiently troubled by the inversion to search in vain for a second crossing of rays before the sensitive surface (Lindberg 1976), but he eventually concluded that there must be further transmission beyond the retina that was not optical but that, instead, belonged to "the realm of the wonderful" (quoted in Hatfield & Epstein 1979, p. 373).

From that point, gradual developments often foreshadowed von Helmholtz's theory of unconscious inference of 1867. The predominant view transitioned from viewing sense data as direct knowledge to the modern notion of sense data as abstract information, containing evidence that needs interpretation. For instance, Descartes (1596–1650) likened the signs through which nature makes us feel the sensation of light to words that make conceivable the things that they represent but that in no way resemble them (Meyering 1989), quite in contrast to the Aristotelian view discussed above. In this period, considerable theorizing on perception centered on the problem of seeing three-dimensional space. Berkeley (1685–1753), for instance, echoed Ptolemy by stating, "Distance. . . immediately cannot be seen. . . . Being a line directed endwise to the eye it projects only one point [which] remains invariably the same, whether the distance be longer or shorter" (Berkeley 1709, p. 1). He then argued, preceding von Helmholtz in emphasizing the role of experience, that depth perception needs to be learned based on repeated co-occurrence of visual ideas with tactual ones (Hatfield & Epstein 1979).

With von Helmholtz, this gradual development culminated in a conception of perception that embraces the ambiguity of sensory data and that emphasizes the need for learned psychological processes that work toward informed but tentative perceptual interpretations. Von Helmholtz's

theory incorporated the then-recent idea of specific sense energies, according to which stimulation of a given nerve has specific effects. Arguably a distant descendant of Descartes's idea of signs (and an ancestor of the modern principle of univariance; Rushton 1972), this theory implies that a nerve's effects do not relate directly to the external stimulus: A given effect can result from various external stimuli that may excite the nerve (e.g., seeing light when pressing one's eyeball), and a given external stimulus can have various effects depending on which nerves it manages to excite (e.g., literally feeling a low bass tone). Von Helmholtz concluded that relating nerve signals back to plausible external stimuli requires a fast and unconscious process of interpretation, and he posited a prominent role for learning and development in shaping this process (Meyering 1989).

It is noteworthy that von Helmholtz's ideas did not immediately find broad acceptance. Her- ington, for instance, held a contrasting position that emphasized innate mechanisms and a direct link between retinal sensations and perception (see Boring 1942, Meyering 1989). More recently, the Gestaltists opposed von Helmholtz's views (Hatfield 2002, Hochberg 1981, Meyering 1989). Nevertheless, key aspects of von Helmholtz's theory are visible in present-day theories of vision and of perception more generally.

Modern theories that may be considered part of von Helmholtz's legacy include predictive coding theories (Rao & Ballard 1999), Bayesian inference theories (Aitchison & Lengyel 2017, Lee & Mumford 2003), and adaptive resonance theory (Carpenter & Grossberg 2003). These theories all share the central feature that the neural process of perception is viewed fundamentally as a process that matches incoming sensory data to existing knowledge of the world. According to several such theories, this matching process takes place via a combination of feedforward and feedback signals throughout a hierarchy of cortical processing stages, in which each successive stage encodes more abstract or more general aspects about the perceptual world than the stage before it (Friston 2005, Mumford 1992, Penny 2012, Rao & Ballard 1999). The fundamental principle is that each stage uses the feedforward signal it receives from the prior stage to inform a model of the cause of that signal. This model, in turn, forms the basis of a feedback signal that encodes a prediction of what the prior stage may be receiving as its input. Then, the feedforward signal sent from the prior stage encodes the discrepancy between this prediction and the actual input. This process continues as the system iteratively adjusts its models (across all stages) to converge on a good match between feedforward signals and feedback predictions. There are clear traces of von Helmholtz's ideas in these modern conceptualizations. First, the central role of existing knowledge fits well with von Helmholtz's emphasis on prior experience. Second, the process of iteratively adjusting an internal model of the world to fit the available data can be seen as one way of implementing von Helmholtz's inference process, and it is a small step to think of the product of this process—an internal model that fits the data well—as the computational counterpart of perception.

## 4. PERCEPTUAL RESOLUTION OF AMBIGUITY FROM CONTEXT

An emphasis above is on the idea that local retinal signals are insufficient for reliably inferring their distal source. What other information, then, provides the context that is combined with those signals on the way to perception?

### 4.1. Context from a Different Time

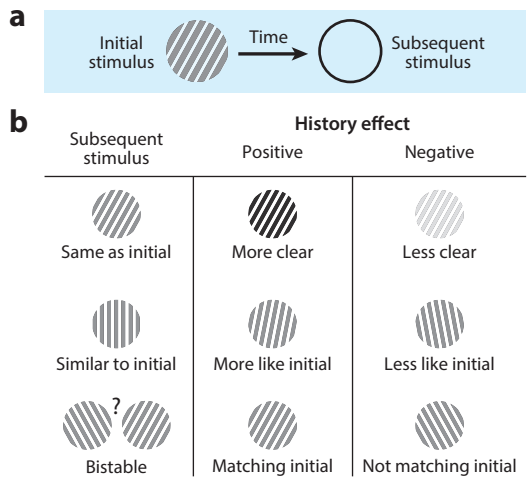
Visual perception depends on prior history. Arguably the best-known example of this is visual adaptation: Prior exposure to a stimulus with certain features (e.g., a grating tilted 10° clockwise from vertical) can cause subsequent stimuli to be perceived as more dissimilar to that initial

stimulus than they actually are (e.g., a vertical grating might appear to have a 4-degree anticlockwise tilt) (Parker 1972). Similarly, such prior exposure often causes the same stimulus, upon repeated presentation, to become more difficult to detect. In bistable perception, a similar repulsive phenomenon is observed: Following exposure to a nonbistable stimulus that induces a given interpretation, an observer is less likely to experience a similar interpretation when faced with a bistable stimulus (Chopin & Mamassian 2012, Kanai & Verstraten 2005, Long et al. 1992, Pastukhov et al. 2013b, Pearson & Clifford 2005a, Wolfe 1984).

Interestingly, when the prior stimulus is, itself, also bistable, an attractive effect often occurs instead: Prior perception of a given interpretation usually increases the probability of an observer experiencing that same interpretation again (Brascamp et al. 2008, Kanai & Verstraten 2005, Leopold et al. 2002, Pastukhov & Braun 2008, Pearson & Clifford 2005a). Outside of the realm of bistable stimuli, attractive effects occur also in some conditions (e.g., improved detection or attractive perceptual shifts) (Chopin & Mamassian 2012, Fischer & Whitney 2014, Fritsche et al. 2017, Maus et al. 2013, Oruç & Barton 2010, Tanaka & Sagi 1998).

The existence of both positive and negative forms of history dependence in perception (for a schematic overview, see **Figure 5**) suggests that the visual system is organized to meet multiple competing demands in the way in which it incorporates prior information in its disambiguation of present input. To better understand what these demands might be, consider the differences between positive and negative forms of history dependence in terms of the stimuli and tasks that give rise to them.

The first difference relates to what might be called stimulus strength. As mentioned above, for bistable stimuli, prior stimulation has a repulsive effect, while prior perception typically has an attractive effect. However, there are instructive exceptions. Brief and/or faint nonbistable stimuli can facilitate subsequent perception of the corresponding percept during bistability (Brascamp



**Figure 5**

Schematic depiction of positive and negative history effects in perception. (a) Exposure to an initial stimulus is followed by a subsequent stimulus. (b) The prior exposure to an initial stimulus can have the positive effect (*center column*) of making that same stimulus more easily detectable or clear when presented again (*top row*); of making subsequent stimuli appear more like the initial stimulus than they actually are (*center row*); or, in cases where the subsequent stimulus is bistable, of prompting perceptual dominance of the interpretation that matches the prior stimulus (*bottom row*). However, prior exposure to an initial stimulus can also have negative effects (*right column*), which are the opposite of the positive effects.

et al. 2007, Kanai & Verstraten 2005, Long & Toppino 1994, Long et al. 1992). Conversely, extensive perception of the same interpretation of a bistable stimulus can reduce the probability of experiencing that interpretation yet again (Brascamp et al. 2009, Noest et al. 2007, Pastukhov & Braun 2008). A similar rule appears to apply outside the realm of bistability: Attractive effects on, for example, detection are often limited to cases where prior stimulation was brief and/or faint (Fischer & Whitney 2014, Huber & O'Reilly 2003, Oruç & Barton 2010, Tanaka & Sagi 1998).

There is also evidence that attractive and repulsive effects tend to be associated with different processing stages. For instance, attractive effects on orientation judgments have less retinal specificity than repulsive effects (Fischer & Whitney 2014, Fritsche et al. 2017), consistent with the former being associated with a higher processing stage. In addition, in at least one case of bistable perception, attractive effects carry over best between objects that match in terms of shape (a property encoded at moderately advanced processing stages) (Pastukhov et al. 2013a), whereas repulsive effects do not (Pastukhov et al. 2013b). Also consistent with this division between lower-level repulsion and higher-level attraction is the finding that repulsive effects that result from previewing a stimulus can become attractive effects if the participant imagines the stimulus instead (DeBruine et al. 2010, Pearson et al. 2008; although see Winawer et al. 2010). Computational modeling of attractive history effects also suggests that they arise uniquely from processing stages where alternative representations compete, whereas repulsive effects primarily originate at levels that feed into those stages (Noest et al. 2007). Finally, functional imaging results show that stimulus repetition in bistable perception is associated with altered early visual responses, whereas percept repetition is reflected in higher visual and integration areas (Schwiedrzik et al. 2014).

## 4.2. Context from a Different Place

Just as sensory data are structured over time, they are also structured spatially, and these regularities can help disambiguate local data. Spatial dependencies in perception show marked similarities to temporal ones (Schwartz et al. 2007); again, both attractive and repulsive effects are observed (Paffen et al. 2006, Polat & Sagi 1993, Tadin et al. 2003, Xing & Heeger 2001).

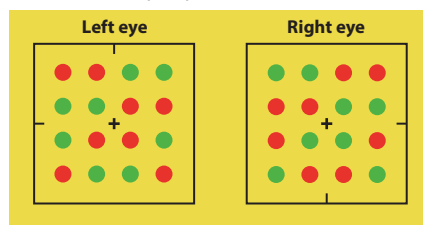
For perception of bistable stimuli, spatial interactions can be readily observed in displays such as that in **Figure 6**. Showing several copies of the same stimulus side by side (Lorenz and Tinbergen's hawk and goose; Schleidt et al. 2011) tends to cause correlated perceptual alternations across the display, rather than multiple independent perceptual sequences in parallel (Attneave 1968, Freeman & Driver 2006, Grossmann & Dobbins 2003, Klink et al. 2009, Ramachandran & Anstis



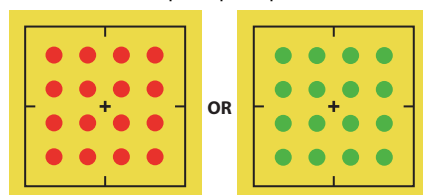
**Figure 6**

When multiple similar bistable stimuli are shown side by side, perception is often correlated across stimuli, consistent with the visual system's use of spatial context for perceptual disambiguation. In this example the observer may perceive either a flock of waterfowl lifting off or four raptors in a hunting stoop but rarely a group made up of raptors and waterfowl going in opposite directions.

### Dichoptic patchwork stimuli



### Frequent percepts



**Figure 7**

(*Top*) Rivalrous chromatic objects. In each eye, eight objects have one chromaticity, and the other eight have another chromaticity (patchwork rivalry). Chromatic rivalry is maintained at all 16 retinotopic locations. (*Bottom*) The resulting percept often consists of all 16 objects appearing the same color, sometimes all red and sometimes all green.

1983). This suggests that disambiguation in this case involves integration of information over relatively large visual distances.

When multiple binocularly rivalrous stimuli (Section 2.3) occupy different parts of the visual field simultaneously, each resulting in the same pair of bistable percepts, they too can undergo correlated perceptual alternations (Shevell 2019). **Figure 7** shows an example with 16 separate chromatically rivalrous objects, each with the same pair of rivalrous chromaticities (so each object sometimes appears red and sometimes green). The fused percept is often all objects of the same color, either all red or all green (Kovács et al. 1996, Slezak & Shevell 2018). Two extensions of standard rivalry point to a binocularly driven neural mechanism of disambiguation. First, note that, in **Figure 7 (top)**, half of the discs in each eye are red, and the other half are green, with stimuli in the opposite eye arranged for chromatic rivalry at every retinotopic location (so-called patchwork rivalry; see Kovács et al. 1996). In this case, monocular eye dominance would not lead to seeing all 16 objects with the same color. Second, the stimuli in each eye can be swapped between the two eyes approximately eight times per second (chromatic interocular switch rivalry) (Christiansen et al. 2017). Despite rapid swapping between the eyes, the percept often is sustained, with the 16 objects all appearing red or green (Slezak & Shevell 2018) (**Figure 7, bottom**). This is an attractive (positive) form of spatial correlation (**Figure 5**).

A repulsive (negative) form of spatial interaction also is found with rivalry in that two bistable objects in view may appear maximally different from each other, rather than alike. Consider two objects in view with rivalry in both color and grating orientation (Peiso & Shevell 2020), thus allowing four possible percepts: (*a*) Both objects are the identical color and orientation (attractive spatial correlation), (*b*) the two objects mismatch in both color and orientation (repulsive spatial interaction), (*c*) the objects match in color but not orientation, and (*d*) the objects match in orientation but not color. All four percepts sometimes occur, but the most frequent one, seen well above chance, is the two objects being mismatched in both color and orientation, thus enhancing their perceived dissimilarity, as would be advantageous for discriminating them from each other.

### 4.3. Context from Other Senses

The use of contextual information for interpreting sensory data is not limited to information from within the same sensory system. A striking example is the McGurk effect: The speech that we hear a person utter can change depending on how we see the person's mouth move (McGurk & MacDonald 1976). Influences in the opposite direction, from auditory information to visual perception, also occur, as when a visual flash is presented simultaneously with a rapid volley of auditory beeps. This can result in the visual perception of multiple flashes, matching the number of beeps (Shams et al. 2000). In general, which sense will most strongly drive the perception of a polysensory stimulus depends on the reliability of the signal in each sense, with the sense with the highest reliability factoring most heavily into the percept (Alais & Burr 2004, Ernst & Banks 2002, Ernst & Bühlhoff 2004). This same principle also applies to the combination of different cues within a single modality (Atkins et al. 2001, Hillis et al. 2004, Young et al. 1993).

### 4.4. Context from Prior Knowledge

In one sense, the category of context from prior knowledge covers much of what is discussed above: The very organization of perceptual systems reflects the structure of the data that they process (Barlow 1961) and, in that sense, reflects prior knowledge. Similarly, general principles of perceptual organization, such as the Gestalt laws, can be interpreted simply as appropriate use of knowledge about statistical regularities of the world (Brunswik & Kamiya 1953, Elder & Goldberg 2002, Geisler et al. 2001).

However, perception is also dependent on more specific knowledge about objects and scenes. One example comes from the perception of Necker cubes, discussed above in relation to **Figure 1a**. The two percepts that we do typically perceive are those that correspond to frequently encountered objects, while other interpretations, including those illustrated in **Figure 1a**, have little a priori plausibility. Another example is the well-known hollow mask illusion (Gregory 1980): When shown the inside of a mask, with its nose pointing away from the observer, there is a strong tendency to perceive it as the outside of a mask, with its nose pointing toward the observer. This perception reflects a lifetime of experience that faces are convex.

Whereas these examples demonstrate that perceptual principles often make sense given known regularities in our environment, there is also direct evidence that experience with those regularities can alter those principles. For instance, perception of object shape incorporates the assumption that illumination usually comes from above (from the sun or overhead lighting), which means that brighter parts of objects are usually perceived as facing upward, while darker parts are perceived as facing down (Kleffner & Ramachandran 1992). However, this assumption is malleable: Exposing observers to scenes in which light consistently comes from a different direction changes the visual system's expectations and, therefore, the perception of object shape (Adams et al. 2004).

## 5. RESOLUTION OF AMBIGUOUS REPRESENTATIONS FOR HOLISTIC OBJECTS OR FOR THE OBJECTS' FEATURES?

The result of ambiguity resolution is typically the perception of objects and other coherent scene elements, each of which has a complete set of features (e.g., size, location, color, orientation, direction of motion). A fundamental question is whether perceptual resolution of ambiguity occurs for such coherent units or, instead, whether particular features can be disambiguated individually.

For the case of binocular rivalry (Section 2.3), this question has received substantial attention, leading to the conclusion that neural competition can occur at several distinct levels of the visual



system (Blake & Logothetis 2002, Kim et al. 2020, Lee & Blake 1999, Silver & Logothetis 2007, Zhang et al. 2014; see also Section 4.1). In some cases, perceptual dominance is driven entirely by signals originating from one eye at a time (eye dominance), which implies that neural competition results in a holistic percept based on the dominant eye's retinal image. Such percepts, however, are not evidence of competition at the level of coherent units because eye dominance leads to the same percept regardless of whether dominance is at the level of a coherent object or, instead, at the level of all of the features composing the object. The phenomenon of eye dominance, therefore, cannot address this question.

When a coherent global object or scene is presented in a mosaic fashion during rivalry, such that separate spatial pieces of the object are shown to different eyes, observers may perceive the coherent global object, implying that a different eye dominates perception in separate retinotopic zones (Diaz-Caneja 1928; for a translation, see Alais et al. 2000). This reveals that a holistic percept may drive the localized retinotopic zones where one eye or the other is dominant but does not answer the broader question of whether visual competition within each zone is at the holistic object level or at the feature level.

An alternative to eye-based rivalry is resolution of conflict between neural representations of different stimuli, regardless of their eye of origin (stimulus rivalry) (Bartels & Logothetis 2010, Brascamp et al. 2020, Logothetis et al. 1996, Pearson & Clifford 2005b, Stuit et al. 2011). Stimulus rivalry can account naturally for some striking examples of dichoptic, rivalrous stimuli that lead to percepts that integrate some signals from each eye, even within a single retinotopic zone, implying that ambiguity resolution can take place at the level of individual features. For instance, one feature, such as the color of a percept, can alternate over time due to rivalrous chromaticities in the two eyes, while at the same time, the two monocularly presented forms are integrated to give stereoscopic depth perception (Treisman 1962). A similar result is found with rivalrous dichoptic orientations or spatial frequencies instead of rivalrous color (Andrews & Holmes 2011). Furthermore, color competition can accompany stereoscopic depth perception even when identical (nonrivalrous) chromatic stimuli are presented to the two eyes if rivalry between different hues is induced by chromatic contrast (Hong & Shevell 2008), showing that stereo depth perception can accompany the bistability of color that originates at a neural level beyond encoding rivalrous stimuli.

Conversely, form rivalry can exist while color perception remains stable. A classic report using two postage stamps with different colors and forms, one stamp presented to each eye, revealed that the seen forms alternate in perception, while a perceived hue, determined by both stamps' chromaticities, is stable (Creed 1935). A more modern study with dichoptically presented gratings that are rivalrous in both orientation and color revealed conditions resulting in a percept with the orientation presented to one eye and the color presented to the other (Holmes et al. 2006). Furthermore, with dichoptic orthogonal chromatic gratings that are equiluminant (for instance, a magenta/gray horizontal grating in one eye and a green/gray vertical grating in the other), the orientation is perceived to alternate, while the colors from both eyes are seen simultaneously in separate bars of the grating (e.g., a magenta/green grating alternating in orientation between horizontal and vertical) (Hong & Shevell 2009). A similar result occurs for rivalrous directions of motion (clockwise or counter-clockwise) generated by dichoptic rotating windmills of a different color in each eye: The percept is often a windmill rotating in one eye's direction but with some vanes at each of the colors presented to the two eyes (Maloney et al. 2013). In sum, the two eyes' competing representations of a particular feature can maintain their distinct neural representations for later binocular integration despite complete eye dominance for another feature, providing particularly clear evidence that, at least in the case of binocular rivalry, ambiguity resolution can take place at the level of individual features.

## 6. OPEN QUESTIONS

### 6.1. How Is Disambiguation Related to Attention and Awareness?

Almost inevitably, one thinks of the disambiguating processes discussed in this review as leading to perceived interpretations, thus shaping what we consciously see. Does this imply a special relationship between disambiguation and awareness? Is input still disambiguated when it does not produce a percept, for instance, due to inattention? Existing work addresses these questions, although a consistent picture is yet to emerge. For instance, some stimuli that normally give rise to bistable perception apparently fail to elicit bistability's characteristic neural signature when attention is directed elsewhere (Brascamp & Blake 2012, Zhang et al. 2011). This seems not to apply to other such stimuli, however (Dieter et al. 2016), raising the possibility of bistable processing of nonperceived stimuli (Zou et al. 2016). Another relevant body of work concerns surface segmentation. As an early step on the way from retinal data to interpreted scene, surface segmentation transforms retinal neural signals into representations of regions corresponding to distinct surfaces and also establishes those surfaces' depth relations. Segmentation is strongly affected during inattention (Mack et al. 1992, Poort et al. 2012, Roelfsema 2006) or lack of awareness of the distinct regions (Lamme et al. 1998). Nonetheless, some segmentation does occur preattentively and, indeed, seems to be required to define the units to which attention can be allocated (Driver et al. 2001, Scholte et al. 2006, Von der Heydt 2004). In sum, work on these topics hints that disambiguation, attention, and awareness are intertwined, but it cannot yet be said with certainty whether this is generally the case.

### 6.2. Is There Optimal Ambiguity for Visual Neural Representations?

Perceptual interpretation limits the possible percepts that we see, even when it does not determine a unique percept. Depending on the degree of suppression, there may be many different percepts that reach consciousness (i.e., multistability) or only one, which raises the question of whether there is an optimal degree of ambiguity. Too little flexibility undermines recognition because different objects can deliver an identical retinal image (Section 2). Too much flexibility may result in failure to generate percepts with sufficient specificity to be functionally advantageous. Is there an optimal level of ambiguity for seeing? This question is related to issues raised in the context of Bayesian theories of perception, according to which representations need to encode not just a singular judgment as to what is seen, but also a measure of the reliability of this judgment and the plausibility of alternatives (Ernst & Banks 2002, Fiser et al. 2010, Jazayeri & Movshon 2006).

Studies of language provide a perspective. Linguists have long grappled with ambiguity in phrases having multiple meanings. (Consider this advice during the COVID-19 pandemic about leaving one's house: "No mask is better than staying at home.") Perhaps surprisingly, some degree of ambiguity is considered beneficial in language:

First, where context is informative about meaning, unambiguous language is partly redundant with the context and therefore inefficient; and second, ambiguity allows the reuse of words and sounds which are more easily produced or understood. (Piantadosi et al. 2012, p. 281)

Ambiguity, therefore, can enhance the efficiency of communication when the role of context is included. In the no-mask example above, context may lead to the intended meaning that there exists no mask that protects someone as well as staying at home, rather than that leaving home without a mask is preferable to not going out at all. Given the importance of context for perceptual resolution of ambiguity in vision (Section 4), insights from language may be instructive.

Assuming that both speakers and listeners aim to minimize effort (Zipf 1949), speakers should prefer the shortest possible utterance (maximal ambiguity) because it takes the least exertion, while listeners should want a clear and necessarily longer statement with a unique meaning that requires no effort to work out the speaker's message (minimal ambiguity). A compromise between speaker and listener is a degree of ambiguity that improves efficiency. In vision, the neural representation of a proximal stimulus is the speaker, and the percept (inferred meaning) is the listener. An advantage of two or more alternating visual percepts from a constant stimulus (bistability), instead of a single percept, is that viewers experience more than one percept at the same retinal location over a short period of time, which may have more functional value than a single stable percept that may mask useful information. At the same time, however, this introduces the cost (or perhaps benefit) of instilling doubt.

## SUMMARY POINTS

1. Perceptual resolution of ambiguity is pervasive in vision because perception requires constructing three-dimensional object percepts from two-dimensional retinal images formed by the eyes' optics (the inverse optics problem) and also because properties of objects in a scene must be disentangled from the light illuminating them.
2. In the history of vision research, the idea of pervasive ambiguity and its resolution is famously associated with von Helmholtz's theory of unconscious inference, but a gradual movement toward that theory can be discerned starting in the seventeenth century with the recognition of a separation between retinal image and percept. Elements of this theory are visible even earlier, in the work of Alhazen more than 800 years before von Helmholtz.
3. Although plurality (the existence of multiple interpretations) characterizes the retinal impression of many everyday scenes, perceptual resolution is typically so effective and fast that it precludes doubt (conscious recognition of the existence of alternative interpretations) except in the case of various illusions that lay bare this plurality and provide ways to study the processes of ambiguity resolution (notably, stimuli causing perceptual bistability).
4. Studies of binocular rivalry, which causes a form of visual neural competition that can lead to perceptual bistability, provide insight into the processes of ambiguity resolution by, for instance, providing evidence that the visual system can, under some circumstances, disambiguate various features of a scene independently.
5. The visual system resolves ambiguity using various sources of contextual information, thereby exploiting redundancies in sensory data across space, across time, or across different senses.
6. The idea that ambiguity resolution is central to seeing is fundamental to several modern-day theories of perception, which frame the neural process of perception as one that matches inconclusive sensory data to existing knowledge of the world.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant EY-026618 to S.K.S.

## LITERATURE CITED

- Adams WJ, Graf EW, Ernst MO. 2004. Experience can change the “light-from-above” prior. *Nat. Neurosci.* 7(10):1057–58
- Adelson EH. 1993. Perceptual organization and the judgment of brightness. *Science* 262:2042–44
- Aitchison L, Lengyel M. 2017. With or without you: predictive coding and Bayesian inference in the brain. *Curr. Opin. Neurobiol.* 46:219–27
- Alais D, Burr D. 2004. The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14(3):257–62
- Alais D, O’Shea RP, Mesana-Alais C, Wilson IG. 2000. On binocular alternation. *Perception* 29:1437–45
- Andrews TJ, Holmes D. 2011. Stereoscopic depth perception during binocular rivalry. *Front. Hum. Neurosci.* 5:99
- Arnold DH. 2011a. I agree: binocular rivalry stimuli are common but rivalry is not. *Front. Hum. Neurosci.* 5:157
- Arnold DH. 2011b. Why is binocular rivalry uncommon? Discrepant monocular images in the real world. *Front. Hum. Neurosci.* 5:116
- Aston S, Hurlbert A. 2017. What #theDress reveals about the role of illumination priors in color perception and color constancy. *J. Vis.* 17(9):4
- Atkins J, Fiser J, Jacobs R. 2001. Experience-dependent visual cue integration based on consistencies between visual and haptic percepts. *Vis. Res.* 41(4):449–61
- Attneave F. 1968. Triangles as ambiguous figures. *Am. J. Psychol.* 81(3):447–53
- Banks MS, Burge J, Held RT. 2011. The statistical relationship between depth, visual cues and human perception. In *Sensory Cue Integration*, ed. J Trommershauser, K Kording, MS Landy, pp. 1–33. Oxford, UK: Oxford Univ. Press
- Barlow H. 1961. Possible principles underlying the transformations of sensory messages. In *Sensory Communication*, ed. WA Rosenblith, pp. 217–34. Cambridge, MA: MIT Press
- Bartels A, Logothetis NK. 2010. Binocular rivalry: a time dependence of eye and stimulus contributions. *J. Vis.* 10(12):3
- Berkeley G. 1709. *An Essay Towards a New Theory of Vision*. Dublin: Aaron Rhames
- Blake R, Logothetis NK. 2002. Visual competition. *Nat. Rev. Neurosci.* 3:13–21
- Bloj MG, Kersten D, Hurlbert AC. 1999. Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature* 402:877–79
- Boring EG. 1942. *Sensation and Perception in the History of Experimental Psychology*. New York: Appleton-Century-Crofts
- Brascamp JW, Blake R. 2012. Inattention abolishes binocular rivalry: perceptual evidence. *Psychol. Sci.* 23(10):1159–67
- Brascamp JW, Cuthbert P, Ling S. 2020. Conflict defined by global gestalt can modulate binocular rivalry suppression. *J. Vis.* 20 (13):3
- Brascamp JW, Knapen T, Kanai R, Ee R, Berg A. 2007. Flash suppression and flash facilitation in binocular rivalry. *J. Vis.* 7(12):12
- Brascamp JW, Knapen T, Kanai R, Noest A, Ee R, Berg A. 2008. Multi-timescale perceptual history resolves visual ambiguity. *PLOS ONE* 3(1):e1497
- Brascamp JW, Pearson J, Blake R, Berg A. 2009. Intermittent ambiguous stimuli: implicit memory causes periodic perceptual alternations. *J. Vis.* 9(3):3
- Breese BB. 1899. On inhibition. *Psychol. Rev. Monogr. Suppl.* 3:1–65
- Brunswick E, Kamiya J. 1953. Ecological cue-validity of “proximity” and of other Gestalt factors. *Am. J. Psychol.* 66:20–32
- Carpenter G, Grossberg S. 2003. Adaptive resonance theory. In *The Handbook of Brain Theory and Neural Networks*, ed. MA Arbib, pp. 87–90. Cambridge, MA: MIT Press. 2nd ed.
- Casati R, Cavanagh P. 2019. *The Visual World of Shadows*. Cambridge, MA: MIT Press

- Castellotti S, Conti M, Feitosa-Santana C, Del Viva MM. 2020. Pupillary response to representations of light in paintings. *J. Vis.* 20(10):14
- Cavanagh P. 2011. Visual cognition. *Vis. Res.* 51(13):1538–51
- Chopin A, Mamassian P. 2012. Predictive properties of visual adaptation. *Curr. Biol.* 22(7):622–26
- Christiansen JH, D'Antona AD, Shevell SK. 2017. Chromatic interocular-switch rivalry. *J. Vis.* 17(5):9
- Creed RS. 1935. Observations on binocular fusion and rivalry. *J. Physiol.* 84:381–92
- DeBruine L, Welling L, Jones B, Little A. 2010. Opposite effects of visual versus imagined presentation of faces on subsequent sex perception. *Vis. Cogn.* 18(6):816–28
- Diaz-Caneja E. 1928. Sur l'alternance binoculaire. *Ann. Ocul.* 165:721–31
- Dieter K, Brascamp J, Tadin D, Blake R. 2016. Does visual attention drive the dynamics of bistable perception? *Atten. Percept. Psychophys.* 78(7):1861–73
- Doerschner K, Fleming RW, Yilmaz O, Schrater PR, Hartung B, Kersten D. 2011. Visual motion and the perception of surface material. *Curr. Biol.* 21(23):2010–16
- Driver J, Davis G, Russell C, Turatto M, Freeman E. 2001. Segmentation, attention and phenomenal visual objects. *Cognition* 80(1–2):61–95
- D'Zmura M, Lennie P. 1986. Mechanisms of color constancy. *J. Opt. Soc. Am. A* 3:1662–72
- Elder J, Goldberg R. 2002. Ecological statistics of Gestalt laws for the perceptual organization of contours. *J. Vis.* 2(4):5
- Ernst M, Banks M. 2002. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415(6870):429–33
- Ernst M, Bühlhoff H. 2004. Merging the senses into a robust percept. *Trends Cogn. Sci.* 8(4):162–69
- Fischer J, Whitney D. 2014. Serial dependence in visual perception. *Nat. Neurosci.* 17(5):738–43
- Fiser J, Berkes P, Orbán G, Lengyel M. 2010. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* 14:119–30
- Foster DH. 2011. Color constancy. *Vis. Res.* 51:674–700
- Freeman E, Driver J. 2006. Subjective appearance of ambiguous structure-from-motion can be driven by objective switches of a separate less ambiguous context. *Vis. Res.* 46(23):4007–23
- Friston K. 2005. A theory of cortical responses. *Philos. Trans. R. Soc. Lond. B* 360(1456):815–36
- Fritsche M, Mostert P, Lange F. 2017. Opposite effects of recent history on perception and decision. *Curr. Biol.* 27(4):590–95
- Geisler WS, Perry JS, Super BJ, Gallogly DP. 2001. Edge co-occurrence in natural images predicts contour grouping performance. *Vis. Res.* 41:711–24
- Gregory R. 1980. Perceptions as hypotheses. *Philos. Trans. R. Soc. Lond. B* 290(1038):181–97
- Gregory RL. 1997. Knowledge in perception and illusion. *Philos. Trans. R. Soc. Lond. B* 352:1121–28
- Grossmann J, Dobbins A. 2003. Differential ambiguity reduces grouping of metastable objects. *Vis. Res.* 43(4):359–69
- Hatfield G. 2002. Perception as unconscious inference. In *Perception and the Physical World*, ed. D Heyer, R Mausfeld, pp. 115–44. Oxford, UK: Oxford Univ. Press
- Hatfield G, Epstein W. 1979. The sensory core and the medieval foundations of early modern perceptual theory. *Isis* 70(3):363–84
- Hillis J, Watt S, Landy M, Banks M. 2004. Slant from texture and disparity cues: optimal cue combination. *J. Vis.* 4(12):1
- Hochberg J. 1981. On cognition in perception: perceptual coupling and unconscious inference. *Cognition* 10(1–3):127–34
- Holmes DJ, Hancock S, Andrews TJ. 2006. Independent binocular integration for form and color. *Vis. Res.* 46:665–77
- Hong SW, Shevell SK. 2008. Binocular rivalry between identical retinal stimuli with an induced color difference. *Vis. Neurosci.* 25:361–64
- Hong SW, Shevell SK. 2009. Color binding errors during rivalrous suppression of form. *Psychol. Sci.* 20:1084–91
- Howard I. 1996. Alhazen's neglected discoveries of visual phenomena. *Perception* 25(10):1203–17
- Howard I. 2012. *Perceiving in Depth*, Vol. 3: *Other Methods of Depth Perception*. Oxford, UK: Oxford Univ. Press

- Huber D, O'Reilly R. 2003. Persistence and accommodation in short-term priming and other perceptual paradigms: temporal segregation through synaptic depression. *Cogn. Sci.* 27(3):403–30
- Jazayeri M, Movshon JA. 2006. Optimal representation of sensory information by neural populations. *Nat. Neurosci.* 9:690–96
- Kanai R, Verstraten F. 2005. Perceptual manifestations of fast neural plasticity: motion priming, rapid motion aftereffect and perceptual sensitization. *Vis. Res.* 45(25–26):3109–16
- Kim I, Hong SW, Shevell SK, Shim WM. 2020. Neural representations of perceptual color experience emerge along the human visual hierarchy. *PNAS* 117:13145–50
- Kleffner D, Ramachandran V. 1992. On the perception of shape from shading. *Percept. Psychophys.* 52(1):18–36
- Klink P, Noest A, Holten V, Berg A, Wezel R. 2009. Occlusion-related lateral connections stabilize kinetic depth stimuli through perceptual coupling. *J. Vis.* 9(10):20
- Kovács I, Papathomas TV, Yang M, Fehér Á. 1996. When the brain changes its mind: interocular grouping during binocular rivalry. *PNAS* 93:15508–11
- Lamme V, Zipser K, Spekreijse H. 1998. Figure-ground activity in primary visual cortex is suppressed by anesthesia. *PNAS* 95(6):3263–68
- Lee SH, Blake R. 1999. Rival ideas about binocular rivalry. *Vis. Res.* 39:1447–54
- Lee T, Mumford D. 2003. Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am. A* 20(7):1434–48
- Leopold D, Wilke M, Maier A, Logothetis N. 2002. Stable perception of visually ambiguous patterns. *Nat. Neurosci.* 5(6):605–9
- Lindberg DC. 1976. *Theories of Vision from Al-Kindi to Kepler*. Chicago: Univ. Chicago Press
- Logothetis N, Leopold D, Sheinberg D. 1996. What is rivalling during binocular rivalry? *Nature* 380:621–24
- Long GM, Toppino TC. 1994. Adaptation effects and reversible figures: a comment on Horlitz and O'Leary. *Percept. Psychophys.* 56(5):605–10
- Long GM, Toppino TC. 2004. Enduring interest in perceptual ambiguity: alternating views of reversible figures. *Psychol. Bull.* 130:748–68
- Long GM, Toppino TC, Mondin G. 1992. Prime time: fatigue and set effects in the perception of reversible figures. *Percept. Psychophys.* 52(6):609–16
- Mack A, Tang B, Tuma R, Kahn S, Rock I. 1992. Perceptual organization and attention. *Cogn. Psychol.* 24(4):475–501
- Maloney RT, Lam SK, Clifford CWG. 2013. Colour misbinding during motion rivalry. *Biol. Lett.* 9:20120899
- Maniatis LM. 2017. Symmetry and uprightness in visually perceived forms. In *The Oxford Compendium of Visual Illusions*, ed. AG Shapiro, D Todorovic, pp. 234–37. Oxford, UK: Oxford Univ. Press
- Marlow PJ, Anderson BL. 2016. Motion and texture shape cues modulate perceived material properties. *J. Vis.* 16(1):5
- Matthews GB. 1978. A medieval theory of vision. In *Studies in Perception: Interrelations in the History of Philosophy of Science*, ed. PK Machamer, RG Turnbull, pp. 186–99. Columbus: Ohio State Univ. Press
- Maus G, Chaney W, Liberman A, Whitney D. 2013. The challenge of measuring long-term positive aftereffects. *Curr. Biol.* 23(10):R438–39
- McGurk H, MacDonald J. 1976. Hearing lips and seeing voices. *Nature* 264:746–48
- Meyerling TC. 1989. *Historical Roots of Cognitive Science: The Rise of a Cognitive Theory of Perception from Antiquity to the Nineteenth Century*. Berlin: Springer
- Mumford D. 1992. On the computational architecture of the neocortex. *Biol. Cybernet.* 66(3):241–51
- Murray RF. 2021. Lightness perception in complex scenes. *Annu. Rev. Vis. Sci.* 7:417–36
- Necker L. 1832. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phenomenon which occurs on viewing a figure of a crystal or geometrical solid. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* 1(5):329–37
- Noest A, Ee R, Nijs M, Wezel R. 2007. Percept-choice sequences driven by interrupted ambiguous stimuli: a low-level neural model. *J. Vis.* 7(8):10
- Oruç İ, Barton J. 2010. A novel face aftereffect based on recognition contrast thresholds. *Vis. Res.* 50(18):1845–54
- O'Shea RP. 2011. Binocular rivalry stimuli are common but rivalry is not. *Front. Hum. Neurosci.* 5:148
- O'Shea RP, Parker A, La Rooy D, Alais D. 2009. Monocular rivalry exhibits three hallmarks of binocular rivalry: evidence for common processes. *Vis. Res.* 49:671–81



- Ossa-Richardson A. 2019. *The History of Ambiguity*. Princeton, NJ: Princeton Univ. Press
- Paffen C, Tadin D, Pas S, Blake R, Verstraten F. 2006. Adaptive center-surround interactions in human vision revealed during binocular rivalry. *Vis. Res.* 46(5):599–604
- Palmer S. 1999. *Vision Science: From Photons to Phenomenology*. Cambridge, MA: MIT Press
- Parker D. 1972. Contrast and size variables and the tilt after-effect. *Q. J. Exp. Psychol.* 24(1):1–7
- Pastukhov A, Braun J. 2008. A short-term memory of multi-stable perception. *J. Vis.* 8(13):7
- Pastukhov A, Füllekrug J, Braun J. 2013a. Sensory memory of structure-from-motion is shape-specific. *Atten. Percept. Psychophys.* 75(6):1215–29
- Pastukhov A, Lissner A, Braun J. 2013b. Perceptual adaptation to structure-from-motion depends on the size of adaptor and probe objects, but not on the similarity of their shapes. *Atten. Percept. Psychophys.* 76(2):473–88
- Pearson J, Clifford C. 2005a. Mechanisms selectively engaged in rivalry: normal vision habituates, rivalrous vision primes. *Vis. Res.* 45(6):707–14
- Pearson J, Clifford C. 2005b. When your brain decides what you see. *Psychol. Sci.* 16:516–19
- Pearson J, Clifford C, Tong F. 2008. The functional impact of mental imagery on conscious perception. *Curr. Biol.* 18(13):982–86
- Peiso J, Shevell SK. 2020. Seeing fruit on trees: enhanced perceptual dissimilarity from multiple ambiguous neural representations. *J. Opt. Soc. Am. A* 37:A255–61
- Penny W. 2012. Bayesian models of brain and behaviour. *Int. Sch. Res. Not.* 2012:785791
- Piantadosi ST, Tily H, Gibson E. 2012. The communicative function of ambiguity in language. *Cognition* 122:280–91
- Polat U, Sagi D. 1993. Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vis. Res.* 33(7):993–99
- Poort J, Raudies F, Wannig A, Lamme V. 2012. The role of attention in figure-ground segregation in areas V1 and V4 of the visual cortex. *Neuron* 75(1):143–56
- Ramachandran VS, Anstis SM. 1983. Perceptual organization in moving patterns. *Nature* 304:529–31
- Rao R, Ballard D. 1999. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2(1):79–87
- Roelfsema P. 2006. Cortical algorithms for perceptual grouping. *Annu. Rev. Neurosci.* 29:203–27
- Rushton WAH. 1972. Pigments and signals in colour vision. *J. Physiol.* 220(3):1–31P
- Sabra AL. 1978. Sensation and inference in Alhazen's theory of visual perception. In *Studies in Perception: Interrelations in the History of Philosophy of Science*, ed. PK Machamer, RG Turnbull, pp. 160–85. Columbus: Ohio State Univ. Press
- Schleidt W, Shalter M, Moura-Neto H. 2011. The hawk/goose story: the classical ethological experiments of Lorenz and Tinbergen, revisited. *J. Comp. Psychol.* 125(2):121–33
- Scholte H, Witteveen S, Spekreijse H, Lamme V. 2006. The influence of inattention on the neural correlates of scene segmentation. *Brain Res.* 1076(1):106–15
- Schwartz O, Hsu A, Dayan P. 2007. Space and time in visual context. *Nat. Rev. Neurosci.* 8(7):522–35
- Schwiedrzik C, Ruff C, Lazar A, Leitner F, Singer W, Melloni L. 2014. Untangling perceptual memory: Hysteresis and adaptation map into separate cortical networks. *Cereb. Cortex* 24(5):1152–64
- Shams L, Kamitani Y, Shimojo S. 2000. What you see is what you hear. *Nature* 408(6814):788
- Shepard RN. 1990. *Mindsights*. New York: W. H. Freeman
- Shevell SK. 2003. Color appearance. In *The Science of Color*, ed. SK Shevell, pp. 149–90. Amsterdam: Elsevier. 2nd ed.
- Shevell SK. 2019. Ambiguous chromatic neural representations: perceptual resolution by grouping. *Curr. Opin. Behav. Sci.* 30:194–202
- Shevell SK, Kingdom FAA. 2008. Color in complex scenes. *Annu. Rev. Psychol.* 59:143–66
- Shimojo S, Nakayama K. 1990. Real world occlusion constraints and binocular rivalry. *Vis. Res.* 30(1):69–80
- Silver MA, Logothetis NK. 2007. Temporal frequency and contrast tagging bias the type of competition in interocular switch rivalry. *Vis. Res.* 47:532–43
- Slezak E, Shevell SK. 2018. Perceptual resolution of color for multiple chromatically ambiguous objects. *J. Opt. Soc. Am. A* 35:B85–91

- Smith VC, Pokorny J. 2003. Color matching and color discrimination. In *The Science of Color*, ed. SK Shevell, pp. 103–48. Amsterdam: Elsevier. 2nd ed.
- Stuit SM, Paffen CLE, van der Smagt MJ, Verstraten FAJ. 2011. What is grouping during binocular rivalry? *Front. Hum. Neurosci.* 5:117
- Tadin D, Lappin J, Gilroy L, Blake R. 2003. Perceptual consequences of centre-surround antagonism in visual motion processing. *Nature* 424(6946):312–15
- Tanaka Y, Sagi D. 1998. A perceptual memory for low-contrast visual signals. *PNAS* 95(21):12729–33
- Tong F, Nakayama K, Vaughan J, Kanwisher N. 1998. Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* 21(4):753–59
- Toscani M, Gegenfurtner KR, Doerschner K. 2017. Differences in illumination estimation in #thedress. *J. Vis.* 17(1):22
- Treisman A. 1962. Binocular rivalry and stereoscopic depth perception. *Q. J. Exp. Psychol.* 14:23–29
- Troje NF. 2017. The Kayahara silhouette illusion. In *The Oxford Compendium of Visual Illusions*, ed. AG Shapiro, D Todorovic, pp. 582–85. Oxford, UK: Oxford Univ. Press
- Tsutsui K-I, Sakata H, Naganuma T, Taira M. 2002. Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298:409–12
- Von der Heydt R. 2004. Image parsing mechanisms of the visual cortex. In *The Visual Neurosciences*, ed. LM Chalupa, JS Werner, pp. 665–80. Cambridge, MA: MIT Press
- von Helmholtz H. 1867. Concerning the perceptions in general. In *Treatise on Physiological Optics*, Vol. III, transl. JPC Southall. New York: Dover. 3rd ed.
- Webster MA. 2020. The Verriest Lecture: adventures in blue and yellow. *J. Opt. Soc. Am. A* 37:V1–14
- Webster MA, Mollon JD. 1997. Adaptation and the color statistics of natural images. *Vis. Res.* 37:3283–98
- Wheatstone C. 1838. Contributions to the physiology of vision. Part the first. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philos. Trans. R. Soc. Lond.* 128:371–79
- Winawer J, Huk A, Boroditsky L. 2010. A motion aftereffect from visual imagery of motion. 114:276–84
- Wolfe J. 1984. Reversing ocular dominance and suppression in a single flash. *Vis. Res.* 24(5):471–78
- Xing J, Heeger D. 2001. Measurement and modeling of center-surround suppression and enhancement. *Vis. Res.* 41(5):571–83
- Yang JN, Maloney LT. 2001. Illuminant cues in surface color perception: tests of three candidate cues. *Vis. Res.* 41:2581–600
- Young M, Landy M, Maloney L. 1993. A perturbation analysis of depth perception from combinations of texture and motion cues. *Vis. Res.* 33(18):2685–96
- Zhang P, Jamison K, Engel S, He B, He S. 2011. Binocular rivalry requires visual attention. *Neuron* 71(2):362–69
- Zhang X, Qiu J, Zhang Y, Han S, Fang F. 2014. Misbinding of color and motion in human visual cortex. *Curr. Biol.* 24:1354–60
- Zipf G. 1949. *Human Behavior and the Principle of Least Effort*. New York: Addison-Wesley
- Zou J, He S, Zhang P. 2016. Binocular rivalry from invisible patterns. *PNAS* 113(30):8408–13