An Integrative Framework for the Appraisal of Coloration in Nature

Darrell J. Kemp,†,,* Marie E. Herberstein, Leo J. Fleishman, John A. Endler, Andrew T. D. Bennett, Adrian G. Dyer, Nathan S. Hart, Justin Marshall, and Martin J. Whiting

1. Department of Biological Sciences, Macquarie University, North Ryde, New South Wales 2109, Australia; 2. Department of Biological Sciences, Union College, Schenectady, New York 12308; 3. Centre for Integrative Ecology, School of Life and Environmental Sciences, Deakin University, Warrn Ponds, Victoria 3216, Australia; 4. College of Marine and Environmental Sciences, James Cook University, Townsville, Queensland 4811, Australia; 5. Department of Physiology, Monash University, Clayton, Victoria 3800, Australia; and School of Media and Communication, RMIT, Melbourne, Victoria 3001, Australia; 6. School of Animal Biology and The Oceans Institute, University of Western Australia, Crawley, Western Australia 6009, Australia; 7. Queensland Brain Institute, The University of Queensland, Brisbane, Queensland 4072, Australia

Submitted March 9, 2014; Accepted January 13, 2015; Electronically published April 16, 2015

Online enhancement: appendix.

Abstract: The world in color presents a dazzling dimension of phenotypic variation. Biological interest in this variation has burgeoned, due to both increased means for quantifying spectral information and heightened appreciation for how animals view the world differently than humans. Effective study of color traits is challenged by how to best quantify visual perception in nonhuman species. This requires consideration of at least visual physiology but ultimately also the neural processes underlying perception. Our knowledge of color perception is founded largely on the principles gained from human psychophysics that have proven generalizable based on comparative studies in select animal models. Appreciation of these principles, their empirical foundation, and the reasonable limits to their applicability is crucial to reaching informed conclusions in color research. In this article, we seek a common intellectual basis for the study of color in nature. We first discuss the key perceptual principles, namely, retinal photoreception, sensory channels, opponent processing, color constancy, and receptor noise. We then draw on this basis to inform an analytical framework driven by the research question in relation to identifiable viewers and visual tasks of interest. Consideration of the limits to perceptual inference guides two primary decisions: first, whether a sensory-based approach is necessary and justified and, second, whether the visual task refers to perceptual distance or discriminability. We outline informed approaches in each situation and discuss key challenges for future progress, focusing particularly on how animals perceive color. Given that animal behavior serves as both the basic unit of psychophysics and the ultimate driver of color ecology/evolution, behavioral data are critical to reconciling knowledge across the schools of color research.

Keywords: biophysics, neural processing, perception, optics, sensory ecology, vision, color signaling.

* Corresponding author; e-mail: darrell.kemp@mq.edu.au.

Introduction

Color is an exquisite natural phenomenon and an enduring source of inspiration for poets, artists, philosophers, and scientists. This allure has not escaped biologists, who have long sought to study color in many ecological and evolutionary contexts (Johnsen 2012). In recent times, growing appreciation that most animals perceive color differently to humans (Endler 1990; Bennett 1994) has created a new surge of interest. This has motivated widespread effort to quantify both color traits and their visual environments. Increased affordability and portability of spectroradiometers has assisted by placing the basic technology for color measurement within the reach of most researchers. Simultaneously, efforts to elucidate perception in nonhuman species have generated a range of analytical approaches (e.g., Vorobyev and Osorio 1998; Endler and Mielke 2005; Pike 2012a). These efforts draw variously on principles derived from human psychophysics that are known to operate similarly in limited animal models (e.g., honeybees; de Ibarra et al. 2014). Effective studies of color in nature require not only appreciation of these principles and how they have been derived but also how they factor in to the available color analyses and what assumptions apply. The need for an accessible intellectual basis at all levels of inquiry presents a fundamental challenge for the field.

Color traits are studied for many different objectives, such as understanding morphological adaptation (e.g., Stoddard and Prum 2008), visual orientation (e.g., Kelber 1999), communication (e.g., Arnold et al. 2002), and deception (e.g., Chiao et al. 2009), as well as for exploring processes such as speciation (e.g., Chamberlain et al. 2009) and mimicry (e.g., Jiggins et al. 2001). Precisely because these traits are assessed
by eyes and processed by brains, such studies are ultimately informed by the cognate research fields of vision and neurophysiology. We conceptualize the intellectual breadth of color research in terms of two schools of inquiry (which, more accurately, represent endpoints along a continuum). The first school—hereafter "top-down"—seeks to use color as a trait in tests of ecological and/or evolutionary hypotheses. Typical top-down studies may target particular species (e.g., mate choice studies; Kemp and Rutowski 2007), ecological guilds (e.g., predators and prey; Endler 1991; Heiling et al. 2003), or phylogenetic groups (Stoddard and Prum 2008; Maia et al. 2013). We envisage this school to include many researchers with newly acquired means to study color as a biological trait in diverse (often novel) species and ecological contexts. The second school—hereafter "bottom-up"—seeks to understand the proximate basis of color propagation, reception, and perception. This school encompasses disciplines such as visual anatomy and physiology (Hardie 1986; Stavenga 2010), neural processing and psychophysics (Kelber et al. 2003; Dyer et al. 2011; Dyer 2012), and molecular genetics (Hunt et al. 2009). Empirical work generally proceeds in model systems (e.g., honeybees and birds; Hart 2002; de Ibarra et al. 2014) but can extend to higher taxonomic levels (e.g., insects; Briscoe and Chittka 2001). Ultimately, bottom-up research delivers the intellectual basis for developing color analyses and perceptual models (e.g., Chittka 1992; Vorobyev and Osorio 1998; Endler and Mielke 2005; Pike 2012a), which are the tools for reaching conclusions in top-down studies.

We propose the top-down/bottom-up terminology as a simplified heuristic basis for addressing what we sense as an intellectual disconnect within the field of color research. Given the rapid expansion of researchers addressing top-down questions, this disconnect is most evident via a frequent lack of informed analytical choices and properly considered conclusions. We aim to redress this by first synthesizing the fundamentals of animal color perception and then placing these principles into the context of top-down research questions. We conclude by exploring the challenges for future empirical progress and for ensuring synergistic development across the schools of color research. We refer readers to table 1 for a glossary of terms and to recent published reviews for more detail on visual processing in model animals (e.g., Osorio and Vorobyev 2005, 2008; Bennett and Théry 2007; Hart and Hunt 2007; de Ibarra et al. 2014; Lunau 2014).

Fundamental Principles of Color Perception

Vision occurs via the detection of incident light propagated through an environment, reflected from and/or transmitted through a surface, and captured by an eye (Lythgoe 1979). Color vision refers to the ability to detect, discriminate, and analyze wavelength distributions of light (Lythgoe 1979; Wyszecki and Stiles 1982). Animals capable of distinguishing different visual stimuli based on their wavelength distributions independent of total intensity are said to possess color vision (for a more considered definition, see Kelber and Osorio 2010). Aside from yielding greater overall information (sensu Osorio et al. 2004), color vision enables the identification of surfaces and objects over a wide range of intensities and despite variable lighting conditions (Kelber et al. 2003). However, understanding and studying this sensory capacity in animals is inherently challenging (Bennett and Théry 2007; Kelber and Osorio 2010). This is because color is a perceptual experience, that is, a subjective property ultimately expressed in the brain of an individual (Cornsweet 1970; Lythgoe 1979).

Most of what we know about color perception is based on more than a century of detailed work in humans, involving the cooperation of conscious experimental subjects (Cornsweet 1970; Kaiser and Boynton 1996; Kelber et al. 2003). In the past several decades, scientists have succeeded in relating much of the detail of human perception to specific anatomical and physiological features of the visual system (e.g., Gegenfurtner et al. 1999). This has provided a guiding framework for the comparative investigation of visual perception in nonhuman animals. However, the complexities of color perception have been elucidated for very few nonhuman model systems, including primates (Osorio et al. 2004), goldfish (Neumeyer 1992; Gehres and Neumeyer 2007; Stojev et al. 2011), bees (von Helversen 1972; Backhaus 1991; Chittka and Menzel 1992; Giurfa et al. 1997; Dyer et al. 2011; Dyer 2012; de Ibarra et al. 2014), pigeons, and chickens (Bowmaker 1977; Bowmaker and Knowles 1977; Okano et al. 1992). The sum of this work offers two important conclusions. First, although there is great complexity (Osorio and Vorobyev 2005), all systems exhibit key perceptual features that relate in similar ways to the basic anatomy and physiology of visual systems. Second, because the studied species represent a diverse sample of the animal world, we can, with some degree of confidence, assume basic principles that apply broadly across color perception systems. An understanding of these principles, which we explore in detail below, is fundamental to any appraisal of animal coloration.

Retinal Photoreception

Visual processing begins with the capture of photons by light-sensitive organs (eyes), generally via photopigments expressed within dedicated photoreceptor cells. Color vision requires photoreceptors with at least two different classes of spectral sensitivity. Interestingly, visual pigments are almost universally comprised of opsin proteins, whose absorption properties are highly conserved (Dartnall 1953).
This means that knowledge of the peak absorption wavelength of a photopigment ($\lambda_{\text{max}}$) allows one to calculate its theoretical absorbance spectrum according to existing templates (e.g., Govardovskii et al. 2000; Stavenga 2010). The light reaching these photopigments—hence photoreceptor sensitivity—may be modified by spectral filtration or reflection elsewhere within the eye and by additional features such as photocell size and structure. These features have particular importance to color vision when they apply at the individual photoreceptor (i.e., intracellular) level, as with vertebrate oil droplets (Liebman and Granda 1975) and the screening pigments of arthropods (Arikawa et al. 1999). Intracellular filters appear to have evolved in many groups as an avenue for tuning spectral sensitivity and/or improving color discrimination (Cronin and Caldwell 2002; Vorobyev 2003; Hunt et al. 2009; Saarinen et al. 2012). If enough is known (or can be extrapolated) about such features, they can be incorporated in predictions of photoreceptor sensitivity and, hence, color vision. For more detail, see the appendix, available online.

Knowledge (or educated estimates) of photoreceptor spectral sensitivity is generally the minimum requirement for a sensory-based analysis of color. Such information can be gained via microspectrophotometry and/or electrophysiology (e.g., Salcedo et al. 1999) or by identifying molecular genetic sequences known to code for visual opsins pigments (see appendix). Published estimates of photoreceptor characteristics such as $\lambda_{\text{max}}$ continue to accumulate (see, e.g., Théry and Gomez 2010). Importantly, there is evidence for great evolutionary conservatism for some features, such as the number and sensitivity of photoreceptor classes in birds (Hart 2001; Hart and Hunt 2007), lizards (Loew et al. 2002), and many insect groups (Briscoe and Chittka 2001; Dyer et al. 2011). Such conservatism is not universal, however. Fish, for example, exhibit a large range of sensitivities that appear more closely related to ecology rather than phylogeny (e.g., Terai et al. 2006). Butterflies also show extraordinary diversification of receptor $\lambda_{\text{max}}$, possibly promoted by sexual signaling (Osorio and Vorobyev 2008).

A well-established principle from psychophysical research is that color perception in humans arises via the comparison of three neural input channels (Shapley and Hawken 2011). This and analogous findings in model animal systems (e.g., honeybees; Backhaus 1991; Chittka 1996; de Ibarra et al. 2014) has informed several key principles of visual perception. As we explore below, different photoreceptor channels feed directly into the opponent neurons that underlie the processing involved in color perception (Shapley and Hawken 2011).

**Chromatic versus Achromatic Visual Channels**

Color in humans is commonly described in terms of three dimensions: hue, saturation, and brightness (Kelber and Osorio 2010). Hue refers to the category of color (red, green, blue, etc.), and saturation refers to its deepness or spectral purity (e.g., pink is a less saturated version of red or purple). Brightness refers to the perceived intensity of a stimulus, independent of hue and saturation. Cornsweet (1970) defines brightness as that aspect of color perception that changes most dramatically with variation in total stimulus intensity. Hue and saturation are considered chromatic properties (i.e., aspects of perception most sensitive to changes in the stimulus spectrum), whereas brightness is considered an achromatic property. Brightness is a complicated phenomenon because it is influenced by both the total intensity (i.e., quantal flux) and spectral quality of a stimulus. Humans judge the apparent brightness of an object largely on its intensity relative to its surroundings and do so largely using visual channels not involved in color perception (Cornsweet 1970; Bowmaker and Dartnall 1980; Dowling 1987).

Studies in nonhuman animals have also revealed the existence of separate chromatic and achromatic channels (Kelber 2005; Osorio and Vorobyev 2005; de Ibarra et al. 2014), which are often used in different ways and for different tasks (Livingstone and Hubel 1988; Giurfa et al. 1997; Osorio and Vorobyev 2005; Schaefer et al. 2006; Zhou et al. 2012). Achromatic information is known to mediate the detection of motion, form, and pattern in a variety of vertebrate and invertebrate species. In flies, this information is sensed by a dedicated class of photoreceptors (retinular cells R1–R6; Hardie 1986) and used to judge movement, orientation, and edges in the visual field (Heisenberg and Buchner 1977; Zhou et al. 2012). Intriguing recent findings suggest that this receptor class also contributes to how flies perceive color (Kelber and Henze 2013). There is reasonable evidence that bees and reptiles detect achromatic information using a single class of LWS photoreceptors (Giurfa et al. 1997; Fleishman and Persons 2001), and many birds do so using double LWS or LWS/MWS receptors (Osorio et al. 1999). In other groups, such as nonhuman primates, the achromatic channel may rely on the summation of outputs across multiple photoreceptor classes (Livingstone and Hubel 1988). An achromatic channel
Table 1: Overview and brief definition of the principal terms related to the study of coloration

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absorption Fraction and spectral distribution of incident light that is neither transmitted nor reflected by an object. Though they are perceived simultaneously.</td>
<td></td>
</tr>
<tr>
<td>Achromatic/chromatic channels</td>
<td>Terms used in reference to the fact that many animals are thought to process color and luminance information separately, even though they are perceived simultaneously.</td>
</tr>
<tr>
<td>Ambient intensity</td>
<td>Broadly speaking, the total amount of light, measured either as the hemispherical sum of light reaching a particular location—see irradiance—or the sum of light arriving to a point (such as an eye) from a particular solid angle—see radiance; the appropriate units are micromoles per square meter per second per nanometer (μmol m⁻² s⁻¹ nm⁻¹) for the former case and μmol m⁻² s⁻¹ nm⁻¹ sr⁻¹ for the latter (where intensity is expressed per steradian [sr]); see Endler (1990) for more detail.</td>
</tr>
<tr>
<td>Brightness</td>
<td>Human-based perception of the overall intensity of light emitted or reflected from a stimulus; this term is used inconsistently and often incorrectly in the animal coloration literature, sometimes confused with color saturation (chroma); we advocate use of the term luminance instead for describing stimulus intensity (see text).</td>
</tr>
<tr>
<td>CIE</td>
<td>Commission internationale de l’éclairage (International Commission on Illumination), a body founded in 1913 that defines standards for the human perception of light and lighting across the arts, sciences, and image technology (see, e.g., Luo et al. 2001).</td>
</tr>
<tr>
<td>City-block/Euclidean distances</td>
<td>Metrics of Cartesian distance between points in a color space; using the analogy of a right-angled triangle, city-block (or Manhattan) distance is the sum of the two sides, and Euclidean distance is the hypotenuse.</td>
</tr>
<tr>
<td>Color constancy</td>
<td>Ability to classify or perceive a color stimulus as largely invariant despite widely varying illumination.</td>
</tr>
<tr>
<td>Color opponency</td>
<td>Key perceptual principle in which the output of different photoreceptor classes is processed in an antagonistic manner; e.g., humans and honeybees process color information according to opponent channels of red versus green and blue versus yellow.</td>
</tr>
<tr>
<td>Color space/chromaticity diagram</td>
<td>Methods for representing color stimuli, typically in relation to a vector coordinate system (e.g., x-y-z) in n-dimensional space; for humans, CIE diagrams (fig. 2b) seek to map the perception of equiluminant colors (Luo et al. 2001); various color spaces exist for animals (see text; figs. 1, 2).</td>
</tr>
<tr>
<td>Hue</td>
<td>Perceptual dimension describing the category of color (e.g., blue, red, green); hue perception is determined by the spectral shape of a stimulus, visual features such as the number/sensitivity of photoreceptor classes, and neural processes such as color categorization (e.g., Lunau 2014).</td>
</tr>
<tr>
<td>Intensity</td>
<td>Absolute amount of light in quantal flux (μmol m⁻² s⁻¹ nm⁻¹); for a particular stimulus, this represents the rate of photons reaching the eye, which for objects in natural habitats will be a product of ambient (A), reflected (R), and transmitted (T) light; ( Q = ART(x) ) (see Endler 1990 for more detail).</td>
</tr>
<tr>
<td>Illumination/ambient light/habitat light</td>
<td>Amount and spectral distribution of light in a particular environment, related to ambient intensity and irradiance; in natural habitats, this will be determined by sunlight and radiance from sources such as the sky, clouds, substrates, and vegetation (Endler 1993).</td>
</tr>
<tr>
<td>Irradiance</td>
<td>Number of photons (all sources of radiance summed) measured over a hemisphere normal to a surface over a particular time, measured in μmol m⁻² s⁻¹ nm⁻¹ (see Endler 1990); the form of irradiance depends on geometry: vector irradiance is proportional to the cosine of the angle of incident light to the perpendicular, while scalar irradiance is independent of incident angle and measures general illumination levels (Johnsen 2012).</td>
</tr>
<tr>
<td>Light</td>
<td>Energy in a narrow band of the electromagnetic spectrum, characterized for animals by the wavelength range of ~300–730 nm (for humans: ~400–700 nm); quanta of light energy are called photons, which propagate with wavelike properties.</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Luminance</td>
<td>Perception of stimulus intensity, which in humans is a function of both the amount and spectral quality of light; in studies of nonhuman animals, luminance is often estimated by the stimulus spectrum multiplied by overall spectral sensitivity.</td>
</tr>
<tr>
<td>Optical/ocular media*</td>
<td>Materials including but not limited to the lens, cornea, vitreous humor (vertebrate eyes), and waveguide structures (invertebrate compound eyes); of particular importance for color vision are screening pigments such as the primate macula lutea, the oil droplets of bird eyes, and the within-waveguide pigments of arthropods.</td>
</tr>
<tr>
<td>PCA and ordination diagrams</td>
<td>Techniques for summarizing and representing spectral data according to the principal orthogonal axes of variation; PCA refers to principal components analysis; parametric inference for spectral data using these approaches is limited (Endler and Mielke 2005; see table 2); their utility is explored in detail by Grill and Rush (2000).</td>
</tr>
<tr>
<td>Photon</td>
<td>Quantum of light energy, which is indivisible when captured by a photopigment.</td>
</tr>
<tr>
<td>Photon capture*</td>
<td>Ability of photopigments in the eye or specific photoreceptors to capture incident light, determined by factors including the area of a pupil or lens, field of view, neural summation (e.g., superposition; Hardie 1986), absorption and/or reflectance by ocular media, integration time, and quantum transduction efficiency (Johnsen 2012).</td>
</tr>
<tr>
<td>Photopigment/visual pigment*</td>
<td>Fundamental basis of light reception, comprised of a transmembrane opsin protein covalently bound to a vitamin A–derived chromophore; most simply, a photopigment absorbs wavelength-sensitive light and generates a subsequent neural signal.</td>
</tr>
<tr>
<td>Photoreceptor/spectral sensitivity*</td>
<td>Spectral range of photon capture that can be achieved by an individual photoreceptor, a class of photoreceptors, a retinal substrate, or an entire eye; defined by the minimal saturation at which a color can be discriminated from a white background.</td>
</tr>
<tr>
<td>Psychophysics</td>
<td>Branch of psychology dealing with the relationship between the physical properties of a stimulus (in this case, light) and an associated &quot;psychological experience&quot;; in animals, the stimulus experience must be assayed via behavior.</td>
</tr>
<tr>
<td>Radiance</td>
<td>Light emitted from (or transmitted through) a unit area of a surface measured over a small defined solid angle (typical units are photons m⁻² s⁻¹ nm⁻¹ sr⁻¹); radiance is often used to describe the light reaching an eye from a surface or object; see Endler (1993) for examples of radiance for the sun, blue sky, and various background objects.</td>
</tr>
<tr>
<td>Reflectance</td>
<td>Probability that photons of each light wavelength will be scattered in a defined direction; for more detail, refer to Endler (1990).</td>
</tr>
<tr>
<td>Receptor noise</td>
<td>Random fluctuations in the rate that photons arrive at photoreceptors; key to receptor-noise-limited threshold modeling (Vorobyev and Osorio 1998).</td>
</tr>
<tr>
<td>Reflectance spectroradiometer</td>
<td>Device capable of measuring the radiation of light from a surface; commonly referred to as a spectrometer; see Johnsen (2012) for further detail.</td>
</tr>
<tr>
<td>Saturation</td>
<td>Dimension of color perception that refers to spectral purity, i.e., the deepness or richness of a color stimulus; the saturation of a stimulus decreases when mixed with white light; also referred to as chroma.</td>
</tr>
<tr>
<td>Screening pigments*</td>
<td>Ocular pigments that absorb light and modify the spectral sensitivity of photopigments; these exist as the macula lutea (primates), oil droplets (vertebrates such as birds), or within the rhabdoms of compound eyes (invertebrates).</td>
</tr>
<tr>
<td>Spectral reflectance</td>
<td>Fraction of each light wavelength that is reflected from an object or surface.</td>
</tr>
<tr>
<td>Threshold (visual)/threshold modeling</td>
<td>Minimal stimulus, i.e., photon flux, required for the perception of a visual stimulus; this has been modeled for color stimuli according to the limits set by noise in receptor channels (Vorobyev and Osorio 1998).</td>
</tr>
<tr>
<td>Transmission</td>
<td>Probability that light photons of each wavelength will pass through an object.</td>
</tr>
<tr>
<td>Transmission media</td>
<td>Environmental media such as air and water that propagate light; propagation is subject to absorption or scattering by suspended particles (such as water vapor or dust), and light passing between different media is subject to refraction.</td>
</tr>
</tbody>
</table>

Note: Terms in boldface link to definitions elsewhere in the table, and those marked with an asterisk are treated in detail in the supplementary appendix.
underlying at least motion detection appears to be widespread, if not universal, and will often merit consideration in studies of animal vision (we refer interested readers to the in-depth review by Osorio and Vorobyev 2005).

The term brightness refers very generally to the subjective appearance of a stimulus, which in humans is known to depend on a complex set of factors. Brightness perception in nonhuman animals has not been elucidated (Kelber and Osorio 2010), although knowledge is accumulating in select systems such as bees and birds (e.g., Lind et al. 2013). Studies of color perception are often concerned with the fact that equally radiant yet spectrally different stimuli may affect perceived brightness due to the spectral sensitivity of photoreceptors involved in the achromatic channel. In human vision, the perception of stimulus intensity is a function of both the amount and spectral quality of light and is described by the term luminance (Wyszecki and Stiles 1982). It is convenient and reasonable to also refer to luminance in studies of animal vision when one is concerned with the perception of stimulus intensity. This is advisable because it is difficult to assess the complex factors that may determine brightness perception, whereas luminance can be more easily quantified. Knowledge of an animal’s luminance function is generally critical for estimating or testing color differences among stimuli because one must control for the effects of brightness differences arising from spectral variation.

**Color Opponency**

Color-matching experiments in humans have established that, over a wide range of intensities, a given ratio of stimulation of the three retinal receptor classes (SWS, MWS, and LWS) produces the same sensation of hue and saturation. This equivalent sensation breaks down only at very low and very high stimulus intensities. How do neural systems calculate ratios? In humans this involves two steps: (1) the neural elements in the retina produce a response that is approximately equal to the log of the rate of quantal capture (Dowling 1987); and (2) the neural outputs from different photoreceptor channels are then subtracted, thereby generating a difference signal. Given two neural signals (a and b), this is approximately

\[ \log(a) - \log(b) = \log\left(\frac{a}{b}\right) \].

Nervous systems approximately respond to the log of the ratio of stimulation among different photoreceptors (Dowling 1987). The elegance of this solution is that since \( a/b \) does not change with equal changes of intensity of a and b, the output signal is largely independent of overall intensity. The relationship between quantal capture and log neural output is not precisely linear, which explains the failure to produce identical color responses at very high and low intensities (Kaiser and Boynton 1996).

The neural processing of color has been examined across various nonhuman animals, and in each case some basis of opponent processing has been found (e.g., Marchiafava and Wagner 1981; Schiller and Logothetis 1990; Yang et al. 2004). Despite variation in details such as which receptor classes are compared, how they are weighted, and so on, opponent processing appears to be at the heart of animal color perception. It is represented accordingly in color analysis (table 2).

**Chromatic Adaptation and Color Constancy**

It is widely believed that an important driving force for the evolution of color vision is the ability of an animal to consistently identify objects in the environment despite highly variable illumination conditions. This ability—known as color constancy—calls on a mechanism by which the brain and/or eyes account for differences in the ambient illumination of objects and their relative surrounds. Most visual systems are thought to accomplish this through a process of chromatic adaptation (Webster 2011). For example, humans exhibit color constancy under a wide range of conditions (Foster 2011); for a given object and its setting, such as a red apple on a white bench, we perceive the apple as red even despite changes in illumination. This applies to both the categorization of hue and the perception of saturation (Reeves et al. 2008). Exceptions to this arise with drastic variation in the spectrum of illumination (e.g., a complete absence of red light would cause this apple to appear black) or if spectral illumination varies between the object and its immediate setting (as explored by Endler and Théry 1996).

Given that ambient viewing environments in nature vary greatly according to habitat, time of day, and weather conditions (Lythgoe 1979; Endler 1993), it is not surprising that some system of color constancy is ubiquitous in both vertebrates (e.g., Neumeyer 1981) and invertebrates (e.g., Chittka et al. 2014). How it is achieved is not completely understood (Foster 2011), and it is also known to be far from perfect (Dyer 1999). One simple yet extremely useful candidate model of color constancy is provided by the von Kries mechanism (von Kries 1905). Procedurally, the average spectrum falling on the eye is measured and multiplied by the absorption spectra of each photoreceptor class. The outputs of each class are then adjusted relative to one another such that the illumination spectrum produces equal outputs across all classes (Ender and Mielke 2005). This correction emulates the process of chromatic adaptation in humans viewing a broadly lit background area (Webster 2011) and provides a simple basis for incorporating color constancy into perceptual analysis.
Receptor Noise

Vision depends on the rates of photon capture by ocular photopigments, but these rates will, to some degree, vary randomly over time. At low light levels, these random variations are an important source of noise in the neural signal that results (Rovamo et al. 2001). This is why, for example, a scene appears grainy at very low light levels (Lythgoe 1979). At higher light levels, these quantal fluctuations become less important, and noise from other sources related to neural processing dominates. These two sources of random noise can be estimated from knowledge of a number of factors, including the intensity of light, photoreceptor size, and the density of photopigment contained in the receptors (Johnsen 2012). Studies in humans have shown that noise in visual channels leads to variation in how given color stimuli are perceived and limits the precision with which given colors can be identified (Kaiser and Boynton 1996). Specifically, discrimination between two similar color stimuli will only be possible if their spectral distributions are different enough in relation to the degree of noise (Wyszecki and Stiles 1982; Goldsmith 1990), a classical problem in signal detection theory. Notably, however, the relative importance of visual versus neural noise in limiting human color perception is poorly known (Kaiser and Boynton 1996).

In an influential article, Vorobyev and Osorio (1998) argued that if one assumes that animal color vision is based on opponent interactions constructed from different receptor class outputs, then the noise in each opponent channel will largely be determined by that arising in the relevant receptors. The limit to discriminating color stimuli can then be modeled according to the difference signals arising from constituent opponent channels in relation to noise. From these assumptions, Vorobyev and Osorio (1998) were able to predict the outcome of a number of earlier studies on different organisms (including humans) that measured how much chromatic stimuli had to differ from gray backgrounds in order to be detectable. While limited to a rather specific set of stimulus conditions, these results were nonetheless exciting because behavioral performance could be predicted from knowledge of the spectral sensitivity of the different photoreceptor classes and based on estimates of the noise in each channel. The model is generalizable for dichromatic, trichromatic, and tetrachromatic animals and provides detailed predictions for behavior of bees and birds (Vorobyev et al. 2001). We discuss the applicability and use of this important approach to color analysis later in the article.

Reconciling Perception within the Study of Coloration

A Question-Driven Framework

To be most effective, top-down color research faces the challenge of reconciling the breadth of perceptual knowledge and its founding assumptions in choosing how to analyze spectral data. Studies in most biological contexts (e.g., signaling, crypsis, deception, resource orientation) invoke animal viewers, and hence some form of sensory analysis will often be desirable (Endler et al. 2005). Sensory-based approaches require varying levels of visual knowledge (estimates of photoreceptor $\lambda_{\text{max}}$ and any important ocular screening media at a minimum) and make implicit assumptions about color vision and perception. They will be best applied when ecologically important viewers can be defined and prove particularly accurate in systems for which vision and/or perception is well characterized (e.g., bees; Giurfa 2004; Dyer and Neumayer 2005; Avarguès-Weber et al. 2010; de Ibarra et al. 2014). Even outside of these situations, incorporating whatever relevant information is available for visual systems and/or viewing environments may often enhance biological conclusions. However, this is not to say that a sensory-based approach will always be most appropriate. There will be many instances for which requisite visual data and/or assumptions cannot be reasonably met and still other studies whose goals do not depend at all on how the traits are viewed.

We present a generalizing framework for top-down research that is founded on clear articulation of the question in relation to color. The nature of the research question is critical to informing two key sets of decisions. First, consideration of whether ecologically relevant viewers can be identified and what visual data are subsequently available will determine whether a sensory approach is necessary and justifiable. Second, explicit consideration of the visual task, that is, the behavioral context in which viewers encounter the focal color trait(s), will define the most appropriate sensory analyses. Although the top-down color literature addresses a breadth of specific questions across many species and taxa, we recognize three broad categories that, in turn, correspond to three discrete analytic approaches, namely, (1) spectral/physical, (2) perceptual distance, and (3) discriminatory. We detail this framework below in summary via table 2. The broader goal is to highlight key considerations for top-down research and, ultimately, to illustrate how the complexity of visual perception precludes a singular or strongly prescriptive approach to color trait analysis.

Spectral/Physical Questions

Many studies seek to explain color variation at population, species, community, or higher ordinal levels, often in novel or unstudied systems. The research may address questions or situations that are largely independent of sensory systems. Examples include the use of color in taxonomy or phylogenetic reconstruction, efforts to understand the proximate basis of color production (e.g., Vukusic et al. 2000),
Table 2: Overview of the primary approaches to analyzing color data as classified according to research context (study question, visual task, and available sensory knowledge)

<table>
<thead>
<tr>
<th>Approach</th>
<th>Purpose of analysis</th>
<th>Technique/scheme</th>
<th>Assumptions</th>
<th>Required data</th>
<th>Works best on/was designed for</th>
<th>Potential strengths</th>
<th>Potential weaknesses</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>To physically summarize and/or classify spectral data</td>
<td>Segment analysis (Endler 1990)</td>
<td>Does not assume normality, homoscedasticity, and sphericity for statistical analyses</td>
<td>Raw spectra</td>
<td>Identifying trends, e.g., evolution of color pigments, clustering of spectra</td>
<td>No visual knowledge required; can detect broad trends; suitable for all data distributions</td>
<td>Unknown biological relevance; tendency to overinterpret data</td>
</tr>
<tr>
<td>S</td>
<td>As above</td>
<td>Principal components analysis (e.g., Bennett et al. 1997); independent component analysis; cluster analysis (Grill and Rush 2000)</td>
<td>No assumptions for goals of representing and exploring variance</td>
<td>As above</td>
<td>As above</td>
<td>No visual knowledge required; can detect broad trends</td>
<td>Unknown biological relevance; tendency to overinterpret data; parametric statistical inference invalid (Endler and Mielke 2005)</td>
</tr>
<tr>
<td>P</td>
<td>To ordinate and/or visualize color in &quot;perceptual space&quot;</td>
<td>Chromaticity diagrams, e.g., triangle (Maxwell 1860), tetrahedron (Endler and Mielke 2005), generalized n-dimensional spaces (Pike 2012a)</td>
<td>Color perception determined by outputs of all photoreceptor used in analysis</td>
<td>Photon capture by receptor cells (including $\lambda_{max}$ estimates); can include light conditions</td>
<td>Estimation of information available to the visual system</td>
<td>Few assumptions, few measures; simplicity; can compare qualitative differences (e.g., color categories); has some analogy to perceptual mechanisms</td>
<td>Unclear how perception scales with Cartesian distance; this limitation (here and in subsequent table rows) may be mitigated by nonparametric inference (e.g., Endler and Mielke 2005)</td>
</tr>
<tr>
<td>P and D</td>
<td>To ordinate and/or visualize color in &quot;perceptual space&quot;; to test</td>
<td>Animal-specific receptor models based on choice assays, e.g., the fly and butterfly species</td>
<td>Specific for flies and butterflies; assumes color choice is based on either a single</td>
<td>As above</td>
<td>Explaining color choice in butterflies and flies</td>
<td>Based on detailed behavioral data; makes explicit assumptions on</td>
<td>Unclear if it can be used to make inferences about discrimination in addition to choice; unclear</td>
</tr>
<tr>
<td>Approach</td>
<td>Purpose of analysis</td>
<td>Technique/scheme</td>
<td>Assumptions</td>
<td>Required data</td>
<td>Potential strengths</td>
<td>Potential weaknesses</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>---------------------</td>
<td>------------------</td>
<td>-------------</td>
<td>---------------</td>
<td>---------------------</td>
<td>----------------------</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Color discrimination</td>
<td>Receptor noise and other ideal observer models (e.g., Vorobyev and Osorio 1998; Vorobyev et al. 2001)</td>
<td>Discrimination is limited by noise originating in photoreceptor; information from photoreceptors is processed in ideal way</td>
<td>Photon capture; anatomical and/or physiological data to estimate relative noise in different receptor classes; irradiance of environment</td>
<td>Discrimination of signals at or near threshold; for a wide range of animals and viewing conditions</td>
<td>No assumptions of specific opponency mechanisms; can be modified for low- and high-light scenarios; powerful if combined with behavioral knowledge</td>
<td></td>
</tr>
<tr>
<td>P and D</td>
<td>To ordinate and/or visualize color in &quot;perceptual space&quot;; to test specific hypotheses about color discrimination</td>
<td>Color hexagon (Chittka 1992)</td>
<td>Color is coded by two or more unspecified color opponent mechanisms; the output is combined to calculate perceptual differences, uses Euclidian metric</td>
<td>Photon capture by receptor cells (including $\lambda_{\text{max}}$ estimates); can include light conditions</td>
<td>Method is based on COC and is mostly used for honeybees but can be applied to any trichromat</td>
<td>Uses unspecified opponency mechanisms; unclear how perception scales with Cartesian distance, including whether distances in all directions are equally informative</td>
<td></td>
</tr>
<tr>
<td>P and D</td>
<td>As above</td>
<td>CIE color space for humans (Luo et al. 2001); COC color space for honeybees (Backhaus 1991)</td>
<td>Color is coded by opponent mechanisms known for focal species; output combined to estimate perceptual difference using city-block metric</td>
<td>Humans and honeybees</td>
<td>Behavioral studies in humans and honeybees allows a meaningful link between physical data and perception</td>
<td>Position of points in color space change with light intensity</td>
<td></td>
</tr>
<tr>
<td>P and D</td>
<td>butterfly color choice model (Kelber 2001)</td>
<td>receptor or a linear interaction between receptors</td>
<td>how receptors interact</td>
<td>how Cartesian distances relate to perceptual distances</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: Spectral/physical (S) approaches analyze raw spectra (reflectance or radiance data), perceptual distance (P) approaches recruit available visual knowledge and incorporate some principles of color perception, while discriminatory (D) approaches are based on receptor-noise-limited modeling. The basic assumptions and data requirements are summarized. Descriptions of what questions the techniques were originally designed to address, and their potential strengths and weaknesses, indicate their most appropriate application(s). Literature citations identify illustrative examples. With the exception of segment analysis, all other methods assume homoscedasticity and sphericity for statistical analyses (unless nonparametric methods are used). CIE = Commission internationale de l'éclairage; COC = color opponent coding.*
and many studies of readily categorized color morphs (e.g., Clarke and Sheppard 1972). Alternatively, the primary viewers of focal color trait(s) may be species for which nothing is known or can be inferred about visual sensitivity, such as an unstudied butterfly species. Still further, studies may involve traits with a large and/or unknown suite of relevant viewers, such as the potential range of arthropod prey attracted by the color lures of orb spiders (Théry and Casas 2009). These questions are best addressed via nonsensory analyses, which require no explicit assumptions about vision or perceptual processing.

Color data in these cases are summarized and analyzed according to raw spectral curves (i.e., reflectance or radiance data; Endler 1990; Grill and Rush 2000). Several techniques exist for doing this, including integration over discrete wavelength ranges, analysis of variance, and principal component analysis (PCA), and ordination cluster analyses (tables 1, 2). These techniques can be used to represent objects in charts of spectral space (Endler 1990), a nonsensory analogy to the color spaces described in the next section. Given the absence of sensory data, in applying this approach it is important to acknowledge the limitations to inferences about color ecology and evolution. Spectral data also generally violate the assumptions of parametric statistics, which limits the ability for statistical inference (e.g., Endler and Mielke 2005). Analysis of achromatic information is relatively more straightforward and can be achieved by integrating across entire spectral curves or across wavelength intervals of specific interest (e.g., Andersson et al. 1998; Kemp and Rutowski 2007). However, biological inference will again be limited by knowledge about how achromatic information is received and processed by the viewer(s) of interest (e.g., Schaerer and Neumeyer 1996; Lind and Kelber 2011; Stojcev et al. 2011; Zhou et al. 2012; Lind et al. 2013). At a minimum, knowledge of the spectral range or function of achromatic sensitivity in viewers will be useful to guide how to best summarize achromatic information (see, e.g., Prudic et al. 2007).

Perceptual Distance Questions

A range of studies seek to understand how differently two (or often more) colors will appear to particular viewers. For example, one might wish to appraise the chromatic contrast between adjacent elements of a color pattern or in the range of colors present within and/or among different species, populations, sexes, or morphs (e.g., Andersson et al. 1998; Endler et al. 2005; Stoddard and Prum 2008; Maia et al. 2013). Alternatively, one might wish to determine how greatly specific colors or color patterns differ from viewing backgrounds (e.g., Heinsohn et al. 2005). Whenever we inquire about the magnitude of difference between different colors, we are essentially asking how far apart these stimuli appear in perceptual space. These questions, therefore, explore spectral variation relevant to defined viewers and are best served by analyses that incorporate sensory information. For most systems of interest to top-down research, sensory knowledge will be limited to the peak sensitivities of the different classes of retinal photoreceptors (i.e., as informed by estimates for $\lambda_{\text{max}}$ and ocular screening media). This level of information makes it possible to plot the spectral reflectance of objects in a photon-capture-based color space—a form of chromaticity diagram (see figs. 1, 2). Such diagrams represent color stimuli according to the relative stimulation of relevant photoreceptor classes (i.e., those known to contribute to chromatic perception in a particular viewer). Examples for trichromats such as primates and most insects include the color triangle (Maxwell 1860; fig. 2a) and hexagon (Chittka 1992). In the case of tetrachromats such as fish, birds, reptiles, and some salticid spiders, the appropriate diagram is a tetrahedron (Goldsmith 1990; Neumeyer 1992; Endler and Mielke 2005; fig. 1b). Here we focus largely on the triangle and tetrahedron.

In chromaticity diagrams, the outputs of all photoreceptor classes are scaled such that a white object under average habitat illumination stimulates each class equally (as per the principles of chromatic adaptation; see below). For a given color stimulus, the photon capture of each photoreceptor class is calculated (as the stimulus spectrum $\times$ spectral sensitivity) and divided by the sum across all classes, thus generating values ranging from zero to one. Following geometric transformation (see, e.g., equation [20] in Endler and Mielke 2005), specific stimuli are henceforth represented by three (color triangle) or four (tetrahedron) values and plotted in a diagram whose apices represent the values of each photoreceptor class (and therefore sum to one). Any number of colors can be plotted in a single diagram, which makes this approach useful for illustrating the position of different stimuli in relative photon-capture space. One can plot the overlap or distribution of different populations of points to explore the differences between whole color patterns (Endler and Mielke 2005; Endler et al. 2005; Stoddard and Stevens 2011; fig. 1b) and calculate Cartesian distances between any pair of colors as an estimate of their potential perceptual difference (fig. 1c).

A general criticism of chromaticity diagrams is that they only represent stimuli as delivered to the visual system and do not account for sensory and/or neural processing. However, this argument is not entirely correct. First, given that relative excitation among photoreceptor classes is the physical basis of color perception, these diagrams depict the information available to the visual system and are therefore at least useful for identifying the potential limits of perception (Lythgoe 1979). Second, by assuming that a
### Figure 1: Example of a sensory-based approach to analyzing and representing color.


#### b. Stereo and unpacked views of the color tetrahedron, indicating the location of each point in three-dimensional space. Points are calculated according to the U-type avian visual system (Endler and Mielke 2005) and assume objects are viewed under open/cloudy habitat light (Endler 1993) by an eye chromatically adapted to such light. The position of an achromatic (white or gray) object is indicated by the open circle in each plot.

#### c. Example metrics for representing potential chromatic contrast and discriminability of bird colors against either leaf or trunk backgrounds. Chromatic contrast estimates are Euclidean distances between each pair of points in tetrahedral space, whereas just noticeable difference (JND) values are obtained from receptor-noise-limited modeling (Vorobyev and Osorio 1998). Existing behavioral evidence only supports inferences about discriminability at values around 1.0 JND (assuming that noise in chromatic channels is decisive; also see fig. 3 for how equivalent inferences may be drawn from tetrahedral data). Labels at tetrahedral and triangular space vertices refer to the relative stimulation of ultraviolet (U), shortwave (S), midwave (M), and long-wave (L) photoreceptor channels. Eclectus photograph courtesy of Doug Janson.

<table>
<thead>
<tr>
<th></th>
<th>Viewed against:</th>
<th>Leaves</th>
<th>Trunks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contrast (Euclidean distance)</strong></td>
<td>Cheek (C)</td>
<td>0.414</td>
<td>0.382</td>
</tr>
<tr>
<td></td>
<td>Bill (Bi)</td>
<td>0.286</td>
<td>0.252</td>
</tr>
<tr>
<td></td>
<td>Back (B)</td>
<td>0.102</td>
<td>0.192</td>
</tr>
<tr>
<td></td>
<td>Tail (T)</td>
<td>0.320</td>
<td>0.225</td>
</tr>
<tr>
<td></td>
<td>Mantle (M)</td>
<td>0.407</td>
<td>0.302</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Discriminability (JND)</strong></th>
<th>Viewed against:</th>
<th>Leaves</th>
<th>Trunks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cheek (C)</td>
<td>39.8</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>Bill (Bi)</td>
<td>37.9</td>
<td>33.3</td>
<td></td>
</tr>
<tr>
<td>Back (B)</td>
<td>22.3</td>
<td>30.3</td>
<td></td>
</tr>
<tr>
<td>Tail (T)</td>
<td>50.7</td>
<td>41.3</td>
<td></td>
</tr>
<tr>
<td>Mantle (M)</td>
<td>68.4</td>
<td>58.2</td>
<td></td>
</tr>
</tbody>
</table>
white object will produce equal photoreceptor class outputs, chromaticity diagrams effectively incorporate the principle of chromatic adaptation, that is, a von Kries–type mechanism of color constancy. Third, as noted earlier, color perception has been modeled in terms of opponent channels (Shapley and Hawken 2011) that are approximated by the log ratio of stimulation among different photoreceptor classes (eq. [1]). Stimuli are plotted in chromaticity diagrams according to ratios of photon capture across receptor classes; hence, the way they represent spectral information is closely analogous to neural processing. It is, therefore, not entirely surprising that distances in color spaces have been shown, in at least some cases, to reasonably approximate behaviorally determined perception. In Anolis lizards, for example, the probability of detecting a colored stimulus moved against a colored background is directly proportional to the distance in tetrahedral space between stimulus and background colors (Fleishman and Persons 2001). Similarly, the extent of color overlap in tetrahedral space between the eggs of hosts and nest parasites is directly related to the likelihood of parasite egg rejection by the hosts (Stoddard and Stevens 2005; see further examples in Endler and Mielke 2005 and Endler et al. 2005).

Chromaticity diagrams provide a useful means for first approximation of color differences because they incorporate available sensory information and follow calculations somewhat analogous to perceptual principles. Their use and interpretation implicitly assume that Cartesian distances among points plotted in color space scale in some way with biological perception; that is, they inform how stimulus differences might be perceived in a viewer’s brain. A critical consideration here concerns the magnitude of the Cartesian distances between plotted points relative to the magnitude of discriminatory thresholds. For distances similar in magnitude to the threshold distance, perceptual inferences may be guided by the principles of receptor-noise-limited modeling (Vorobyev and Osorio 1998). Discrimination based on thresholds may be represented within color space (e.g., MacAdam 1942; Endler and Mielke 2005; Stoddard and Stevens 2011). We illustrate an analytic example for how this may be achieved in figure 3 but defer detailed discussion of discrimination to following sections. Stimuli that are more widely separated in color space represent suprathreshold variation, and how animals perceive such variation is largely unknown (Kelber and Osorio 2010). Despite evidence that color space units scale proportionately with behavioral responses in a few specific contexts (e.g.,

![Figure 2](image_url)

**Figure 2:** Heuristic for how psychophysical insights can be used to scale the representation of spectral stimuli in color space. 

\(a\), A Maxwell triangle (Maxwell 1860) indicating the colors of *Eclectus roratus* and their backgrounds (from fig. 1a). Colors are plotted simply according to their relative stimulation of the three human cone classes (fig. 1b). 

\(b\), The same points represented in CIELUV chromaticity space. Diagrams of this nature are modified (i.e., calibrated) according to psychophysical data of human discrimination between equiluminant color stimuli (for an average human viewer under idealized daylight viewing conditions, e.g., the CIE standard illuminant D65; Judd et al. 1964). Cartesian (Euclidean or city-block) distances between stimulus points will more accurately scale with perceptual distance in \(b\) versus \(a\), as indicated by the differences in the vector \(BC\) between panels. Unless verified against behavior, Euclidean distance will provide a reasonable first approximation of relative perceptual differences among spectral stimuli. Abbreviations are as in figure 1.
Fleishman and Persons 2001), attempts to reconcile perception across color space are limited to categorization in humans (the Commission internationale de l’éclairage [CIE] color space; e.g., MacAdam 1942; Luo et al. 2001; fig. 2b) and honeybees (the color opponent coding [COC] color space; Backhaus and Menzel 1987; Backhaus 1991). Even for humans, these efforts have provided imperfect and highly context-dependent results (Kaiser and Boynton 1996; Luo et al. 2001). The problem, as perhaps best exemplified by the ongoing discoveries in honeybee psychophysics (e.g., Giurfa et al. 1997; Avarguès-Weber et al. 2010; Dyer et al. 2011) is that we only barely understand the complexity of higher-order processing (Kelber and Osorio 2010; Skorupski and Chittka 2011; Avarguès-Weber and Giurfa 2014). As we explore in the final section, how to draw inferences about suprathreshold perception from chromaticity diagrams presents a great ongoing challenge for the field.

**Discriminatory Questions**

A third broad category of top-down research considers whether color stimuli can be reliably discriminated by spe-

---

**Figure 3:** Relationship between geometric distances in tetrahedral color space and just noticeable difference (JND) as modeled according to receptor noise (Vorobyev and Osorio 1998). 

- **a,** Positions of four colors in tetrahedral space, showing their relative stimulation of the four cone classes (black bars for U, S, M, L; Endler and Mielke 2005). The gray boxes at the tops of the cone stimuli bars indicate ±1 standard deviation (SD); receptor noise causes the level of cone outputs to vary at random around their predicted (mean) value with a SD of $w/\sqrt{g_i}$, where $w = 0.05$ for most vertebrates in full daylight and $g$ is the relative abundance of cone class $i$ (for birds, a typical example for the $g$ set is 1, 2, 2, 4). This captures the fact that rarer and smaller shorter-wavelength-sensitive cones are associated with more receptor noise; hence, their output levels 1, 2, 2, 4). This captures the fact that rarer and smaller shorter-wavelength-sensitive cones are associated with more receptor noise; hence, their output levels vary at random around their predicted (mean) value with a SD of $w/\sqrt{g_i}$. 

- **b,** An illustration of the relationship between the geometry of tetrahedral color space and just noticeable difference (JND ($\Delta S$)). When two mangos, with radii defined by 1 SD instead of 2 SD, just touch, their centroids (mean stimuli, indicated by black dots) are 2 SD or 1 JND apart (indicated by the lines). This indicates ≤5% chance of not distinguishing the two stimuli. The radii choice depends on the purpose: using the 95% criterion, we use 1 SD ellipsoids to compare two stimuli and a 2 SD ellipsoid to define a zone of confusion. Behavior discrimination tests often use a 75% criterion ($1.15 \approx 0.575$ SD radii) instead of the 95% criterion for 1 JND, and this means smaller ellipsoids that are closer together when touching. Note the lack of homogeneity of scale in the tetrahedron with respect to JNDs, which arises due to differences in noise in the different cone classes. If two stimuli differ horizontally, they can be closer together (more spectrally similar) and still be distinguishable ($\Delta S_1$) than if they differ vertically (vertical axes). If colors were equally spaced throughout tetrahedral space, more of them would be distinguishable toward the L-M edge than toward the U-S edge and even fewer would be distinguishable higher on the vertical axis. This inhomogeneity of scale invalidates assumptions that distances are equally discriminable throughout the tetrahedron, which supports permutation-based rather than parametric statistical analyses of such data. Abbreviations are as in figure 1.
cific viewers. By definition, any study of variation in a very restricted chromatic range deals with the question of discriminability. This will apply in many biological contexts, including mate choice, cryptism, mimicry, and brood parasitism (e.g., Endler 1991; Endler and Houde 1995; Heiling et al. 2003; Spottiswoode and Stevens 2011). Short of using behavior to verify discriminability, studies of color traits that deal with this question ultimately require a sensory-based analysis (where possible). The most popular approach is the receptor-noise-limited model of Vorobyev and Osorio (1998).

The receptor-noise-limited model represents differences in detectability between two color stimuli in terms of just noticeable difference (JND) units, which under bright light are equivalent to Weber fractions (Wyszecki and Stiles 1982). One JND is taken to approximate the minimum stimulus difference required to produce detectable variation by a given receiver. This value has been defined empirically (e.g., 75% correct responses in a color choice test; Vorobyev and Osorio 1998) or predicted in terms of the number of standard deviations separating two stimuli (fig. 3). The threshold of discrimination is therefore modeled as a relative rather than absolute quantity. JND could change, for example, if an animal tolerates more or fewer incorrect identifications, if the intensity of ambient illumination changes greatly (such as a bright, sunny day versus late twilight), or if a viewer examines a stimulus for a longer time period (i.e., greater integration time, as analogous to increased shutter speed in photography; e.g., Narendra et al. 2013). However, for a given criterion, the relative discriminability of different pairs of colors has been shown to remain largely consistent (Vorobyev and Osorio 1998). Emerging evidence in bees suggests that there may also be a plastic component to chromatic discrimination thresholds; for further detail, we refer interested readers to the primary literature (e.g., Giurfa 2004; Avarguès-Weber et al. 2010; Dyer 2012).

Effective application of receptor-noise-limited modeling demands consideration of issues such as ambient illumination and signal intensity (Burnham et al. 1957; Newhall et al. 1957; Dyer and Chittka 2004). Different calculations are required to estimate receptor noise under bright versus dim ambient viewing conditions (for details, see Vorobyev and Osorio 1998). The model makes specific assumptions with regard to visual adaptation and assumes that all relevant information resides in chromatic channels. It is therefore not applicable to achromatic variation or for situations where the achromatic channel is understood to dominate, such as motion detection in many groups (as discussed earlier). Opponent processing is unspecified; that is, a visual system with \( n \) receptor classes is assumed to contribute \( n - 1 \) opponent channels (Vorobyev and Osorio 1998). Similarly, higher-level neural processing is not considered. This means that the model is potentially widely generalizable, but its relevance will be questionable when receptor noise is not the dominant influence on color discrimination. In humans, for example, threshold discrimination is most relevant when colors are viewed simultaneously rather than when successively encountered stimuli must be coded and retrieved from memory (e.g., Newhall et al. 1957; Uchikawa and Shinoda 1996; Perez-Carpinell et al. 1998). Honeybees also discriminate with greater accuracy under simultaneous viewing conditions (Dyer and Neumeyer 2005), which is important considering that foraging bees are most likely to encounter flowers sequentially. This does not render the modeling of thresholds automatically invalid, but it does warrant cautious interpretations and explicit statements about viewing conditions. Unfortunately, little is known about how the nature of the visual task influences color discrimination outside of humans and honeybees (see also Giurfa et al. 1997).

**Future Challenges**

**Informing a Basis for Sensory Analysis**

The increased means and motivation to quantify spectral data has expanded the breadth of taxa and ecological scenarios for which biologists seek to study coloration. We have noted how sensory-based analysis is often desirable, but the increasing use of novel study species will mean that requisite visual data (e.g., at a minimum, receptor sensitivity estimates) are frequently unavailable. This will present the temptation to extrapolate from what is known in related taxa or well-characterized model organisms. However, given the complexity of visual perception, adopting surrogate parameters will almost inevitably lead to a loss of analytical accuracy. This presents as a key judgment call for the field of color research as it expands across and into an increasing range of novel taxa: Under what circumstances can using surrogate parameters enhance the biological insights gained from a purely nonsensory analysis?

In practice, this is a question that will require careful consideration on a case-by-case basis. For intensively studied groups such as hymenopterans (Dyer et al. 2011) and mammals (particularly primates; Osorio et al. 2004), the fundamentals of vision and even perception are well established. In less-studied groups, this decision can be guided by knowledge of evolutionary conservatism in key visual parameters. For example, among most birds (Hart and Vorobyev 2005; Hart and Hunt 2007) and most insect groups (Briscoe and Chittka 2001), the evidence suggests relative invariance in photoreceptor sensitivity. While this is not nearly grounds for assuming uniformity of visual perception (given that \( \lambda_{max} \) is just one of the many constituents of color perception; e.g., see below and the appendix),
it does offer a basis for likely improvement on a nonsensory analysis. The confidence attached to surrogate parameters declines with both the depth of knowledge for particular taxa as well as the established degree of within-taxon variance in key visual parameters. In Anoline lizards, for example, many species exhibit similar visual sensitivity, yet there are occasional exceptions (e.g., Loew et al. 2002), which implies some need for caution. Yet other taxa such as fish and butterflies (Osorio and Vorobyev 2008) exhibit evolutionary diversification to the extent that intrafamilial and even intrageneric extrapolation is clearly unwise. These considerations place onus on individual researchers to explicitly account for their choice against the evidence provided in the relevant literature (e.g., Goldsmith 1990; Douglas and Marshall 1999; Briscoe and Chittka 2001; Loew et al. 2002; Kelber et al. 2003; Newman and Robinson 2005; Osorio and Vorobyev 2005, 2008; Hart and Hunt 2007; Kelber and Osorio 2010; Lunau 2014).

Another avenue for inferring visual parameters is to identify the gene sequences known to code for receptor photopigments (i.e., opsins). This can also be problematic. First, the presence of particular opsin sequences does not necessarily inform whether or how photopigments are expressed in the retina. Expression may be inhibited by mutation (Newman and Robinson 2005), depend conditionally on which other opsin genes are present (Archer et al. 1995), or differ with age and/or gender (Laver and Taylor 2011). Different opsins may be coexpressed in single photoreceptors (Röhlich et al. 1994; Arikawa et al. 2003) or conjugate with different chromophores in vivo, all of which cannot be informed by genetic data. A second complication is that molecular data usually do not inform intraocular transmission. At the basic level, this is due to filtration or reflection by ocular media (e.g., the cornea, aqueous humor, lens, and vitreous humor; Walls and Judd 1933), including screening by pigments such as the macula in humans; oil droplets in fishes, reptiles, and birds; and intrarhabdom pigments in arthropods. Although such effects will often be mitigated by chromatic adaptation, absorption of shortwave light (Douglas and Marshall 1999) may restrict the overall range of spectral sensitivity (as in the lack of human ultraviolet sensitivity). Intracellular filtration is more important because it may alter the sensitivity of individual receptors or receptor classes, hence modifying the basis of color opponent processing. Knowledge of spectral biases at the intracellular level is largely incomplete and confined to model systems, which limits generalizations for populations, for species, and at higher ordinal levels. See the appendix for a more detailed treatment of this issue.

Overall, the assumptions arising from extrapolating visual parameters need careful and explicit justification. It will be important to consider that different analytic approaches may be more sensitive to error in some parameters than others. For example, Linde and Kelber (2009) demonstrated that receptor-noise-limited modeling is considerably more robust to misestimation of the spectral sensitivity of receptors than the relative noise in receptor channels. Receptor noise is unfortunately very difficult to quantify, but the widespread use of surrogate data carries the danger of generating systematic biases in the literature. Robustness to particular assumptions can and ideally should be explored by contrasting the conclusions gained from different analytical choices (sensu Grether et al. 2005). This could include comparing the outcomes of approaches (i.e., sensory- versus nonsensory-based analysis) or investigating results for envelopes of values for key parameters (such as photoreceptor \( \lambda_{\text{max}} \)). The more that assumptions, educated guesses, or surrogate inferences are built into an analysis, the more its outcomes should be considered working hypotheses as opposed to strongly supported conclusions.

Quantifying and Representing Perception

For animal systems where visual data are available, perhaps the greatest challenge is how to represent the perception of suprathreshold colors (Kelber and Osorio 2010). As noted earlier, the biological basis for inferring this decision is actually rather limited. Complexity in neural processing— which continues to emerge even in well-characterized systems (e.g., Kelber and Henze 2013)—argues against simple generalizable solutions across taxa, different visual tasks, or different viewing situations. Animals may show biases arising from higher-order processes such as innate preferences, learning, and memory that do not reflect their relative ability to see different stimuli (Kelber 2005; Kelber and Osorio 2010; Skorupski and Chittka 2011). For example, color categorization is prominent in humans—as evidenced by our perception of hue—and there is some evidence for categorization as a learned property in fish (Poralla and Neumeyer 2006), birds (Ham and Osorio 2007), and flies (Lunau 2014). On the whole, however, whether and how animals categorize color is largely unknown (Kelber and Osorio 2010). Still further, animals make use of only some of the available sensory information for particular tasks (Giurfa et al. 1997) or prioritize information in some channels over others (Kelber 2005). This is seen in behavior under natural settings; for example, fruit-foraging crows switch between prioritizing chromatic versus achromatic information across different habitats depending on which channel proves more informative (Schaefer et al. 2006). Evolutionarily, such flexibility in the use of different available visual channels should often prove adaptive.

For top-down studies, an important question is whether suprathreshold perception is best appraised by simple color
space metrics (e.g., Euclidean distances) or if biological accuracy is increased by scaling in relation to discrimination thresholds (i.e., JNDs). We indicated earlier how such thresholds can be represented in chromaticity diagrams (fig. 3); conversely, several workers have considered that JNDs may present appropriate units for scaling supra-threshold variation (Siddiqi et al. 2004; Ham and Osorio 2007; Pike 2012b). Although estimates in JND units should broadly correspond to Euclidean color space distances, there can actually be considerable disparity between the two (see fig. 1c). The rationale for JND-based scaling follows from simple optical principles such as Fechner’s law, which predicts a logarithmic relationship between stimulus intensity and detection (Schrödinger 1920; Stevens 1957). However, these are principles of vision and therefore say nothing about the higher-level processing that characterizes perception (see further discussion of this point in Kelber and Osorio 2010). Vorobyev and Osorio’s (1998) model, for example, seeks to predict JNDs according to photoreceptor noise, a visual limitation that need not necessarily influence or relate to how suprathreshold color is perceived. By the same token, it is also true that color spaces and metrics such as Euclidean distance do not explicitly model perceptual mechanisms. Ultimately, the question of scaling will need to be confirmed by behavioral studies (e.g., Fleishman and Persons 2001; Ham and Osorio 2007). In the absence of such data, JND scaling should not be considered by default to increase the biological validity of chromaticity diagrams, except in cases nearing threshold variation. As an alternative to using Euclidean, JND, or other methods of distance scaling, one can analyze the geometry of color space using nonparametric statics that do not assume homogeneity of spatial scale, as in Endler and Mielke (2005).

The Ultimate Importance of Behavior

Behavior has a special role in the study of coloration because it resides at the intellectual interface of top-down versus bottom-up research. As clearly elucidated by Endler et al. (2005), the ecology and the evolution of color traits in nature will ultimately be driven by behavioral responses of individual viewers (e.g., the behavior of predators to prey or pollinators to flowers). Likewise, behavioral assays offer the ultimate basis for validating our understanding of animal color perception (Kelber et al. 2003; Kelber and Osorio 2010). Approaches such as receptor-noise-limited modeling (Vorobyev and Osorio 1998), for example, draw explicit links between sensory features and perceptual capacity, but it is important to realize they are based on principles generalized largely from human psychophysics and tested against behavior in relatively few animals (Vorobyev and Osorio 1998; Vorobyev et al. 2001).

Highly controlled psychophysical experiments will ultimately be necessary to define key parameters such as noise thresholds and the nature of perceptual scaling for specific viewers and/or visual tasks. This work is challenging because it requires subjects to perform complex behaviors under highly controlled visual environments (e.g., Chittka et al. 2003; Giurfa 2004; Avargués-Weber et al. 2010). The outcomes are fundamental, but such work will not prove achievable for many study systems. In this sense, the value of insights gleaned through less-stringent approaches should not be overlooked. Best-case scenarios are when quantitative predictions derived from visual physiology, modeling, and/or opsin-based inferences can be tested against behavioral responses specific to the research question (e.g., Nagata et al. 2012). These studies address the generalizability of principles considered as fundamental to color vision and processing in animals. Similarly, knowledge of behavioral responses to color variation in natural settings (e.g., Schaefer et al. 2006; Rojas et al. 2014) may play an important role in linking theoretical prediction to the visual complexity of the real world (see below).

Conclusion

The study of coloration has been transformed over the past several decades by the realization that animals view and perceive their world very differently than humans. Recent advances in understanding and modeling non-human visual perception, coupled with the unprecedented ability for biologists to quantify spectral information (i.e., color traits, viewing backgrounds, and other important features of natural visual environments), have poised the field to achieve rapid progress. This article is motivated by the thesis that synergy and reciprocity at all intellectual levels will determine the rate of such progress and the quality of insights arising. We provide a basis for conceptual alignment across the field by reconciling key principles of vision/perception (i.e., bottom-up knowledge) with the available color analyses and then placing this information into a framework based on the research questions of interest to top-down empiricists.

At the top-down level, a clearly articulated research question is paramount for specifying whether identifiable viewers exist and then—if so—for framing the analysis principally in terms of either near-threshold discrimination or supra-threshold perception (or both, in some cases). We suggest that sensory-based analysis will be desirable for most instances that involve identifiable viewers, even when sensory knowledge is imperfect. However, the effectiveness of this work will critically depend on informed consideration of analytic assumptions (table 2), justified use of surrogate parameters, and explicitly stated caveats to study conclusions (as discussed at length
in the text). At the bottom-up level, priority areas for future research will be informed via comparative analyses and reviews of the top-down literature. Currently, we have identified a particular need to better understand how non-human animals perceive and scale suprathreshold color variation (including the issue of whether color is categorized; Ham and Osorio 2007; Kelber and Osorio 2010). More broadly, it will prove important for bottom-up researchers to resolve the phylogenetic basis of key visual and perceptual parameters. As top-down studies extend across a breadth of novel color traits and species, emphasis at the bottom-up level may usefully switch from elucidating the finer details of model systems to understanding variation at and across higher taxonomic levels.

Ultimately, synergy and reciprocity across the schools of visual perception and color trait ecology/evolution will require an appreciation for how each school can guide development in the other. We envisage a growing opportunity for traversing the bottom-up/top-down boundary via an iterative process where bottom-up considerations are used to generate key—albeit sometimes imperfect—hypotheses for testing in top-down research. In this way, the available color analytics and their underlying assumptions become candidate models for testing in innovative ways rather than mere tools for informing system-specific conclusions. This iterative process can help focus top-down researchers on which of the most critical bottom-up assumptions need to be tested empirically in particular animals and visual scenarios. The outcomes can in turn contribute to the baseline of bottom-up knowledge, thereby providing crucial real-world feedback about the generalizability of visual and perceptual principles. Studies such as Schaefer et al. (2006) and Rojas et al. (2014) exemplify this approach and offer great potential for articulating the link between color variation and perception in nature.

Acknowledgments

We are indebted to E. Brodie III, S. Kalisz, and T. Oakley for guidance and comments on early versions of the manuscript. The material and concepts presented here were informed greatly by participants of the color research workshop at Macquarie University, particularly M. Vorobyev and M. Westoby; and the Macquarie University Genes to Geoscience Research Centre supported the workshop initiative. The Eclectus parrot image featured in figure 1 was kindly provided by Doug Janson. Funding from the Australian Research Council supported D.J.K. (DP-140104107), J.A.E. (DP-110101421), A.G.D. (DP-0878968 and DP-130100015), A.T.D.B. (LP-140100691), and J.M. Institutional support was provided by Macquarie University to D.J.K., M.E.H., and M.J.W. L.J.F. acknowledges support from the National Science Foundation (NSF-1051796), and N.S.H. acknowledges support from the University of Western Australia and the Western Australian government (through a Premier’s Fellowship to S. Collin). Authors A. T. D. Bennett, A. G. Dyer, and N. S. Hart are credited alphabetically.

Literature Cited


Associate Editor: Edmund D. Brodie III

Editor: Susan Kalisz