# THE ROYAL SOCIETY

# Why we see things the way we do: evidence for a wholly empirical strategy of vision

## Dale Purves[*], R. Beau Lotto, S. Mark Williams, Surajit Nundy and Zhiyong Yang

*Box 3209, Duke University Medical Center, Durham, NC 27710, USA*

Many otherwise puzzling aspects of the way we see brightness, colour, orientation and motion can be understood in wholly empirical terms. The evidence reviewed here leads to the conclusion that visual percepts are based on patterns of reflex neural activity shaped entirely by the past success (or failure) of visually guided behaviour in response to the same or a similar retinal stimulus. As a result, the images we see accord with what the sources of the stimuli have typically turned out to be, rather than with the physical properties of the relevant objects. If vision does indeed depend upon this operational strategy to generate optimally useful perceptions of inevitably ambiguous stimuli, then the underlying neurobiological processes will eventually need to be understood within this conceptual framework.

**Keywords:** visual illusions; perception; brightness; colour; angles; motion

## 1. INTRODUCTION

The fundamental problem in vision was stated at the beginning of the 18th century by George Berkeley (1709), who pointed out that the sources underlying visual stimuli are unknowable in any direct sense. In modern terms, since the light returned to the eye from any scene conflates the contributions of reflectance, illumination and transmittance (and a host of other factors that affect these parameters), the provenance of any retinal stimulus (and therefore its significance for visually guided behaviour) is profoundly and inevitably ambiguous. This fundamental fact presents a biological dilemma. Successful behaviour in a complex and potentially hostile environment clearly depends on responding appropriately to the physical sources of visual stimuli; but, if the pattern of retinal activity generated by light cannot uniquely define the underlying reality that the observer must deal with, how then does the visual system routinely generate behaviours that accord with sources of visual stimuli? (See table 1 for definitions.)

The central tenet of the theory of vision that has emerged from the evidence summarized here is that this dilemma is solved by having proximal stimuli trigger patterns of neuronal activity that have been shaped solely by the past consequences of visually guided behaviour. In this concept of vision, the perceived images (which we take to be conscious manifestations of the different patterns of activity elicited by visual stimuli) will necessarily accord with the history of what the same or similar stimuli have turned out to be. This operational determination of what we see can explain the fact that visual percepts do not systematically covary with the characteristics of a light stimulus or the physical properties of the objects that generated the stimulus. Much to the advantage of the observer, percepts covary with the efficacy of past actions in response to visual stimuli, and thus only coincidentally with the measured properties of the stimulus or the underlying objects. This strategy ensures that visually guided responses will usually deal successfully with the objects and conditions that have given rise to retinal stimuli whose sources are, as Berkeley pointed out, unknowable in any direct way.

Of course, this (or any) theory of vision has little value if it cannot account for the perceptual phenomena that vision scientists have long recognized as challenges to rationalizing what we see. These challenges include understanding (i) the relationship between luminance and sensations of brightness; (ii) the relationship between spectral stimuli and sensations of colour; (iii) the relationship between the orientation of objects in space and the orientations perceived; and (iv) the relationship between the direction and speed of moving objects and the motion perceived. The following sections summarize evidence showing how each of these much-debated aspects of vision is explained by the theory outlined here.

## 2. THE RELATIONSHIP BETWEEN LUMINANCE AND BRIGHTNESS

It has long been apparent that the perceived brightness of objects does not correspond in any simple way to their luminance (i.e. to the measured intensity of light corrected for the spectral sensitivity of the human visual system). In particular, two surfaces returning the same amount of light to the eye can look differently bright if the surfaces are observed in different contexts, a phenomenon called simultaneous brightness contrast (figure 1a). (Psychologists prefer to call this 'lightness' contrast to distinguish the appearance of reflected light from the appearance of endogenous sources of light; for present purposes, this distinction is not critical, and the term 'brightness' is used here in its ordinary meaning of perceived intensity).

[*]Author for correspondence (purves@neuro.duke.edu).

Table 1. *Definitions of some key terms*

*objects* the physical entities that give rise to a distal stimulus by reflecting light (or by emitting light, if, as more rarely happens, they are themselves generators of light)

*combine stimulus* the light emanating from objects that reaches the eye and interacts with photoreceptors, thus setting up a pattern of retinal activity

*scene* the arrangement of objects and their illumination with respect to the observer that gives rise to the distal and proximal stimuli ('source' is used as a synonym)

*visually guided response*s the actions undertaken by the observer to deal with the objects and conditions in the scene giving rise to the proximal stimulus

*empirical* on the basis of past experience

*empirical significance* what the proximal stimulus has typically turned out to be

*visual percepts* the patterns of central neuronal activity triggered by the proximal stimulus

The explanation of this remarkable effect found in most textbooks is predicated on lateral interactions among retinal ganglion cells or other lower-order visual neurons, which demonstrably cause distorted rates of neuronal firing at contrast boundaries, presumably to enhance the detection of edges (Kuffler 1953; Barlow 1953). This interpretation implies that the relative intensities perceived in response to such stimuli are, in effect, 'readouts' of the relative firing rate of neurons at the input stages of the visual system (e.g. Mach 1914; Ratliff 1965; Cornsweet 1970). In short, any target predominantly surrounded by an area of higher luminance should look darker than the same target predominantly surrounded by an area of lower luminance.

Despite the apparent concordance of perception and retinal physiology in this instance, a number of investigators have shown that identical targets embedded in scenes that have exactly the same local contrast relationships with their surrounds can nonetheless look differently bright (e.g. Wertheimer 1912; Benary 1924; Gilchrist 1977; Chubb *et al*. 1989; Adelson 1993; Agostini & Proffitt 1993; Todorovic 1997). Indeed it is even possible to construct a stimulus in which a target in a predominantly higher luminance surround looks brighter than an identical target in a predominantly lower luminance surround (White 1979).

How, then, can these seeming contradictions in the relationship of luminance and brightness be explained?

### (a) *An empirical explanation of simultaneous brightness contrast*

In terms of the empirical strategy of vision outlined here, the explanation of the difference in perceived brightness of the two equiluminant targets in figure 1a is as follows (Williams *et al*. 1998a,b; see also Lotto & Purves 1999). Since the amount of light returned to the eye from any portion of a scene depends on both the illumination of the relevant surfaces and their reflectances, the equiluminant returns from the targets in figure 1a are inherently ambiguous. Such stimuli will often have been generated by similarly reflective surfaces on differently reflective surrounds under the same illuminant (figure 1b); the same luminance profiles, however, will often have signified differently reflective target surfaces under different amounts of illumination (figure 1c).

Since dealing successfully with this or any stimulus depends on responding appropriately to the sources of the

returns rather than the returns themselves, we have argued that the visual system solves this problem in a wholly empirical fashion. Thus, to the extent that the stimulus in figure 1a is consistent with similarly reflective target surfaces under the same illuminant (see figure 1b), the targets will tend to appear similarly bright. However, insofar as the stimulus is consistent with the experience of the visual system with differently reflective objects in different levels of illumination (see figure 1c), the targets will tend to appear differently bright. Because the stimulus in figure 1a is consistent with either of these possible sources (i.e. the scenes illustrated in figure 1b,c), the pattern of neural activity elicited—that is, the percept experienced when looking at figure 1a—is a manifestation of both possibilities (and indeed all of the many other possibilities not illustrated) in proportion to their relative frequency of occurrence in past experience with stimuli of this general sort.

In keeping with this explanation, crafting the stimulus in figure 1a to be more consistent with differently reflective surfaces in different illuminants (figure 1c) increases the 'illusion' of brightness contrast, whereas making the stimulus less consistent with this possibility, and more consistent with the source being similarly reflective objects under similar illuminants (figure 1b) causes the targets to appear more alike, even if all the luminance relationships in the scene are preserved (Williams *et al*. 1998a,b; Lotto & Purves 1999, 2001).

### (b) *An empirical explanation of the Craik–O'Brien–Cornsweet effect*

According to this wholly empirical way of understanding the relationship between luminance and brightness, whenever a given stimulus is consistent with the experience of equiluminant targets signifying differently reflective objects, the brightnesses of the returns should appear different. If this idea is correct, then the same perceptual effect elicited by the stimulus profile in figure 1a should be generated by any stimulus in which regions with the same luminance would have typically turned out to be differently reflective objects in different amounts of light. We therefore sought to test this prediction by examining other sorts of brightness 'illusions'.

In figure 2a, the two equiluminant territories adjoining an opposing pair of luminance gradients that meet along a linear boundary appear differently bright. This stimulus, called the Cornsweet edge, belongs to a
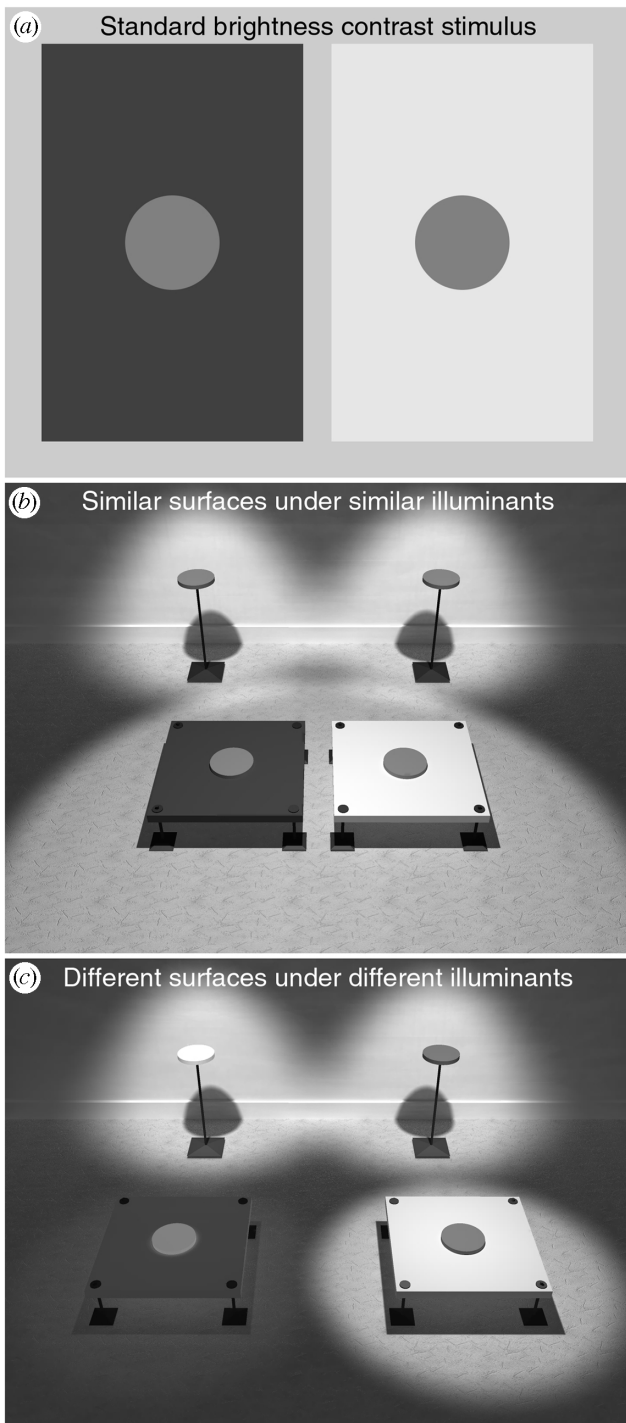
Figure 1. Simultaneous brightness contrast and its empirical explanation. (*a*) Example of a standard simultaneous brightness contrast stimulus. (*b*) The stimulus in (*a*) could have been generated by similarly reflective targets (the circular patches) on differently reflective backgrounds under the same illuminant. (*c*) Alternatively, the targets could be differently reflective surfaces on similarly reflective backgrounds under different amounts of illumination. The theory outlined here contends that the visual system deals with this inevitable ambiguity by incorporating into the percept these (and other) possibilities according to the frequency of their occurrence in the past experience of the species and the individual. (After Williams *et al.* 1998*b*.)

larger class of edges referred to as Craik–O'Brien–Cornsweet edges, after the several psychologists who devised such stimuli (O'Brien 1959; Craik 1966; Cornsweet 1970). Although this illusion has also been rationalized in terms of lateral interactions among neurons in the input stages of the visual system (see, for example, Ratliff 1965; Cornsweet 1970), the region that appears brighter in figure 2*a* actually borders the gradient of higher average luminance, whereas the region that appears darker is next to the gradient of lower average luminance, a profile that is in this respect opposite to the arrangement of luminances in standard simultaneous brightness contrast stimuli (see figure 1*a*).

Despite these apparent differences, the effects elicited by both types of stimuli can be understood in the same empirical terms (Purves *et al.* 1999). The common denominator of the Cornsweet stimulus in figure 2 and the simultaneous brightness contrast stimulus in figure 1 is simply what the equiluminant 'targets' in the two luminance profiles have typically turned out to be. The empirical reason for the perception elicited by the Cornsweet stimulus in figure 2*a* is that the surfaces bordering the gradients will sometimes have been generated by similarly reflective surfaces under the same illuminant (figure 2*b*); however, the same stimulus will often have been generated by differently reflective surfaces under different intensities of illumination (figure 2*c*). These possible sources influence the reflex response to the stimulus according to their relative frequencies of occurrence, with the result that the two equiluminant regions adjoining the Cornsweet edge look different: the territory adjoining the light gradient appears brighter than the territory adjoining the dark gradient. Consistent with this explanation, the effect can be increased, or decreased, by altering the relative probabilities of the possible sources of the stimulus, as in the case of simultaneous brightness contrast (see figure 1) (Purves *et al.* 1999). In both cases, the target that appears brighter is the one more consistent with the experience of a more reflective surface in relatively less light, whereas the target that appears less bright is more consistent with a less reflective surface in relatively more light.

### (c) *An empirical explanation of Mach bands*

A particularly challenging problem for this way of explaining the anomalous relationship between luminance and brightness is Mach bands, the name given to the light and dark zones seen at the onset and offset, respectively, of luminance gradients that lack any photometric basis for this effect (Mach 1865; Ratliff 1965) (figure 3). It is at first difficult to imagine what empirical (or historical) facts about human interactions with the sources of luminance gradients could explain this gratuitous addition of light and dark bands to the percepts elicited by such stimuli. Nevertheless, there is an entirely empirical explanation of Mach bands that closely parallels the accounts of the simultaneous brightness contrast and Cornsweet effects (Lotto *et al.* 1999*a,b*) (figure 3). By interacting with the objects that give rise to luminance gradients, observers will have experienced that the underlying source of stimuli such as the gradient in figure 4*a* is sometimes (i) differences in the reflectance properties of flat surfaces (as in the Mach stimulus; see
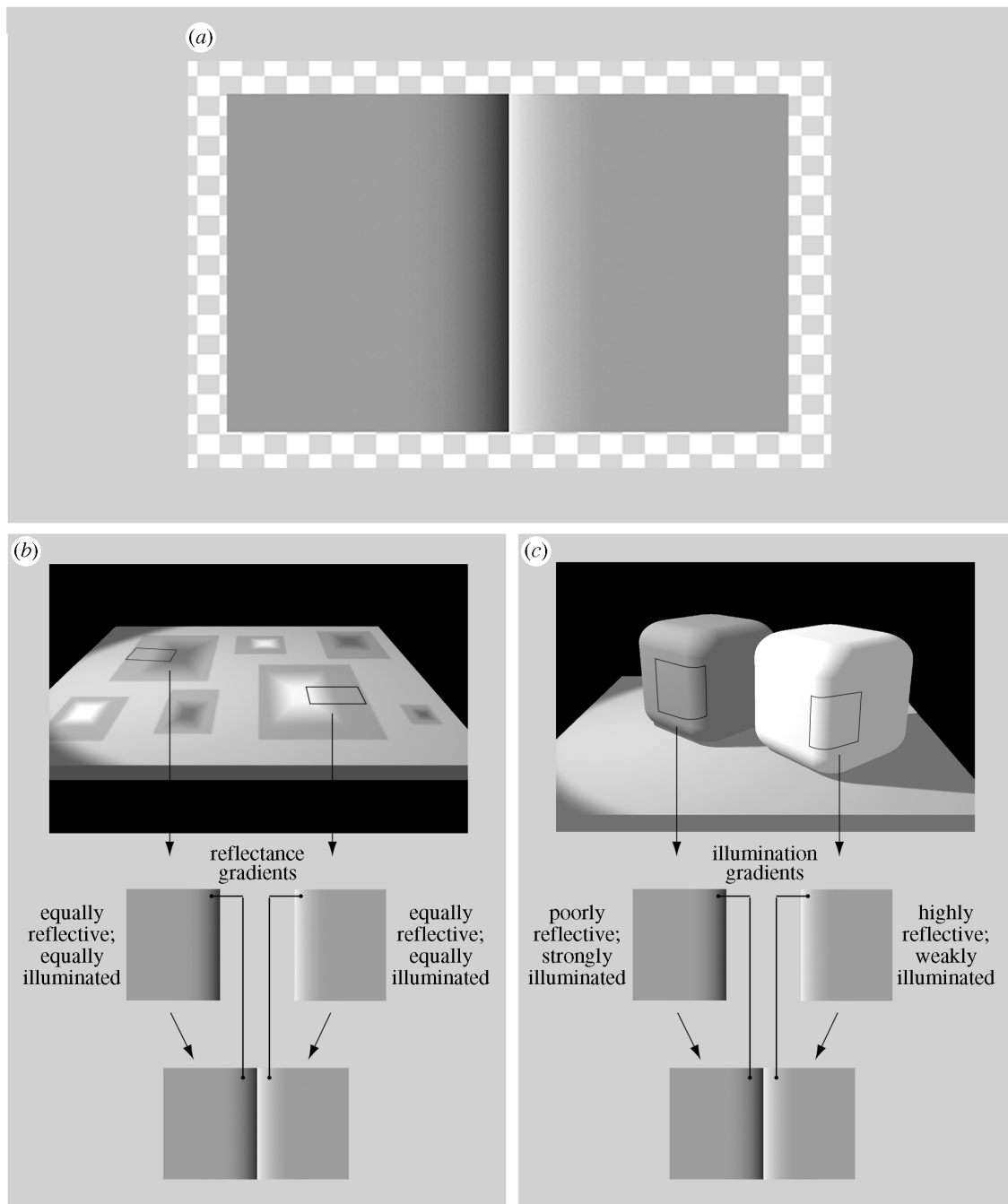
Figure 2. Filling-in elicited by the Cornsweet edge stimulus. (*a*) Standard presentation of the Cornsweet stimulus. Despite the equal luminances of the territories adjoining the two opposing gradients, the territory to the right of the light gradient looks lighter than the territory to the left of the dark gradient. (*b*) The stimulus in (*a*) could have been generated by gradual changes in the reflective properties of the substrate observed under the same illuminant. (*c*) Alternatively, the stimulus could have arisen from gradual changes in the different illumination of two differently reflective surfaces. As a means of contending with this ambiguity, the percept generated by the Cornsweet stimulus incorporates the relative probabilities of the possible sources of the stimulus. In consequence, the two equiluminant territories adjoining the Cornsweet edge appear differently bright. (After Purves *et al*. 1999.)

figure 3), (ii) penumbras (figure 4*b*), and (iii) differences in the illumination of curved surfaces (figure 4*c*) (among other less frequent possibilities not considered here).

   An important aspect of experience derived from interacting with curved surfaces is that the linear luminance gradients associated with such surfaces are frequently adorned with photometric highlights and lowlights at the beginning and end of the gradient. Highlights are a consequence of the relatively greater amount of light returned to the observer from curved surfaces that are to some degree specular; lowlights arise because objects on the surface of the Earth are typically illuminated by indirect as well as direct light (see Lotto *et al*. (1999*a*,*b*) for a fuller explanation of these phenomena). Since the linear luminance gradient in the Mach stimulus in figure 3*a* has attributes associated with the stimuli generated by a curved surface or a penumbra on a flat surface (or the specific surface reflectances, which are the source

(a) stationary disc     spinning disc

(b)
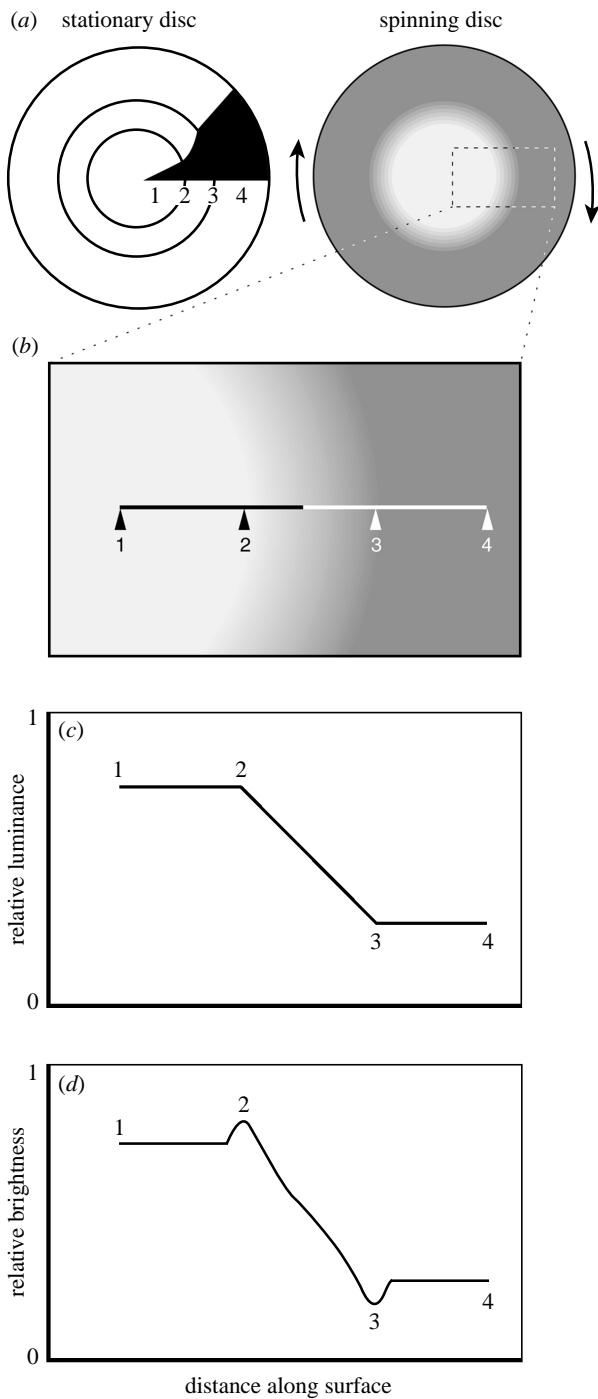
(c)

(d)

distance along surface

Figure 3. Mach bands. (a) Diagram of the painted disk used by Mach to elicit the Mach band illusion. When the disk is spun, a luminance gradient is established between points (2) and (3), which links the uniformly lighter centre of the stimulus (1) and the uniformly darker region at its periphery (4). (b) Blow-up of a portion of the stimulus in (a), indicating the nature and position of Mach bands. A band of illusory lightness is apparent at position (2), and a band of illusory darkness at position (3). (c) Because the portion of the black sector between points (2) and (3) in (a) is a segment of an Archimedean spiral, the luminance gradient generated between the corresponding points on the spinning disk is linear, as indicated by this photometric measurement along the line in (b). (d) Diagram of the perception of the photometric profile in (c), indicating the illusory lightness maximum at the initiation of the linear gradient (2), and the illusory minimum at its termination (3). (After Lotto *et al.* 1999*a*.)
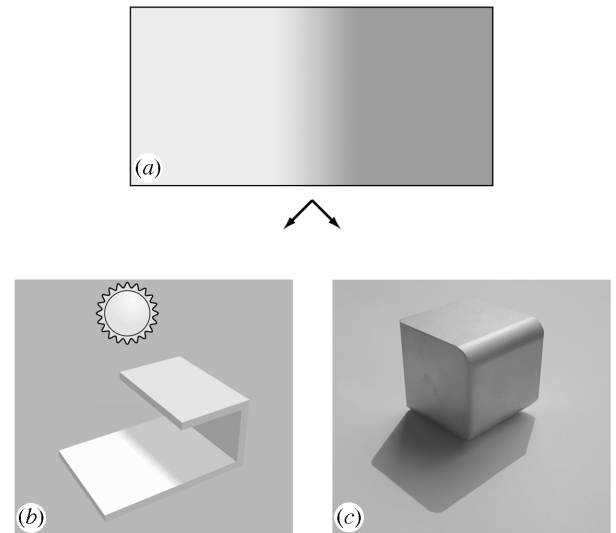
(a)

(b)     (c)

Figure 4. Mach bands explained in the same empirical framework used to rationalize the effect of simultaneous contrast stimuli and Cornsweet edge effects. Like the stimuli in previous figures, the luminance gradient in (a) is ambiguous: the profile could be the penumbra of a cast shadow on a flat surface (b), or a gradient generated by a curved surface (c). The penumbral gradients of cast shadows lack photometric highlights and lowlights, whereas the gradients generated by curved surfaces typically have luminance maxima and minima. In keeping with the parallel explanations of contrast and edge effects, the percepts elicited by luminance gradients incorporate this empirical fact about a major category of real-world luminance gradients, with the result that Mach bands are seen. (After Lotto *et al.* 1999*b*.)

of the printed stimulus) (see figure 4), the percept elicited incorporates highlights and lowlights in proportion to the frequency of their historical occurrence as accompaniments of luminance gradients.

As predicted by this reasoning, proximal stimuli more consistent with a curved surface as the underlying source (which would normally be adorned with highlights and lowlights) elicit a stronger sensation of Mach bands than oppositely biased stimuli, which elicit such sensations weakly or not at all (see Lotto *et al.* 1999*b*). This modulation is similar to the enhancement or diminishment of the simultaneous brightness contrast or edge effects achieved by manipulating the relative probabilities of the possible sources of the proximal stimulus in the examples given earlier.

This empirical explanation of Mach bands is the same, in principle, as the explanations of simultaneous brightness and Cornsweet effects. The common cause is the genesis of visual percepts according to a strategy in which percepts are elicited as reflexes, the network connectivity of which has been wholly determined by the history of human visual experience. Because the perceptual responses to the several stimulus categories considered here manifest this strategy in superficially different ways, the common basis of these effects is less obvious than it might have been.

## 3. THE RELATIONSHIP BETWEEN SPECTRAL RETURNS AND COLOUR

Given that these otherwise puzzling aspects of the sensation of light intensity (brightness) can be understood
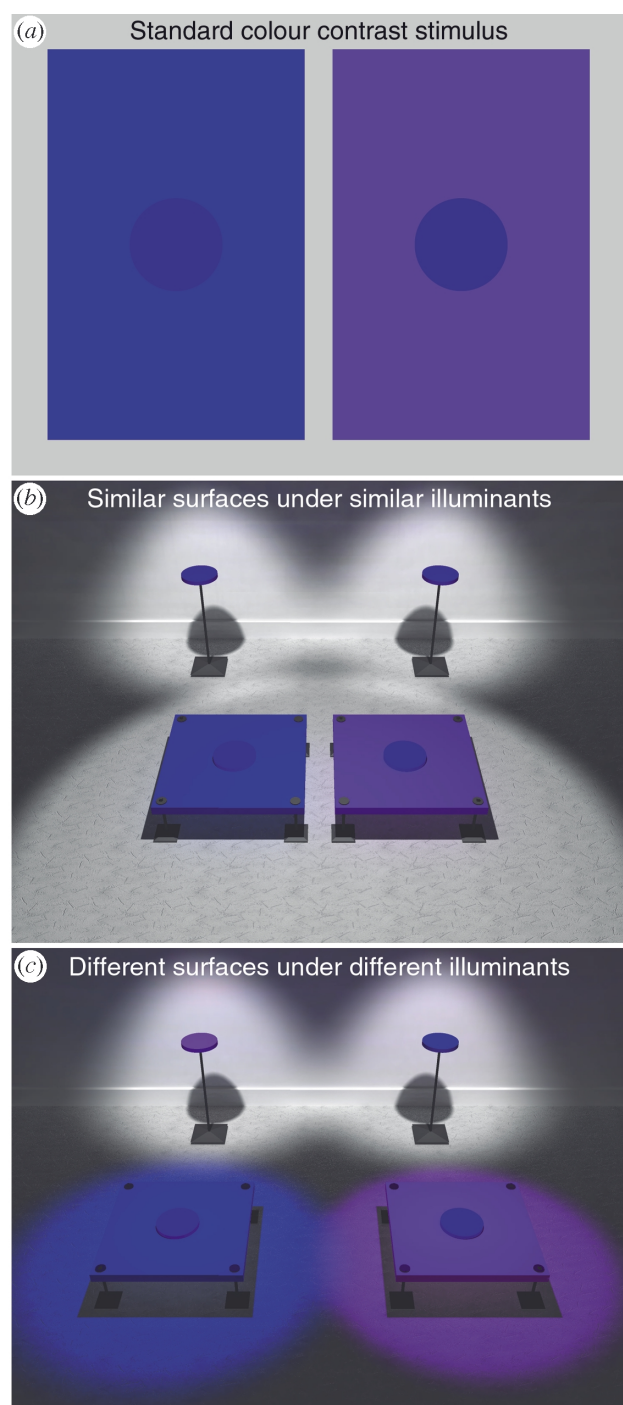
Figure 5. Simultaneous colour contrast and its empirical explanation. (*a*) Example of a standard simultaneous colour contrast stimulus. (*b*) The stimulus in (*a*) could have been generated by similarly reflective targets on differently reflective backgrounds under the same illuminant. (*c*) Alternatively, the targets could be differently reflective surfaces on similarly reflective backgrounds under different amounts of illumination. As a result of the incorporation of both these possible sources into the percept elicited by the stimulus, the identical targets appear to be differently coloured. (After Lotto & Purves 2000.)

reasons as is spectral intensity, to wit the conflation of illumination, reflectance and transmittance in the spectral return. Indeed, if the wholly empirical theory outlined for the perception of luminance has merit, it should apply not only to colour, but to all categories of visual sensation.

A useful starting point in any exploration of the genesis of colour on an empirical basis is simultaneous colour contrast, a phenomenon that shares many similarities with simultaneous brightness contrast. The standard stimulus for eliciting colour contrast is two targets with the same spectral composition on differently chromatic backgrounds (figure 5*a*). As in figure 1*a*, the two targets look different; in this case, however, the perceptual distinction is based on differences in the apparent hue and saturation of the targets rather than brightness alone (colour sensations are generally described in terms of hue, saturation and brightness, all three of which are appreciably different in comparing the two targets in figure 5).

The percepts elicited by the stimulus in figure 5*a*, and similar anomalies in which the same spectral targets elicit different colour sensations, are usually ascribed to 'adaptation' of the colour system to the average spectral content of the overall stimulus (typically at the input stages) (e.g. Von Kries 1905; Jameson & Hurvich 1989; D'Zmura & Lennie 1986; Creutzfeldt *et al.* 1990, 1991; Chichilnisky & Wandell 1995; Walsh 1995; Webster & Mollon 1995), and/ or to computations of spectral ratios across chromatic contrast boundaries (e.g. Land 1986). Both these hypotheses, however, fail to account fully for all the perceptual consequences of such stimuli (Ware & Cowan 1982; Brainard & Wandell 1996; Brown & MacLeod 1997; Lotto & Purves 1999). Moreover, they are really mathematical descriptions rather than explanations, and provide only a limited biological rationale for colour contrast (the truism usually provided is that it makes sense to see an object as having more or less the same colour in different illuminants, and that colour contrast anomalies are the price that must be paid for the supposed benefit of 'colour constancy').

An explanation of colour contrast (and constancy) can, however, be given in fully empirical terms. The sources of the target and surround in the standard colour contrast stimuli shown in figure 5*a* are, like all visual stimuli, profoundly ambiguous: much like the achromatic targets in figure 1*a*, the same spectral patterns could have been generated by many combinations of reflectances, conditions of illumination and influences of transmittance (figure 5*b,c*). The visual system must nevertheless generate appropriate behavioural responses to the enormous variety of the spectral patterns returned to the eye. According to the theory summarized here, the visual system solves this problem by using feedback from the success or failure of these responses to progressively instantiate patterns of neural connectivity that promote ever more appropriate reactions to the stimuli. In this phylogenetic and ontogenetic process, the neuronal activity elicited by spectral stimuli comes to link the spectral profiles of inevitably uncertain provenance with what they typically turned out to be (i.e. with their empirical significance). In this scheme, then, the particular pattern of neuronal activity elicited in response to a given stimulus is ultimately dictated by the relative frequencies
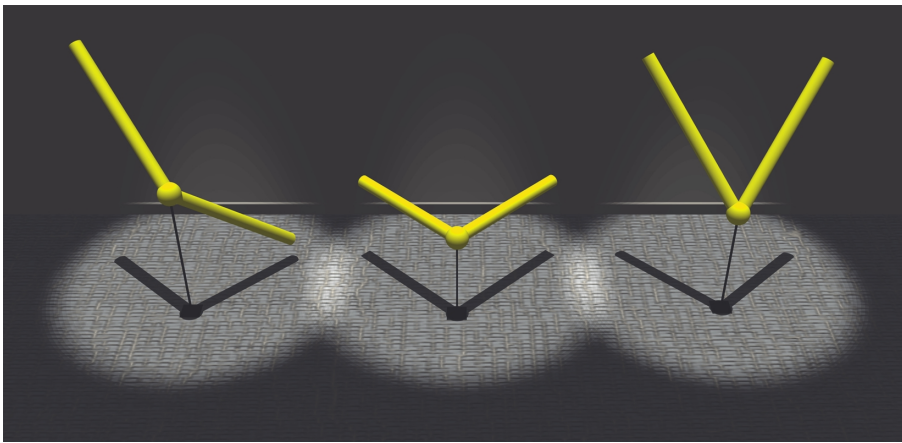
in terms of a wholly empirical concept of how the visual system generates percepts, it is natural to ask whether the colour sensations elicited by different spectra might arise according to the same scheme. After all, the spectral qualities of a stimulus are ambiguous for exactly the same

Figure 6. Diagram illustrating the ambiguity of angle projections. Three different angular objects having real-world subtenses of 120°, 90° and 60°, respectively, can project identically onto a plane (or the retina). (After Nundy *et al*. 2000.)
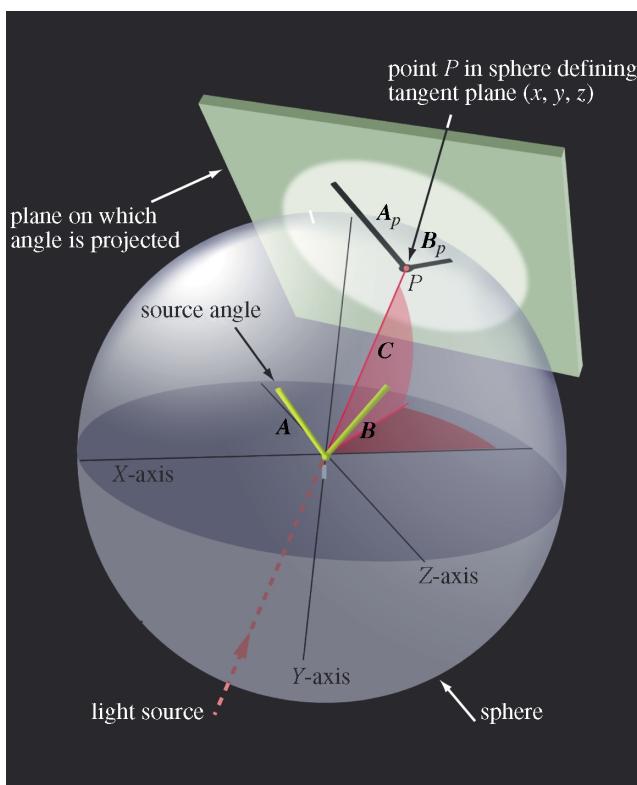


Figure 7. Computing the possible sources of a projected angle. An angular object is arbitrarily placed in the $XY$ plane with its vertex at the centre of a sphere and its arms specified by the two vectors, $\boldsymbol{A}$ and $\boldsymbol{B}$. The set of points within the sphere is first generated by calculating random values for $x$, $y$ and $z$ between $-1$ and $+1$, keeping only those points satisfying $x^2 + y^2 + z^2 \leqslant 1$ (where $x$, $y$ and $z$ are the coordinates along the three conventional Cartesian axes $X$, $Y$ and $Z$ of any point $P$). The projection $(\boldsymbol{A}_p, \boldsymbol{B}_p)$ of each of the arms of the angular object $(\boldsymbol{A}, \boldsymbol{B})$ onto a plane perpendicular to the vector joining $P$ and the centre of the sphere $(\boldsymbol{C})$ is then determined by vector calculus based on the following relationships: $\boldsymbol{A}_p = \boldsymbol{A} + t_A \boldsymbol{C}$, and $\boldsymbol{B}_p = \boldsymbol{B} + t_B \boldsymbol{C}$. Further, $\boldsymbol{A}_p \times (t_A \boldsymbol{C}) = 0$, and $\boldsymbol{B} \times (t_B \boldsymbol{C}) = 0$, where $t_A$ and $t_B$ are scalars. After solving for $t_A$ and $t_B$ and subsequently for $\boldsymbol{A}_p$ and $\boldsymbol{B}_p$, the angle of the projection onto the plane is given as: projected angle $= \cos^{-1}[(\boldsymbol{A}_p \times \boldsymbol{B}_p)/(|\boldsymbol{A}_p||\boldsymbol{B}_p|)]$. Using this method, the projections of all angles subtending 0–180° are determined (in 1° increments) and stored in a data file. The file can then be searched for all the source subtenses and orientations that could give rise to a particular projection, in this way establishing a probability distribution for the sources of any angle projected onto a plane (see figure 8). (After Nundy *et al*. 2000.)

of occurrence of the real-world combinations of reflectances, illuminants and transmittances that have given rise to that spectral stimulus in the past.

If perceptions of colour are indeed generated in this wholly empirical way, then the same spectral target on two differently chromatic backgrounds should give rise to different chromatic sensations (see figure 5*a*). The reason is that, in addition to requiring behaviours appropriate to the same reflectances in the same illuminant (i.e. the stimulus on the printed page) (figure 5*b*), such stimuli will in other instances have required behaviours appropriate to targets that arise from different reflectances in different illuminants (figure 5*c*). Consequently, the pattern of spectral returns in figure 5*a* elicits a pattern of neuronal activity that incorporates these possible underlying sources in proportion to their past occurrence in human experience with spectral stimuli.

The validity of this explanation can be tested by examining (i) whether the colours experienced in response to a full range of colour contrast stimuli in fact accord with (and are predicted by) the physical relationships between reflectances, illuminants and the spectral returns they give rise to; and (ii) whether altering the empirical significance of two targets returning identical spectra changes the colours perceived in the manner predicted by the theory. In both cases, colour experience varies as expected on the basis of this empirical framework (Lotto & Purves 1999, 2000; see also Lotto & Purves 2001).

Evidently the colours we see are, like brightness, linked to the stimuli that generate them by the historical success and/or failure of the interactions of human observers and their predecessors with objects and illuminants in the world.

## 4. THE RELATIONSHIP BETWEEN ORIENTATED LINES AND THEIR PERCEPTION

The evidence for an empirical basis of vision has so far been restricted to the perception of light intensities (luminances) and spectral differences (colours). A very different perceived quality is the form of objects, an aspect of visual experience that depends on a subjective sense of how the edges of objects—lines in the simplest case—are orientated in space with respect to the observer (contrast boundaries that define edges are presumably the critical determinants of the forms we see, and indeed of our ability to see objects at all; see, for example, Yarbus 1967).

It has been known since the middle of the 19th century that the perception of orientated lines does not always accord with the real-world geometry of the underlying objects and their retinal projections. Thus, the angles formed by lines making (or implying) an acute angle are seen as being a few degrees larger in subtense than they really are, whereas obtuse angles are seen as being a few degrees smaller (Hering 1864; Biehler 1896; Wundt 1902; Helmholtz 1924). Despite a great deal of speculation about this anomaly, there has been no consensus regarding its origin. In the modern literature, these discrepancies in the perception of angles have usually been explained in terms of complex inhibitory inter-actions among orientation-selective neurons in the primary visual cortex (e.g. Von Bekesy 1967; Howard 1971; Carpenter & Blakemore 1973; Tyler & Nakayama 1984).

The anomalous way we perceive angles can, however, be explained in empirical terms, similar to the accounts of the way we perceive brightness and colour (Nundy *et al.* 2000). The proximal stimuli that give rise to perceived angles, like the luminances or the spectral content of the returns, are profoundly ambiguous. As illustrated in figure 6, an angle projected onto a surface (the retina, for example) can arise from angular objects having a variety of subtenses and arm lengths, arranged in infinitely many orientations. In interacting with the objects that give rise to particular retinal projections, observers will have found that the real-world angles giving rise to the prox-imal stimuli vary greatly, and, as it turns out, system-atically (see figure 8 below). In consequence, the perception elicited by an angle projected onto the retina should correspond to the frequency distribution of the possible sources underlying the proximal stimulus in phylogenetic and ontogenetic experience.

A particular advantage of considering the merits of a wholly empirical theory of vision *vis à vis* the perception of angles (as opposed to brightness or colour) is the ability to model the cumulative visual experience on which the perceptions of angular stimuli are presumably founded (figure 7). Whereas the frequency distribution of the relevant past experience is difficult to compute for luminance or spectral content, in the case of angles, the major features of the experience that have shaped the rele-vant patterns of neural connectivity elicited by retinal stimulation can be specified by geometrical principles, at least to a first approximation. This information can then be used to predict how angles should be perceived according to the theory proposed here, thus providing a more rigorous test of a wholly empirical basis for vision.

The relative frequency of occurrence of all the possible three-dimensional sources of a projected angle can be assessed by analysing all of the ways a given angular object can project onto a plane (see figure 7). Obviously, a line or any other object can exist in an infinite variety of orientations with respect to the observer. The simplifying assumption in the approach we used is that angular objects occupy these positions with equal probability (in fact, there is a slight bias even in natural scenes toward contours in the cardinal axes, and therefore toward right-angles (Coppola *et al.* 1998)). The distribution obtained in this way can be used in turn to generate the frequency distribution of the subtenses of the objects that could have given rise to any given angular projection. For example,
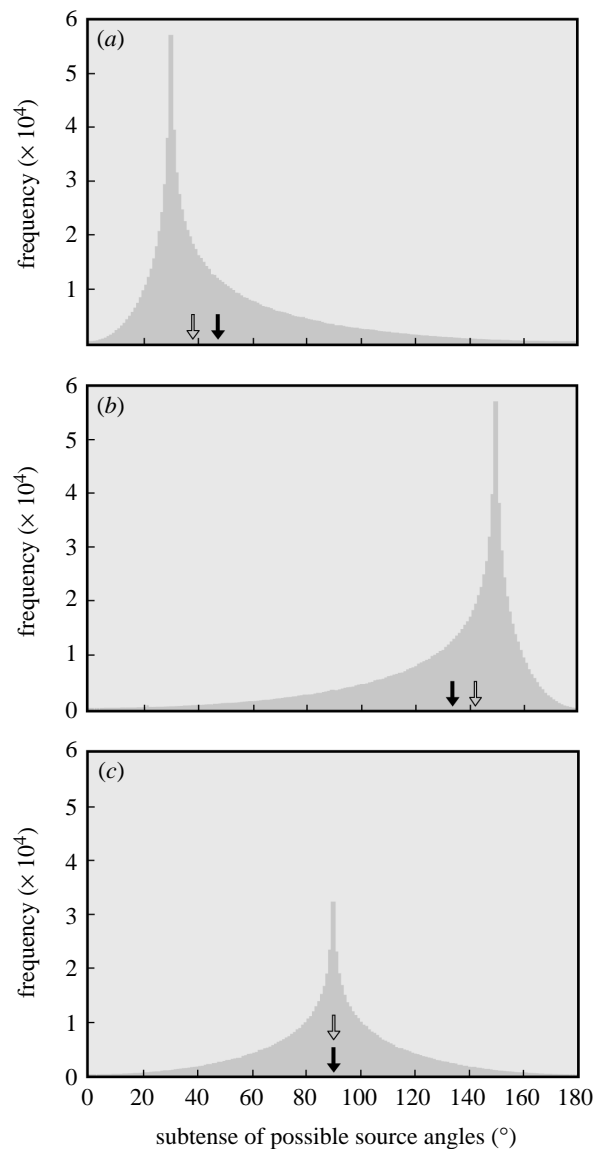


Figure 8. Frequency distribution of the possible source subtenses of representative angles projected onto a plane. (*a*) The frequency distribution of the possible source sub-tenses of a representative acute angle (30°) projection (in 1° increments). The solid arrow indicates the mean, and the open arrow indicates the median. (*b*) The frequency distribution of the possible source subtenses of a representative obtuse angle (150°) projection. (*c*) The frequency distribution of the possible source subtenses of a right-angle projection. (After Nundy *et al.* 2000.)

the most frequently occurring sources of acute angle projections are angles larger than the subtense of the projected stimulus (figure 8*a*). Conversely, the sources of obtuse angle projections will, by the nature of projective geometry, typically have been generated by sources that are somewhat smaller than the projected angle (figure 8*b*). Right-angle projections and straight lines, however, will have been generated by sources that on average have the subtense of the object itself (figure 8*c*). The visual system should, if the theory is correct, generate percepts that incorporate these statistical facts of projective geometry, which have necessarily determined the way visual stimuli generated by angular objects have
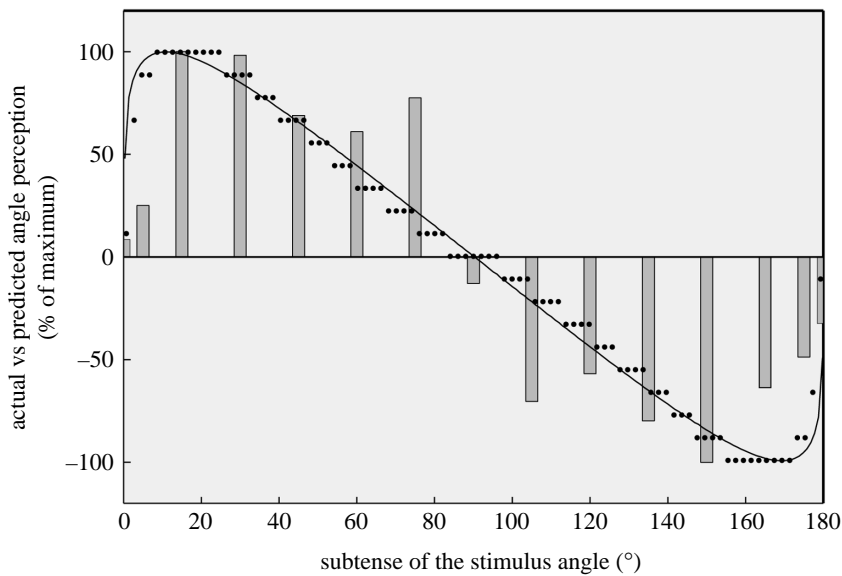
Figure 9. Comparison of the perceptions of the subtenses of angular stimuli and the perceptions predicted by the frequency distributions of the possible sources of the stimuli (calculated in the same way as for the examples in figure 8). Bars show the mean perceptual performance data from three tests of angle perception, normalized for comparison with the functions predicted by the geometrical analysis. The solid line indicates the perceptions predicted by the means of the source frequency distributions for each angle; the dots indicate the perceptions predicted by the medians of the source frequency distributions. (After Nundy *et al.* 2000.)

been experienced throughout human history (and have therefore shaped the circuitry underlying these perceptions).

Although the difference between perceived angles and the actual subtense of their sources is only a few degrees, by careful testing it is possible to plot the direction and magnitude of an observer's perceptions in comparison with the angle projected on the retina. Whereas this analysis is limited in several ways (see Nundy *et al.* 2000), the perceptions predicted by the distribution of the possible sources of the stimuli accord surprisingly well with what people actually see (figure 9).

As expected on this basis, depicting a projected angle such that it is more consistent with one real-world source than another changes the perceived subtense accordingly, often quite strikingly. Thus, whereas all the angles printed in figure 10 subtend 92.5°, the stimulus is consistent with each projection having been generated by angular objects having different subtenses and orientations. As in the domains of brightness and colour, the perceptions of the identical stimuli vary according to their empirical significance.

## 5. THE RELATIONSHIP BETWEEN OBJECT MOTION AND PERCEIVED MOTION

The final aspect of visual experience that we consider here is motion perception. In objective terms, motion is the continuous displacement of an object within a spatial frame of reference; as such, motion is fully described by physical measurements. In subjective terms, however, the corresponding sensation of motion is not so easily defined or understood. Because the real-world displacement of an object is conveyed to an observer by a changing projection on the retinal surface, the velocity field that uniquely defines motion in physical terms is ambiguous with respect to the possible causes of the changing retinal image: an infinite number of physical displacements can generate the same stimulus sequence (Helmholtz 1924; Wallach 1935) (figure 11a). This ambiguity presents a fundamental problem, namely how, in the face of such uncertainty, the brain generates quite definite percepts that

usually (but not always) allow the observer to deal successfully with the real-world cause of the retinal stimulus.

A solution to this problem is to accumulate experience interacting with moving objects, such that motion percepts gradually come to accord with the statistics of the possible displacements underlying the stimulus. In figure 11b, for example, the physical correspondences of the points along the line in any two sequential images cannot be determined directly (because some points come into view, others disappear, and still others could represent deformation as the line expands or contracts). Although it is generally believed that physical correspondence is the basic requirement for a stimulus to be correctly perceived as moving (both motion and stereoscopic vision have long been assumed to be correspondence problems), we have proposed that motion perception is derived empirically from the complete set of possible correspondences and differences between any sequential images (Yang *et al.* 2001). This set entails (i) the identity of some, but not necessarily all, the elements in any two sequential images; (ii) the possible appearance of some, but not necessarily all, of the elements in the second image not present in the first; (iii) the possible disappearance of some, but not necessarily all, of the elements in the first image compared with the second; and (iv) the possible deformation of the stimulus during the interval between the two images (figure 11b). Since the relative contribution of these correspondences and differences to the physical displacements underlying the stimulus cannot be determined by inspection of the stimulus *per se*, we have again argued that the problem posed by this inevitable uncertainty is solved empirically by generating motion percepts based on past experience of what such stimuli typically turned out to be.

Accordingly, we computed the probability distribution of the possible physical displacements underlying simple image sequences. In this way we could predict the percepts that subjects would be expected to see on a wholly empirical basis, and compare them with actual performance (figure 12). Having stated the fundamental ambiguities embedded in the correspondences and differences underlying the sequential images of any motion
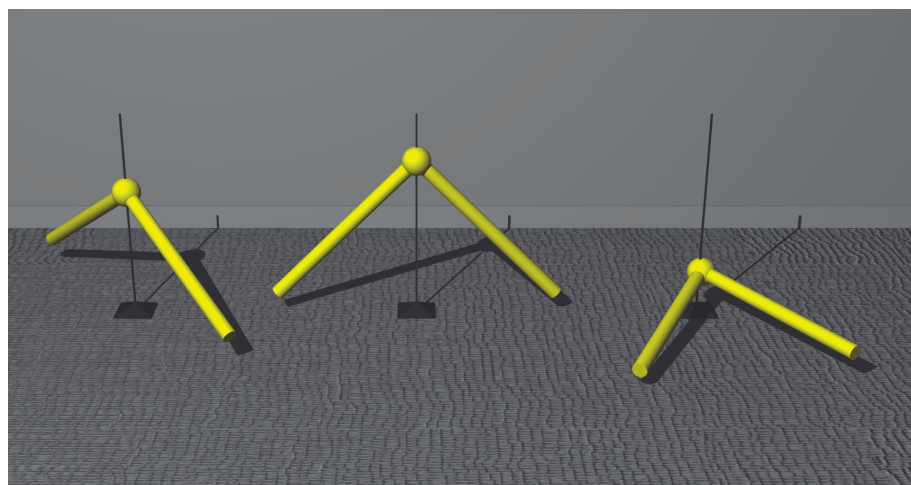
Figure 10. A combination of empirical factors can strongly influence the perception of angles. In this scene, the subtenses of the three angular objects are identical, each measuring 92.5° on the printed page. As a result of the variety of depicted cues about the attributes of the possible real-world sources, the subtenses of the three identical angles look quite different.

stimulus (see figure 11b), this approach entailed (i) describing quantitatively how these correspondences and differences in the image plane are generated in the case of a straight line moving in the fronto-parallel plane; (ii) using this information to derive a set of probability distributions of the possible real-world displacements underlying the stimulus; (iii) deriving a principle for combining these probability distributions based on the statistical structure of the underlying events in the stimulus; and (iv) devising a procedure for predicting motion perception based on this joint probability (see Yang *et al*. 2001).

The way subjects perceived the direction and speed of a line moving in any of a variety of apertures in the fronto-parallel plane could be accounted for on this basis. For instance, the perceived direction of motion of a line orientated at 45° and translating horizontally from left to right should change when constrained by a circular aperture (figure 12a), now being seen moving downward and to the right (a shift in the direction of 45°) (see figure 11a). Moreover, the perceived speed of the line in this circumstance should be *ca*. 30% less than the perceived speed without the aperture in place. The predicted effects of this and a variety of other apertures correspond remarkably well with what subjects in fact see (cf. figure 12b,c).

A merit of this concept of motion perception is that this wholly empirical framework accounts for the full range of phenomena first described by Wallach over 65 years ago (Wallach 1935), whereas other motion theories do not (see Yang *et al*. (2001) for a detailed discussion).

## 6. THE BIOLOGICAL RATIONALE FOR A RADICALLY EMPIRICAL STRATEGY OF VISION

The breadth of evidence regarding the perception of luminance (brightness), spectral differences (colour), projected lines (orientation) and displacement (motion) makes a general case that the visual system generates percepts based entirely on the historical significance of proximal stimuli. Predicating percepts on the probability distributions of what the same or similar stimuli have turned out to be in the past is different in both concept and consequence from seeing stimuli on the basis of the present properties of the distal or proximal stimulus *per*

*se*, or even the most probable source of the stimulus (ideas which have all been, and continue to be, popular in various quarters). This hypothesis about vision is also fundamentally different from the assumption made by many physiologists and others interested in visual 'coding', that visual perception is based on a 'readout' of the firing rates of 'neuronal detectors' of the various attributes of the stimuli that fall on the retina.

The underlying rationale for the evolution of a wholly empirical strategy of vision is that the visual system must generate successful responses to visual stimuli that are unknowable by any direct means. The most straightforward, and perhaps the only biologically feasible way of solving this problem is to depend on the operational success or failure of visually guided behaviour, taking full advantage in this process of the experience of both the species and the individual. Even animals that cannot afford to harbour the rich patterns of connectivity that characterize the cortices of 'higher' mammals would nonetheless be well served by having visual reflexes occur on this same empirical basis, for exactly the same reasons our own visual systems have evidently evolved according to this strategy. By using the full gamut of past experience, the observer—human or otherwise—will in all circumstances generate the response that, on a purely operational basis, is most likely to succeed in the face of the always uncertain sources of the stimulus on the retina.

This way of generating percepts does not mean that the observer has any information about illuminants, reflectances or transmittances *per se*, or that it would be of any use to have such information. Nor do the visual percepts 'represent' any aspect of the objects underlying the stimulus. Rather, the perceptions elicited by any pattern of retinal activity accord with the history of human interactions with the same or a similar stimulus. In consequence, the perceptual qualities experienced by observers covary with the physical laws that govern the relationships between objects and the stimuli they generate, and not with the properties of objects as such (thus explaining the wealth of familiar 'visual illusions'). This general strategy provides the short answer to the implied question in the title of this article: seeing things the way we do is the signature of (and the evidence for) this wholly empirical strategy of vision.
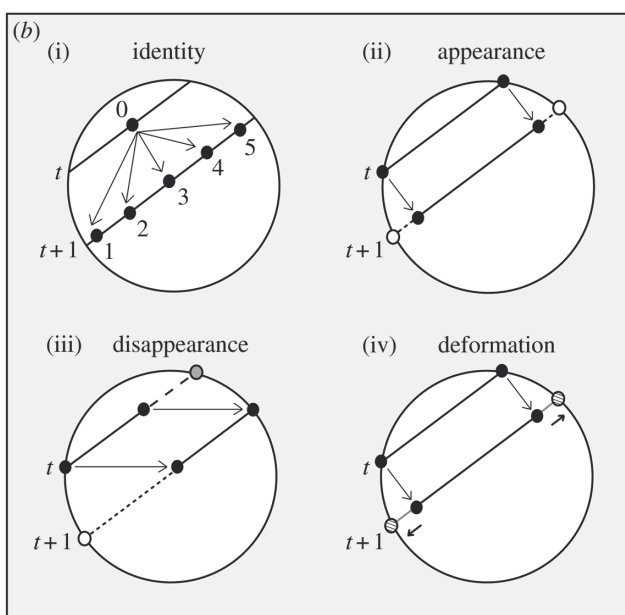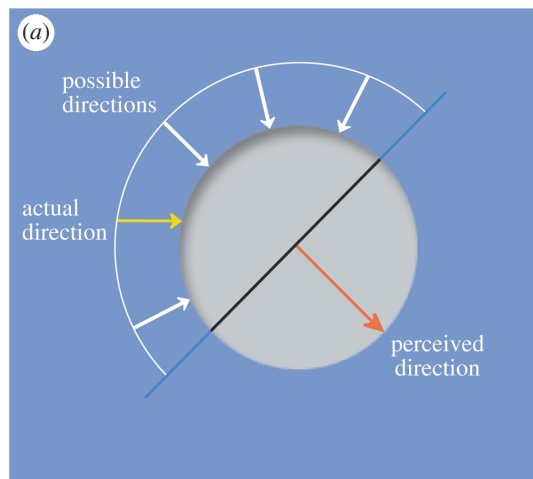
Figure 11. The ambiguity of stimuli that elicit perceptions of motion. (*a*) When a line moves horizontally from left to right in the absence of an occluding aperture, the perceived direction of motion is the same as the actual direction of motion (indicated by the yellow arrow). When the aperture is in place, however, the perceived direction of motion is downward and to the right (indicated by the red arrow), a shift of 45°. As explained in the text, the reason for this altered perception is the empirical strategy used by the visual system to contend with the inherent ambiguity of the stimulus (the line could be moving at different speeds in any of the directions indicated by the white arrows). (*b*) Diagram indicating the factors underlying the complete set of correspondences and differences that describes the ambiguity of the stimulus in (*a*) in any two sequential images of a line as it moves across an aperture. Computation of the joint probability that determines the perception of motion is based on this set. (After Yang *et al.* 2001.)

## 7. MECHANISMS THAT COULD INSTANTIATE VISUAL EXPERIENCE

If this general conception of vision is correct, the underlying circuitry of the primary visual pathway must somehow come to elicit patterns of central neural activity

that correspond to the empirical significance of the proximal stimulus (i.e. what the stimulus has typically turned out to be in the past). To accomplish this feat, the end product of visual system development in each member of the species must incorporate the experience of the species as a whole, as well as the experience accumulated by the individual. Whatever the details of these processes in phylogeny and ontogeny, the neuronal linkages between the retinal effects of a particular proximal stimulus and the central activity that corresponds to the empirical significance of the scenes that have typically given rise to that retinal activity must be strengthened, and those that do not must be weakened. Consequently, when the same or a similar stimulus occurs in the future, the probability of activating the neural circuitry that corresponds to the scene typically underlying the stimulus will be increased.

What, then, are the mechanisms that could instantiate this strategy in the circuitry of the visual system? Although the answer is not known, at least a broad explanation can be suggested. In phylogeny, the mechanism of the feedback that promotes the increasing prevalence of the pertinent networks in the population is presumably the greater success of progeny who, by virtue of random variation, happen to possess associative circuitry that generates visually guided responses (and corresponding percepts) that enable the observer to respond more effectively to visual stimuli than competing individuals. Statistically speaking, these members of the cohort will survive preferentially, leading to the inheritance of the relevant network connectivity in subsequent generations.

In ontogeny, the mechanism of experiential feedback is presumably the conjoint activity of the relevant network elements. In accord with the known rules of cellular and molecular neurobiology, such activity will strengthen the neural connections between the neurons in the primary visual pathway made active by the proximal stimulus, and the neurons in extrastriate (and other) cortical areas pertinent to the generation of successful visually guided actions in response to a particular type of stimulus (see Purves & Lichtman 1985; Purves 1988, 1994; Purves *et al.* (2000*c*) for a review of many of these cellular and molecular processes). Since developing individuals who succeed in visual tasks more often than their competitors will likewise have a reproductive advantage over individuals less frequently successful, the relevant mechanisms of associational plasticity will also be transmitted to each future generation. The parallel operation of these phylogenetic and ontogenetic mechanisms would ensure that the percept generated by any particular proximal stimulus will, over time, incorporate the influence of the possible underlying scenes in proportion to the frequency with which they have turned out to be guides to the generation of successful visual behaviour in the past.

Finally, it is important to emphasize that these mechanisms are not predicated on the symbolic representation of image properties, but simply on operational success derived from trial and error. As a result, the underlying circuitry of the visual system is more likely to have the qualities of a connectionist network determined by feedback than the qualities of a program whose operation is defined *a priori*.
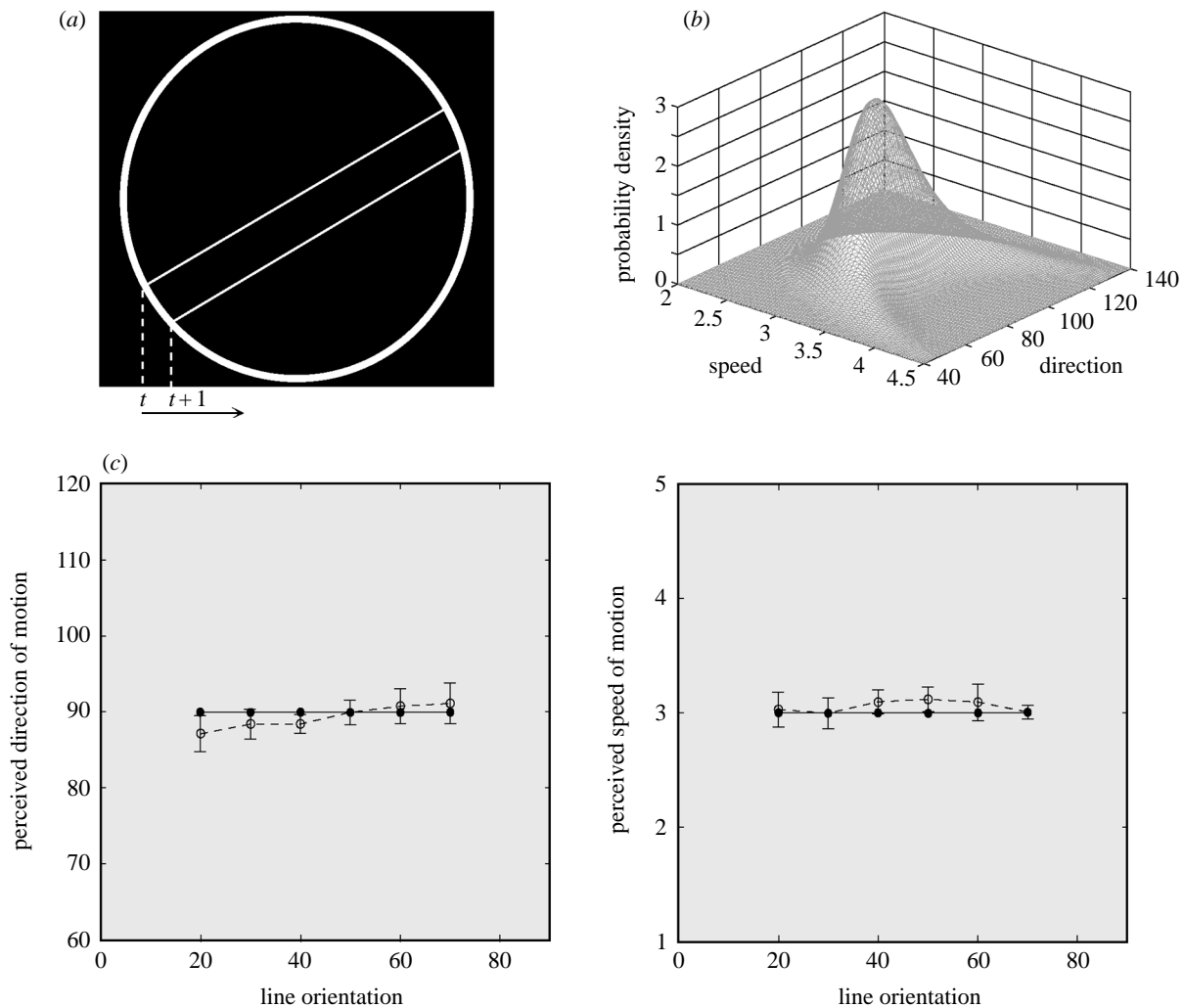
Figure 12. Comparison of the perceived motion of a line moving across a circular aperture and the perceptions predicted by probability of the possible line translations underlying such stimuli. (*a*) A representative stimulus shown at two different times (*t* and *t* + 1) moving in the direction indicated by the arrow. (*b*) The probability distribution of possible line translations underlying the stimulus in (*a*). (*c*) Comparison of the direction (left-hand graph) and speed (right-hand graph) of the motion perceived (dotted line) and the perceptions predicted by probability distribution of the line translations at a variety of different orientations and speeds (solid line) (vertical bars show standard deviation of the performance of three subjects). The close agreement between the observed and predicted behaviour in various contexts is consistent with the hypothesis that the perceived direction and speed of moving objects is generated on a wholly empirical basis. (After Yang *et al*. 2001.)

## 8. CONCLUDING REMARKS

The range of evidence we have described here indicates that percepts correspond to, and are generated by, the historical significance of proximal stimuli. The ability to predict the subjective experiences of luminance (brightness), spectral differences (colour), line orientation (form) and the translocation of objects (motion) on this basis is consistent with the conclusion that the patterns of neural activity that give rise to visual percepts that are essentially reflex responses (i.e. patterns of neuronal activity triggered automatically by appropriate stimuli). These patterns of connectivity have evidently been moulded during phylogeny and ontogeny by the frequency of occurrence of the physical sources and relationships that have been discovered to underlie visual stimuli during interactions of the observer with real-world objects.

If visual percepts are indeed generated on this wholly empirical basis, the notion that the visual system normally represents objects veridically, and that visual illusions are anomalies in which accurate representation fails, is not tenable. Since the measured properties of objects are not perceived, they cannot be misperceived. By the same token, the idea that visual neurons are 'tuned' to extract specific 'features' of the visual scene and that their activation 'represents' these features is likewise illogical, since the gist of these long-held concepts is also based on the representation of veridical information by processing the retinal 'image'. The enormous advantage of generating vision on this wholly empirical basis is solving the problem presented by the unknowability of the sources of visual stimuli by any other means.

## REFERENCES

Adelson, E. H. 1993 Perceptual organization and the judgment of brightness. *Science* **262**, 2042–2044.

Agostini, T. & Proffitt, D. R. 1993 Perceptual orgnanization evokes simultaneous lightness contrast. *Perception* **22**, 263–272.

Barlow, H. B. 1953 Summation and inhibition in the frog's retina. *J. Physiol.* (*Lond.*) **119**, 69–88.

Benary, W. 1924 Beobachtungen zu einem Experiment über Helligkeitskontrast. *Psychol. Forsch.* **5**, 131–142.

Berkeley, G. 1709 *A new theory of vision.* (Everyman's Library.) London: J. M. Dent.

Biehler, W. 1896 Beitrage zur Lehre vom Augenmass für Winkel. PhD thesis, University of Freiburg, Germany.

Brainard, D. H. & Wandell, B. 1986 Analysis of the retinex theory of color vision. *J. Opt. Soc. Am.* **3**, 1651–1661.

Brown, R. O. & MacLeod, D. I. 1997 Color appearance depends on the variance of the surround colors. *Curr. Biol.* **7**, 844–849.

Carpenter, R. H. S. & Blakemore, C. 1973 Interactions between orientations in human vision. *Exp. Brain Res.* **18**, 287–303.

Chichilnisky, E.-J. & Wandell, B. A. 1995 Photoreceptor sensitivity changes explain color appearance shifts induced by large uniform backgrounds in dichoptic matching. *Vision Res.* **35**, 239–254.

Chubb, C., Sperling, G. & Solomon, J. A. 1989 Texture interactions determine perceived contrast. *Proc. Natl Acad. Sci. USA* **86**, 9631–9635.

Coppola, D. M., Purves, H., McCoy, A. & Purves, D. 1998 The distribution of oriented contours in the real world. *Proc. Natl Acad. Sci. USA* **95**, 4002–4006.

Cornsweet, T. N. 1970 *Visual perception.* New York: Academic Press.

Craik, K. J. W. 1966 *The nature of psychology: a selection of papers, essays and other writings by the late Kenneth J. W. Craik* (ed. S. L. Sherwood). Cambridge University Press.

Creutzfeldt, B., Lange-Malecki, B. & Dreyer, E. 1990 Chromatic induction and brightness contrast: a relativistic color model. *J. Opt. Soc. Am.* **7**, 1644–1653.

Creutzfeldt, B., Crook, J., Kastner, S., Li, C.-Y. & Pei, X. 1991 The neurophysiological correlates of colour and brightness in lateral geniculate neurons. *Exp. Brain Res.* **87**, 3–45.

D'Zmura, M. & Lennie, P. 1986 Mechanisms of color constancy. *J. Opt. Soc. Am.* **3**, 1662–1672.

Gilchrist, A. L. 1977 Perceived lightness depends on perceived spatial arrangement. *Science* **195**, 185–187.

Helmholtz, H. L. F. von 1924 *Helmholtz's treatise on physiological optics*, vols I–III. Translated from the third German edition (1909) by J. P. C. Southall. Rochester, NY: The Optical Society of America.

Hering, E. 1864 *Outlines of a theory of the light sense.* Translated by L. M. Hurvich & D. Jameson. Cambridge, MA: Harvard University Press, 1964.

Howard, R. B. 1971 Neurophysiological models of figural aftereffects and visual illusions. *Psychonom. Monogr. Suppl.* **4**, 57–72.

Jameson, D. & Hurvich, L. M. 1989 Essay concerning color constancy. *A. Rev. Psychol.* **40**, 1–22.

Kuffler, S. W. 1953 Discharge patterns and functional organization of mammalian retina. *J. Neurophysiol.* **16**, 37–68.

Land, E. H. 1986 Recent advances in retinex theory. *Vision Res.* **26**, 7–21.

Lotto, R. B. & Purves, D. 1999 The effects of color on brightness. *Nature Neurosci.* **2**, 1010–1014.

Lotto, R. B. & Purves, D. 2000 An empirical explanation of color contrast. *Proc. Natl Acad. Sci. USA* **97**, 12 834–12 839.

Lotto, R. B. & Purves, D. 2001 An empirical explanation of the Chubb illusion. *J. Cogn. Neurosci.* (In the press.)

Lotto, R. B., Williams, S. M. & Purves, D. 1999a An empirical basis for Mach bands. *Proc. Natl Acad. Sci. USA* **96**, 5239–5244.

Lotto, R. B., Williams, S. M. & Purves, D. 1999b Mach bands as empirically derived associations. *Proc. Natl Acad. Sci. USA* **96**, 5245–5250.

Mach, E. 1865 Über die Wirkung der raümlichen Verthelung des Lichtreizes auf die Netzhaut. I. Sber. math.-nat. *Classe kaiserlichen Akad. Wiss.* **52**, 303–322.

Mach, E. 1914 *The analysis of sensation and the relation of the physical to the psychical* (ed. C. M. Williams), translated and revised by S. Waterlow. Chicago, IL and London: Open Court Publishing Co.

Nundy, S., Lotto, R. B., Coppola, D., Shimpi, A. & Purves, D. 2000 Why are angles misperceived? *Proc. Natl Acad. Sci. USA* **97**, 5592–5597.

O'Brien, V. 1959 Contrast by contour-enhancement. *Am. J. Psychol.* **72**, 299–300.

Purves, D. 1988 *Body and brain: a trophic theory of neural connections.* Cambridge, MA: Harvard University Press.

Purves, D. 1994 *Neural activity and the growth of the brain.* Cambridge University Press.

Purves, D. & Lichtman, J. W. 1985 *Principles of neural development.* Sunderland, MA: Sinauer Associates.

Purves, D., Shimpi, A. & Lotto, R. B. 1999 An empirical explanation of the Cornsweet effect. *J. Neurosci.* **19**, 8542–8551.

Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S. & McNamara, J. O. 2000 *Neuroscience*, 2nd edn. Sunderland, MA: Sinauer Associates.

Ratliff, F. 1965 *Mach bands: quantitative studies on neural networks in the retina.* San Francisco, CA: Holden-Day, Inc.

Todorovic, D. 1997 Lightness and junctions. *Perception* **26**, 379–394.

Tyler, C. W. & Nakayama, K. 1984 Size interactions in the perception of orientation. In *Sensory experience, adaptation and perception* (ed. L. Spillman & B. R. Wooton), pp. 529–546. Hillsdale, NJ: Lawrence Erlbaum Associates.

Von Bekesy, G. 1967 *Sensory inhibition.* Princeton University Press.

Von Kries, J. 1905 Die Gesichtsempfindugen. In *Handbuch der Physiologie des Meschen* (ed. W. Nagel), pp. 109–282. Braunschweig, Germany: Viewveg.

Wallach, H. 1935 Über visuell wahrgenommene Bewegungsrichtung. *Psychologische Forscheung* **20**, 325–380. Translation by S. Wuerger, R. Shapley and N. Rubin 1996 On the visually perceived direction of motion by Hans Wallach: 60 years later. *Perception* **25**, 1317–1367.

Walsh, V. 1995 Adapting to change. *Curr. Biol.* **5**, 703–705.

Ware, C. & Cowan, W. B. 1982 Changes in perceived color due to chromatic interactions. *Vision Res.* **22**, 1353–1362.

Webster, M. A. & Mollon, J. D. 1995 Colour constancy influenced by contrast adaptation. *Nature* **373**, 694–698.

Wertheimer, M. 1912 Laws of organization in perceptual forms. In *A sourcebook of gestalt psychology* (translated and edited by W. D. Ellis), pp. 71–88. New York: Humanities Press, 1950.

White, M. 1979 A new effect of pattern on perceived lightness. *Perception* **8**, 413–416.

Williams, S. M., McCoy, A. N. & Purves, D. 1998a The influence of depicted illumination on perceived brightness. *Proc. Natl Acad. Sci. USA* **95**, 13 296–13 300.

Williams, S. M., McCoy, A. N. & Purves, D. 1998b An empirical explanation of brightness. *Proc. Natl Acad. Sci. USA* **95**, 13 301–13 306.

Wundt, W. 1902 *Beitrage zue theorie der sinnewahrnelmung.* Translated by C. H. Judd *Outlines of psychology*, p. 125. New York: G. E. Stechert.

Yang, Z., Shimpi, A. & Purves, D. 2001 A wholly empirical explanation of motion perception. (Submitted.)

Yarbus, A. L. 1967 *Eye movements and vision.* Translated by B. Haigh and L. A. Riggs. New York: Plenum Press.