Color Vision and the Four-Color-Map Problem

Dale Purves, Beau Lotto, and Thomas Polger
Duke University Medical Center

Abstract

■ Four different colors are needed to make maps that avoid adjacent countries of the same color. Because the retinal image is two dimensional, like a map, four dimensions of chromatic experience would also be needed to optimally distinguish regions returning spectrally different light to the eye. We therefore suggest that the organization of human color vision according to four-color classes (reds, greens, blues, and yellows) has arisen as a solution to this logical requirement in topology. ■

An abiding puzzle in human color vision is why chromatic experience is predicated on four classes of color, each defined by a unique hue. Thus, a particular red, green, blue and yellow is seen as being entirely free of any other color, whereas all other hues are perceived as mixtures of these four (Figure 1) (Hurvich, 1981; Hering, 1964; Evans, 1948). Although it is well established that the initial processing of spectral information depends on the different absorption characteristics of three distinct receptor types (short-, medium-, and long-wavelength cones), and that the central processing of this information involves color-opponent mechanisms (Kaiser & Boynton, 1996; Wandell, 1995; Hurvich, 1981; Hurvich & Jameson, 1957), no clear rationale for this organization of human color vision has emerged. It is generally supposed that the perceptual quality of these four-color categories and their unique members is an incidental consequence of the color-opponent channels that Hering first proposed more than a century ago (Hering, 1964). Perhaps as a result, relatively little attention has been paid to understanding why human color experience is organized in this particular way. Here, we suggest that humans perceive four-color categories defined by unique hues because the visual system has evolved to solve a fundamental problem in topology, namely ensuring that no two areas separated by a common boundary in a two-dimensional array will appear the same if they are actually different. In topology, this issue is generally referred to as the "four-color-map problem" (Figure 2).

THE FOUR-COLOR-MAP CONJECTURE

Although cartographers had long known that four colors are needed to make unambiguous maps, the four-color-map problem was first posed as a logical challenge in 1852. In that year, a student at University College London asked Augustus de Morgan, a professor of logic and mathematics, if he knew a proof for the apparent sufficiency of four colors to illustrate any map without having adjacent regions of the same color. After a lapse of some years, a friend of de Morgan's stated the problem formally as a query in the Proceedings of the London Mathematical Society (Cayley, 1878). Proving the conjecture that "four colors are sufficient to color any map drawn in a plane or on a sphere so that no two regions with a common boundary [other than a point] are colored with the same color" was quickly taken up by scholars around the world. Indeed, most mathematicians during the subsequent century are said to have devoted at least some thought to the solution of this conundrum, and many a great deal (Saaty & Kainen, 1986; Appel & Haken, 1977; Ore, 1967). Its apparent simplicity notwithstanding, the four-color conjecture resisted efforts at a formal proof until 1976, when it was finally solved using a computer algorithm that required more than 200 pages to publish (Appel & Haken, 1976). Quite apart from the nature of this proof, the four-color-map problem raises the possibility that the four dimensions of human color experience (red, green, blue, and yellow) may have arisen as a means of dealing with this basic requirement in topology.

RELATIONSHIP OF THE FOUR-COLOR-MAP PROBLEM AND COLOR VISION

Visual perception is necessarily based on a two-dimensional topography. Despite the three-dimensional provenance of most visual stimuli, the scenes we see are derived from two-dimensional projections focused on the retina. Because objects are defined visually by their contrast with other objects, the ability to distinguish one object from another—presumably the central purpose
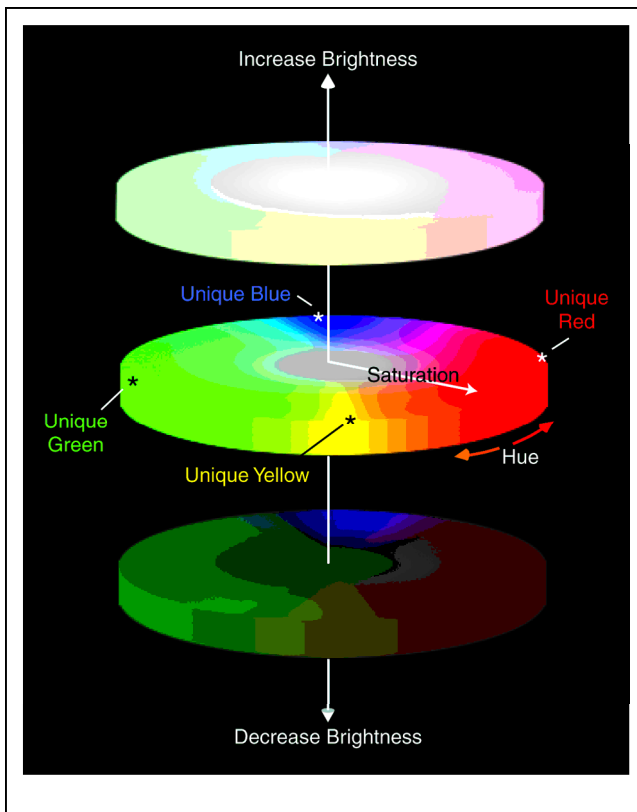
Figure 1. Human color experience is generally described in terms of hue, saturation and brightness. Hue is the quality of the color as such, saturation the degree to which the color differs from a neutral gray, and brightness (or lightness) the perceived intensity of the color. Four primary colors—red, green, blue, and yellow—are characterized by a unique color percept (asterisks), that is, a color experience that cannot be seen or imagined as a mixture of any other colors. Secondary color groupings, such as purples, oranges, cyans, and yellow-greens, are perceptual mixtures of two of the four primary hues, and can always be perceptually reduced to the relative contributions of these four underlying components.

of vision—depends on defining contrast boundaries within this topography. To the extent that object boundaries are not effectively defined, scenes remain ambiguous, thus increasing the chance of responding to a visual stimulus with inappropriate (and potentially detrimental) behavior.

All sighted mammals (some species of moles and bats are effectively blind) can readily distinguish boundaries demarcated by differences in surface luminance. However, mammals or other animals with color vision are also able to distinguish boundaries defined by differences in the spectral distribution of the light stimulus. It therefore follows that a visual system capable of identifying boundaries that entail spectral differences will be more effective in distinguishing objects than one that cannot (see, e.g., Mollon, 1991; Frome, Buck, & Boynton, 1981).

Given this general rationale for the evolution of color vision, the logic of the four-color-map problem implies that humans or other animals would wish to instantiate this sensory modality in a manner that avoids the problem map makers would have if they were limited to a palette of less than four colors. Unlike the cartographer, however, the visual system cannot arbitrarily assign colors to different regions of the visual world to avoid ambiguity; on the contrary, it must deal with the spectral returns from objects, whatever they happen to be (Figure 3A). How, then, can the human visual system, or any other, solve the demands of the four-color-map problem for spectral returns?

AN EXPLANATION IN PRINCIPLE

A variety of psychophysical evidence—most famously the demonstrations presented by Edwin Land in the
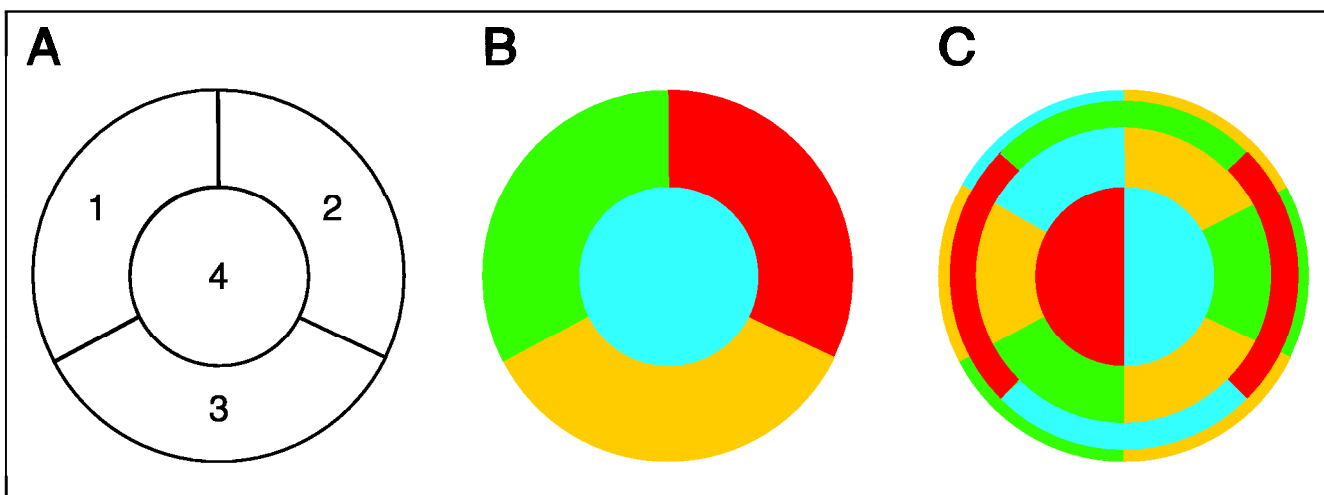

Figure 2. The four-color-map problem. Four regions (numbered 1 through 4) in a two-dimensional topology (A) cannot be unambiguously distinguished using fewer than four colors (B). The challenge in the classical four-color-map problem was to prove that four colors are sufficient to disambiguate any arbitrarily complex two-dimensional map, such as the example in (C). Although the four-color requirement seems empirically obvious, it took more than a century to show that the four-color-map conjecture is, in fact, correct.

late 1950s (Land, 1959a; Land, 1959b)—has shown that the visual system generates color percepts by comparing spectral returns throughout a scene (Kaiser &

Boynton, 1996; Wandell, 1995; Land, 1986; Hurvich, 1981; Evans, 1948). Applying the lesson of the four-color-map problem, this process, whatever its particulars, must involve a comparison of spectral returns in at least four different dimensions. If the number of dimensions for comparison were fewer than four, then, as indicated in Figure 2, some abutting territories would be conflated. An intriguing possibility, therefore, is that the four primary-color categories we experience—reds, greens, blues, and yellows—are the sensations produced in the course of these four comparisons. In this conception, the unique hue that defines each category (that is, the particular color in each group that is perceived as pure and unmixed) represents the maximum perceptual ''distance'' that human visual experience can proceed along one of the four necessary dimensions of comparison (Figure 3B) (the unique hues of red, green, blue, and yellow defining these dimensions in much the same way that the terminus defines a subway line).

Hering (1964) proposed that black and white also be considered polar opposites similar to the red-green or blue-yellow axes. Whatever the merits of this suggestion, neutral shades, ranging from black to white, can be ignored for present purposes, since these percepts do not entail spectral differences, but only differences in luminance (which may explain the diminished ability of humans to discriminate boundaries using luminance information alone; see Mollon, 1991; Frome et al., 1981). Thus the four-color-map problem is not resolved by the visual systems of animals that lack color vision, or by our own visual systems when viewing scenes that are rendered only in shades of gray (as in an old movie). Nor would black and white in conjunction with a single-color axis suffice, since chromatic and achromatic experience are effectively different domains.

WHY NOT ANY FOUR COLORS?

Even if one accepts the conclusion that perceptions of grays cannot solve the four-color-map problem for spectral differences, it is reasonable to ask why any four colors could not do the job. After all, we see hundreds of colors, all of which are relevant to disambiguating spectral returns in the two-dimensional retinal topogra-
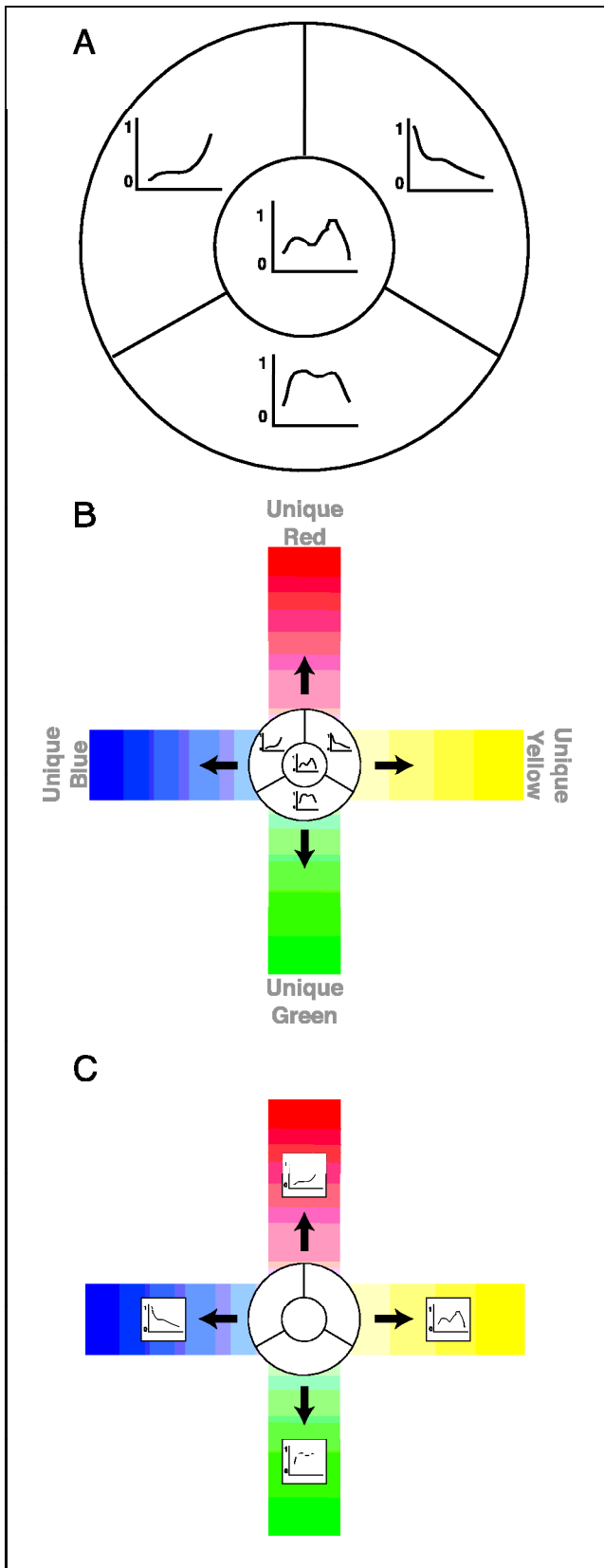
Figure 3. The topological requirements made plain in the four-color-map problem imply that four dimensions of color comparison are needed to unambiguously distinguish spectral returns. (A) Diagram of the different spectral returns of four adjacent regions in a visual stimulus, using the same schema as in Figure 2. (B) By analogy with the four-color-map problem in cartography, comparisons in four different dimensions would be required to optimally disambiguate objects with different spectral returns in any given scene. (C) As a consequence of these four-way comparative processes, the spectral return from any given region is experienced as relatively more reddish, greenish, bluish, or yellowish than the return from any other region in the scene.

phy that initiates vision. If the cartographer can solve the problem with *any* four colors, or indeed with any four distinguishing symbols (four different textures would do as well in cartography, although at some expense to esthetics), why can the visual system not do so?

First, color vision is restricted to the perception of spectral differences, which, as already noted, rules out the use of other visual qualities, such as texture to solve the four-color-map problem in the spectral domain. The goal of color vision is presumably to aid and abet whatever contributions are made to object discrimination by nonspectral differences arising from surface qualities. Second, the many colors we experience are all perceptual mixtures of red, green, blue, and yellow. It is not the spectral returns themselves, or even the resulting color percepts that optimally distinguish color boundaries, but rather the process of comparing spectral returns in four requisite dimensions of color space (Figure 3C).

Of course, the specific colors we experience as primary could have been otherwise, but if our argument with respect to the four-color-map problem is correct, the challenge of optimally distinguishing spectral boundaries could not have been met with less than four dimensions of color experience. If fewer than four comparisons were available, some ambiguity in the boundaries defined by different spectral returns would be inevitable, as is evident in the discussion of color deficiencies in the following section.

HUMAN COLOR DEFICIENCIES AND COLOR VISION IN OTHER SPECIES

Although many species have color vision, among mammals only humans, catarrhine primates, and some platyrrhine primates have a trichromatic color system, most other mammals being dichromats (Mollon, 1991; Neumeyer, 1991). Moreover, in some New World species, females are trichromatic and males dichromatic, whereas tetrachromatic vision is relatively common among birds and fishes (op. cit.). This diversity of color-vision mechanisms, and the relative rarity of trichromatic vision among mammals (it seems reasonable to assume that experiencing four-color categories by virtue of color-opponent neurons is linked to having three cone types), presents a problem for the argument here. If, for the reasons stated, solving the four-color-map problem is essential for maximally effective color vision, why have so many species failed to hit upon the solution that has evolved in humans and some other primates (that is, trichromatic vision)?

The answer may simply be that, whereas maximizing the information from spectral returns by comparison along four different dimensions of color experience is indeed essential to solve the topological problem outlined here, the value of this solution does not add so greatly to the efficacy of visually guided behavior as to

have stimulated the evolution of trichromacy in a wider range of species. In other words, the cost/benefit ratio of solving this problem may be relatively high. The practical consequences of human color deficiencies tend to support this interpretation. The most common form of color deficiency in humans (excluding the minor abnormalities found in anomalous trichromats) arises from abnormalities in one of the three cone pigments (Nathans, Piantanida, Eddy, Shows, & Hogness, 1986). Human dichromats (so called because they require two instead of three variable lights of independent hue to match any spectral stimulus) are deficient either in distinguishing blues and yellows, or, more commonly, reds and greens (depending on which of the three cone pigments is affected). Such individuals (who make up about two to three percent of the male population in the United States) are capable of carrying out only two of the necessary four-color comparisons, and are thereby at a disadvantage in discriminating objects on the basis of spectral differences (Shepard & Cooper, 1992; Dalton, 1798). Nonetheless, human dichromats are not much impeded in carrying out the functions of daily life, and are excluded from only a few types of jobs.

With regard to animals that have more than three cone types (the mantis shrimp presently holds the record with 10 different photoreceptor types; see Neumeyer, 1991), it is simply not known what their color experience might be, and, therefore, how to consider them in the present argument.

CONCLUDING REMARKS

The hypothesis that the human color vision solves a fundamental problem in topology provides a novel way of thinking about an otherwise perplexing feature of color experience, namely why we see four categories of color, each defined by a unique hue.

REFERENCES

Appel, K., & Haken, W. (1976). Proof of 4-color theorem. *Discrete Mathematics*, *16*(2), 179–180.

Appel, K., & Haken, W. (1977). The solution of the four-color-map problem. *Scientific American*, *237*(4), 108–121.

Cayley, A. (1878). On the colouring of maps. *Proceedings of the London Mathematical Society*, *9*, 148.

Dalton, J. (1798). Extraordinary facts relating to the vision of colours: With observations (read in October, 1794). *Memoirs*

and *Proceedings - Manchester Literary and Philosophical Society*, 5, 28–45.

Evans, R. (1948). *An introduction to color* (pp. 77, 78, 107, 130, 151, 164, 166, 224, 231). New York, NY: Wiley.

Frome, F. S., Buck, S. L., & Boynton, R. M. (1981). Visibility of borders: Separate and combined effects of color differences, luminance contrast, and luminance level. *Journal of the Optical Society of America*, *71*, 145–150.

Hering, E. (1964). *Outlines of a theory of the light sense* (L. M. Hurvich & D. Jameson, Trans.). Cambridge, MA: Harvard Univ. Press.

Hurvich, L. (1981). *Color vision* (pp. 180–194). Sunderland, MA: Sinauer Associates.

Hurvich, L. M., & Jameson, D. (1957). An opponent-process theory of color vision. *Psychological Review*, *64*, 384–108.

Kaiser, P., & Boynton, R. (1996). *Human color vision*, 2nd edn. (pp. 250, 287, 353, 500, 507). Washington, DC: Optical Society of America.

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

Mollon, J. D. (1991). Uses and evolutionary origins of primate color vision. In J. R. Cronly-Dillon & R. L. Gregory (Eds.), *Evolution of the eye and visual system (Vision and visual dysfunction, Vol. 2)* (pp. 306–319). Boca Raton, FL: CRC Press.

Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B., & Hogness, D. S. (1986). Molecular genetics of inherited variation in human color vision. *Science, 232*, 203–210.

Neumeyer, C. (1991). Evolution of color vision. In: Uses and evolutionary origins of primate color vision. In J. R. Cronly-Dillon & R. L. Gregory (Eds.), *Evolution of the eye and visual system (Vision and visual dysfunction, Vol. 2)* (pp. 284–305). Boca Raton, FL: CRC Press.

Ore, O. (1967). *The four-color problem*. New York: Academic Press.

Saaty, T. L., & Kainen, P. C. (1986). *The four-color problem. Assaults and conquest*. New York: Dover Publications.

Shepard, R. N., & Cooper, L. A. (1992). Representation of colors in the blind, color-blind and normally sighted. *Psychological Science, 3*(2), 97–103.

Wandell, B. (1995). *Foundations of vision* (pp. 69, 287, 291, 305). Sunderland, MA: Sinauer Associates.