

Object Colours, Material Properties and Animal Signals

Lucas Wilkins^{*1,2} and Daniel Osorio²

¹Department of Zoology, University of Oxford, Oxford, United Kingdom

²School of Life Sciences, University of Sussex, Brighton, United Kingdom

April 6, 2019

Abstract

Colour is commonly regarded as an absolute measure of object properties, but most work on visual communication signals is concerned with colour differences, typically scaled by just noticeable differences (JNDs). Object colour solids represent the colour gamut of reflective materials for an eye. The geometry of colour solids reveals general relationships between colours and object properties which can explain why certain colours are significant to animals and evolve as signals. We define a measure of colour vividness, such that points on the surface are maximally vivid and the ‘grey’ centre is minimally vivid. We show that a vivid colour for one animal is likely to be vivid for others, and highly vivid colours are less easily mimicked than less vivid colours. Further, vivid colours such as black, white, red, blue and light, unsaturated shades are produced by pure or orderly materials. This kind of material needs to be created and maintained against entropic processes that would otherwise degrade or destroy them. Vivid coloration is therefore indicative of ecological affordance or biological function, so that it is valuable to have attentional biases towards these colours regardless of any specific significance.

1 Introduction

Animal and plant communication signals carry varied messages: some are attractive while others are defensive; courtship displays appeal to naïve viewers, whereas flowers and aposematic signals need to be memorable. Despite their various functions, and the diversity of colour vision in their natural receivers, signals generally include a limited palette of colours including black, white, saturated hues and light but unsaturated colours such as pink, whereas greys and browns are infrequent.

Any signal must attract attention and engender a response – it will be ineffective if it is overlooked or ignored. Why then should certain colours attract the attention of an eye with any given set of spectral photoreceptors, and why are the same types of spectra appropriate for animals with different types of colour vision? We show here how answers to these questions might be found if a colour is regarded not simply as a measure of the spectral composition of a light, but rather as a property of physical objects. This requires quantifying colour in new ways and dropping two of the most widespread procedures: the splitting of colour into *independent* chromatic and achromatic components, and, the scaling of colour distances using discriminability.

Most work in biological signaling treats animal colour vision as a means to discriminate between spectra, and is primarily concerned with the magnitudes of colour differences as measured by ‘just noticeable differences’ (JNDs) [Kelber et al., 2003, Kemp et al., 2015, Olsson et al., 2017]. Accordingly, one might predict that a colour signal will attract a receiver’s attention when it differs strongly from the background [Gittleman and Harvey, 1980], or the pattern itself has a high contrast [Rowe and Guilford, 1996, Aronson and Gamberale-Stille, 2008]. In its everyday use ‘colour’ is not a relative or relational term like ‘contrast’, but is absolute, and part and parcel of object recognition. We do not say “the tennis ball is more yellow than the court,” but “the tennis ball is yellow”. Light reflected from a surface depends upon its chemical composition and (nanoscale) physical structure. Colour vision yields information about these properties. We speak of red faces, blue tea-mugs and so forth, and other animals may be similar. It follows that animals could find particular colours significant because they are characteristic of particular kinds of object.

The significance of a colour might be related to the specific coloration mechanism – for example if a pigment is costly to produce [Olson and Owens, 1998] – or to associations with particular

beneficial or harmful objects [Endler and McLellan, 1988, Endler and Basolo, 1998, Palmer and Schloss, 2010]. In contrast with, and complementary to these approaches, we look here more broadly at the relationship between the composition of materials, their reflectances, and their colours. Perhaps surprisingly, we find that there are aspects of object colour that are consistent between different observers, which can be linked to underlying physical properties that are relevant to the psychology of an organism.

In particular, we analyse a property of colour we term ‘vividness’. Highly vivid colours for one organism are highly vivid for any other with the same or more number of photoreceptor types (rules 1 and 2). We then show that highly vivid colours are more informative, in that they correspond to fewer materials, and that those materials will be purer and more ordered than less vivid materials. From this we argue that psychological salience of these colours would be evolutionarily adaptive.

1.1 Object Colour Solids and Vividness

Objects are seen by reflected light, and the gamut of all possible reflection spectra can be represented by their locations within in a Cartesian space known as an object colour solid whose axes correspond to photoreceptor excitations relative to the illumination spectrum (Figs 1, 2 Vorobyev [2003], Koenderink [2010]). Colour solids include the three main aspects of colour namely hue, saturation and brightness, whereas the more familiar chromaticity diagrams, such as Maxwell’s triangle discount brightness. An additional difference lies in the nature of the gamut boundaries. In a chromaticity diagram the boundary is defined by monochromatic spectra and (for trichromats) the purple line. As a monochromatic reflection contains negligible light highly saturated (or pure) colours are dark. By comparison object colour solid boundaries include both black (zero reflection) at the origin, and white corresponding to maximal reflection at all visible wavelengths, with intermediate the boundary surface being well approximated by spectral step-functions (Fig 2). Such spectra can be bright, and are more nearly physically realizable by natural pigmentation.

Following the logic that colour refers primarily to the physical properties of reflective materials (or objects) we propose a measure of colour within the object colour solid which we call vividness. Vividness resembles colorimetric parameters such

as purity or saturation, but achromatic colours – black and white – and light unsaturated colours can be highly vivid. We show mathematically and empirically that the vividness of reflectance spectra is well correlated between different types of colour vision. Consideration of the relationship between vividness and the physical processes that generate object-colour demonstrates that orderly nanostructures or pure pigments are typically more vivid than their less orderly or pure materials. As order emerges against entropic tendencies, vivid or ‘bright’ [Hamilton and Zuk, 1982] coloration is indicative of a functional role, and therefore more “meaningful”, and hence such colours worthy of greater attention.

2 Modelling

We start with an account of why the achromatic colours black and white are maximally vivid, leading to a general model which includes chromatic colours.

2.1 Black and White

Models of colour vision and colour appearance usually treat chromaticity, which combines hue and saturation, as qualitatively distinct from brightness or luminance. This distinction is grounded in physiology and psychophysics [Livingstone and Hubel, 1988, Osorio and Vorobyev, 2005], yet black and white are ‘colours’ in ordinary English usage, and they are common in biological signals. In physiological terms an animal sees black when its spectral photoreceptors all have a low excitation, and white when they all have a high excitation. Hence, spectra that look black have low intensity at all visible wavelengths and spectra that look white have high intensity. As there is only one perfectly black spectrum, and a perfectly white surface must reflect maximally at all wavelengths it follows that spectra responsible for black and white are the same for all observers that share the same range of visible wavelengths, with minimal or maximal excitation across that range. These spectra can be designated as ‘extreme’, because they are at the limits of receptor excitation achievable by a surface.

Unlike black and white, intermediate reflectance spectra can produce different excitations according to the particular set of photoreceptors in a given eye. For the simplest case of two eyes each with a single type of photoreceptor, but

tuned to different wavelengths, a stimulus of intermediate intensity can give different receptor responses in the two monochromatic observers. Consequently the appearance of greys for different types of monochromat is less predictable than are black and white. Similarly, for any given monochromatic eye only one spectrum can give black or white, but many different spectra can produce indistinguishable intermediate responses, a phenomenon which is known as colour metamerism [Logvinenko, 2009].

Black and white, produced by extreme spectra, are maximally vivid colours, while greys have lower vividness.

The best known geometric representations of colour are chromaticity diagrams in which desaturated colours lie near the centre and the most saturated colours at the extremities, with the angle around the centre specifying hue.

One such diagram is Maxwell’s triangle [Maxwell, 1860]. Using a method based on colour mixing, Maxwell specified the colours of monochromatic lights as mixtures of red, green and blue primaries. When the brightness is ignored these results can be drawn in a triangle, with pure primaries at each corner, and mixtures within. The line of all the monochromatic lights, the “monochromatic locus” or “spectral line” forms a rounded Λ shape along two of the edges, the ends joined by a “purple line”. The fractional distance from the centre to the edge is then a measure of spectral purity, which is related to the colour’s perceived saturation [Wyszecki and Stiles, 2000].

For colour vision of non-human species similar diagrams are typically based on photoreceptor spectral sensitivities [Renoult et al., 2017]. For an eye with n spectral types of photoreceptor (contributing independently to colour vision) the chromaticity diagram is $n - 1$ dimensional. Chromatic spaces are good for describing the colours of lights, but less suitable for reflectance spectra, because a *reflectance* can only approach the boundary by reducing the amount of reflected light, making the colour darker – maximal spectral purity is black! Hence chromaticity diagrams typically exaggerate differences between dark colours.

Object colour solids are so-named because they are appropriate for representing ‘object’ or reflectance spectra. They are useful in colour reproduction and the formulation of dyes and pigments, because the available gamut can be compared to the colour range visible to the human eye. The

axes of colour solids [Wyszecki and Stiles, 2000, Schrödinger, 1920, Vorobyev, 2003, Koenderink, 2010] correspond to photoreceptor excitations (or a similar set of primaries) normalised to the illumination intensity. As photoreceptor spectral sensitivities overlap, colour solids do not fill the space defined by the axes, but are roughly ellipsoidal with two pointed corners (Fig 1). Monochromatic spectra lie an infinitesimal distance from the origin which is black, and maximal reflectance (white) is at the opposite vertex. Humans have three types of cone photoreceptor, and hence a 3-dimensional object colour space, but the same geometrical principles apply to any type of colour vision: for most mammals, which are dichromats, the space is 2-dimensional, whereas the spaces of birds are probably 4-dimensional [Kelber et al., 2003, Vorobyev, 2003]. As there are both empirical and mathematical relationships between these spaces we refer to all of them as colour solids.

2.2 Colour Solids and Vividness

We now more formally define the colour solid and vividness. For a given set of photoreceptor spectral sensitivities ($s_i(\lambda)$, $i \in 1 \dots n$) and an illuminant ($l(\lambda)$), reflectance spectra can be organised into an geometric object known as the *object colour solid* [Schrödinger, 1920, Wyszecki and Stiles, 2000], shown in figure 1. This object is formed from the colours (a vector of photoreceptor quantum yields, ($q_1 \dots q_n$)) associated with all theoretically possible reflectance spectra: that is all distributions of reflectance values between zero and one, taken over visible wavelengths (Λ):

$$q_i = \int_{\Lambda} s_i(\lambda) l(\lambda) r(\lambda) d\lambda \quad (1)$$

To obtain the colour object solid, photoreceptor responses (q_i) are normalised to the quantum yield of a perfectly reflecting surface¹ ($q_1^{\max} \dots q_n^{\max}$).

$$x_i = q_i / q_i^{\max} \quad \text{where} \quad q_i^{\max} = \int_{\Lambda} s_i(\lambda) l(\lambda) d\lambda \quad (2)$$

This means that black has x coordinates of $(0, 0, \dots)$ and white has coordinates $(1, 1, \dots)$. Throughout we express the x coordinates in terms of a relative quantum yield function $f_i(\lambda) = s_i(\lambda) l(\lambda) / q_i^{\max}$:

$$x_i = \int_{\Lambda} f_i(\lambda) r(\lambda) d\lambda \quad (3)$$

¹The maximal reflectance could instead be set according to a standard such as BaSO₄ pellet – this does not affect the arguments we make here (see appendix SI 1.3) and should, in fact, result in a more practical measure of vividness.

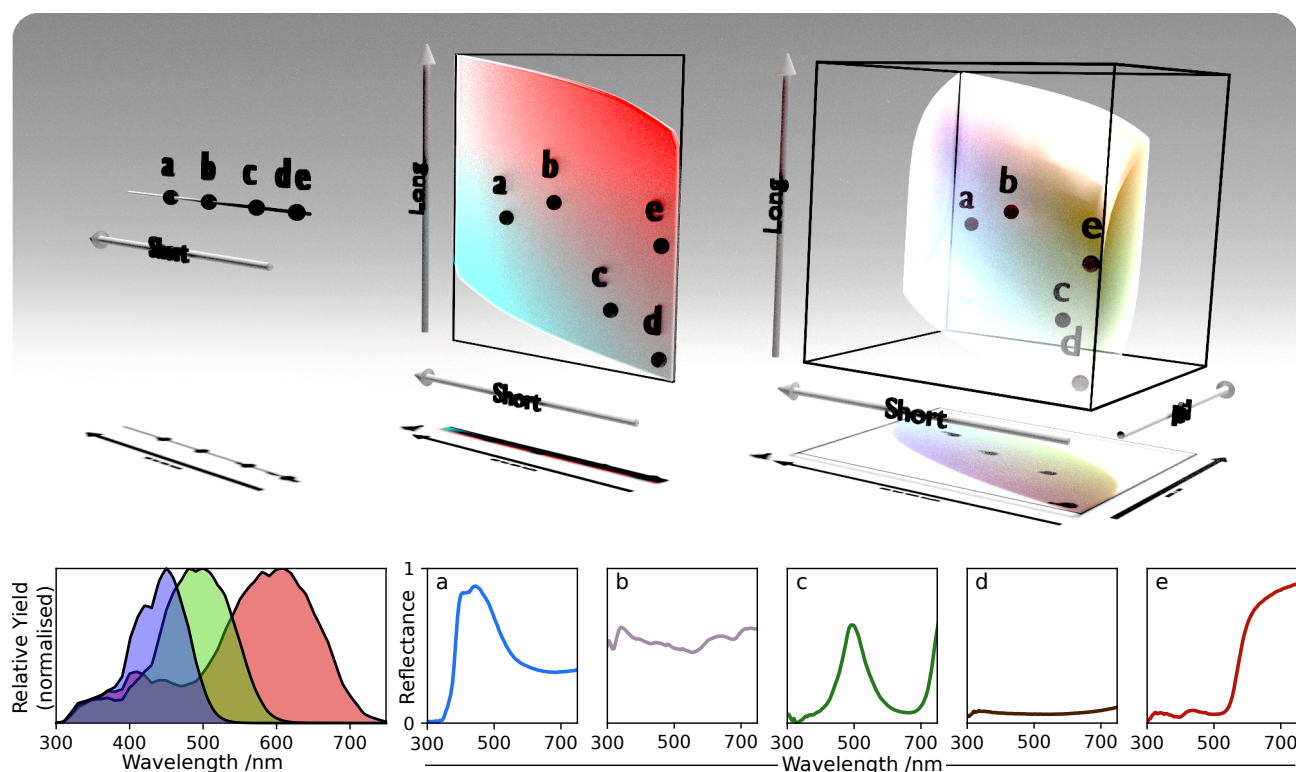


Figure 1: The Object Colour Solid *Top, left to right*: Colour solids for a monochromat, a dichromat and trichromat. *Bottom left*: Relative quantum yield curves (spectral sensitivity multiplied by the illuminant and normalised). The monochromat has just the short wavelength sensitive photoreceptor class, the dichromat contains the short, and the long wavelength sensitive photoreceptor classes, the trichromat contains all of them. Each ‘expands’ the solid outwards into the new dimension. Note how the positions of colours (a-e) in the pre-existing dimensions do not change as more photoreceptor classes are added. *Bottom right*: The natural reflectance spectra responsible for colours (a-e).

The object colour solid is convex, and lies within a unit n -cube (Fig 1). It is pointed at the diagonally opposite black and white corners, but the other corners and edges are smoothed due to the spectral overlap of the photoreceptors. Except for the boundaries, every point in the solid maps to more than one reflectance spectrum, corresponding to colour metamerism [Logvinenko, 2009]. The boundaries represent reflectances with the highest spectral purity for a given luminance. Unlike chromaticity diagrams, the colour solid therefore accounts for the trade-off of saturation against luminance, so light colours can lie on the boundary. This accords with the intuition that we do not see light colours as necessarily less pure than dark.

The exact calculation of the colour solid is rather involved. Numerical solutions can be obtained by dynamic programming [Wyszecki and Stiles, 2000] or our own method (see SI 2 for details of both), but for the current purpose it is better to start with Schrödinger’s approximation [Schrödinger, 1920, Vorobyev, 2003], which

uses spectra formed by step changes in intensity between zero and one. Figure 2 illustrates Schrödinger’s spectra. The number of step changes depends on the number of spectral classes of photoreceptor: for an eye with n spectral classes the maximum number of step changes needed to approximate the colour solid boundary is $n - 1$. For dichromats, they are single steps, the two series (step up and step down) ranging from black through red, orange, then yellow to white, and from black through blue then cyan to white. For trichromats there are two steps in the visual range, while for a tetrachromat the Schrödinger spectra include those that have a reflectance of 0 up to some wavelength λ_1 then 1 until another, λ_2 , then 0 until λ_3 then 1, and in addition their inversions, i.e. those that have values of 1 then 0 then 1 then 0. Whilst the Schrödinger spectra are only approximations to the extreme spectra that lie on the boundary of the solid, the extreme spectra, like Schrödinger spectra, will only ever have either maximal (1) and minimal (0) intensity at

every wavelength²

Schrödinger's approximation of boundaries of the object colour solid holds best when photoreceptor spectral sensitivities have a single peak. Some receptor spectral sensitivities are bimodal (e.g. Fig 3), in which case Schrödinger's spectra typically lie slightly inside the boundaries of the corresponding colour solids. Such deviations are in practice small, because photoreceptor spectral sensitivities are typically smooth with a dominant peak. As long as this approximation holds a conclusion analogous to that for black and white applies; namely that for eyes with the same number of photoreceptor classes, the Schrödinger spectra corresponding to boundary colours are the same. Differences between receptor sensitivities mean that the shape of the object colour solid varies between species, and the locations of individual spectra within the boundary are not directly comparable (Fig 3), but they will be on the boundary nonetheless. It follows that Schrödinger spectra for one eye are Schrödinger spectra for an eye with a larger number of photoreceptor classes: the set of Schrödinger spectra for an n -chromat contains the Schrödinger spectra for an m -chromat, when $m \leq n$.

The tendency for the same spectra to lie at the boundary of the colour solid extends to colours not on the boundary, although, as was the case for black, white and grey, the variability in the position is greater for more central colours (Fig 3). Whilst the theoretical limitations on the position, illustrated in Fig 3, are broad, the spectra that achieve these limits are difficult to realise practically. This is evident in Fig. 4.

In practice, spectra of natural objects are a small subset of physically possible spectra [Maloney, 1986, Osorio and Bossomaier, 1992, Vorobyev et al., 1997], and natural spectra tend to be smoother than Schrödinger spectra [Maloney, 1986] (compare Fig 1 and Fig 2). Also, spectra with multiple transitions within the visible range are unusual, so that higher order Schrödinger spectra are seldom approached in nature; exceptions include some structural colours, such as that found on the nape of the feral pigeon, which can appear greyish to trichromats, but is likely to be vivid for birds [Osorio and Ham, 2002].

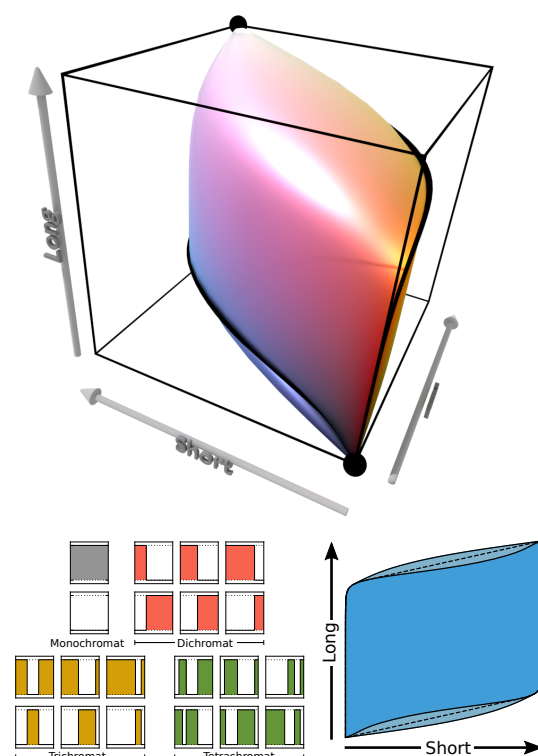


Figure 2: The Schrödinger approximation. *Top:* Loci of Schrödinger spectra for the trichromat in figure 1. The coloured surface shows the trichromatic Schrödinger spectra. Note that for this observer the surface formed by Schrödinger spectra is concave in some places (see main text), this is most easily observed in the blue region. The black ribbon running around the edge of the solid shows the dichromat Schrödinger spectra. The spheres at the tips correspond to the monochromat Schrödinger spectra (i.e. back and white). *Bottom left:* Examples of Schrödinger spectra for monochromats, dichromats and trichromats. The relationship between the colour solid and chromaticity (colour triangle) has been explored by others [Wyszecki and Stiles, 2000, Luther, 1927, Nyberg, 1928]. *Bottom right:* A dichromatic solid (same as in figure 1) showing the locus of the dichromat Schrödinger spectra (innermost shape). The non-convexity is much easier to see in this representation. The convex hull of the Schrödinger spectra's colours is shown as a dotted line, this does not reach the boundary in its entirety either. For this observer the dichromatic Schrödinger spectra have a mean vividness of ≈ 0.943 (all in the range ≈ 0.843 to 1). The colours used to render these solids are only indicative of their true appearance.

2.3 Definition of Vividness

We define vividness as a number ranging from zero at the centre of the colour solid to one at

²When discretised approximations like those in the appendix are used to calculate boundary spectra, a step transition between two discretisation points will appear at an intermediate intensity. This is a discretisation artifact.

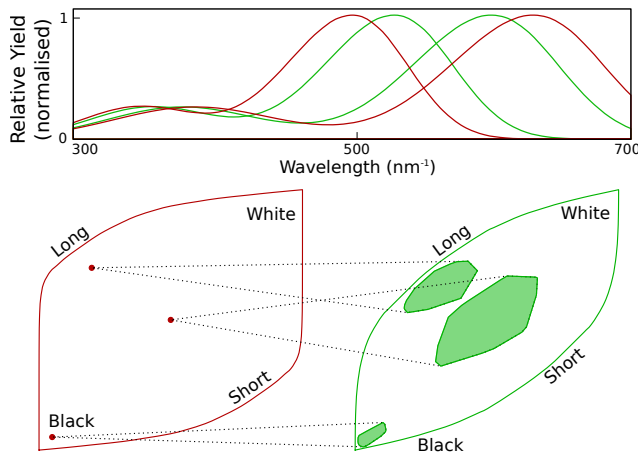


Figure 3: The relationship between the colours in two dichromat object colour solids. Because of metamerism, points on the left correspond to multiple points (areas) on the right. The exact location depends on the particular underlying spectrum. Similarly (though not shown) points on the right correspond to areas on the left. At the edge of the solid the correspondence is one-to-one, and the size of the region in one corresponding to a point in the other increases towards the centre. These mapped areas, like the solids themselves, are based on the assumption that spectra may instantaneously transition between zero and one. Such transitions cannot be realized physically, and the spaces and areas are theoretical bounds. The filled areas on the right are bounded by spectra with a greater number of transitions than the solid. These filled areas are conservatively large estimates of the degree of metamerism that is physically realisable.

The relative quantum yields were produced from an A1 pigment template which has a pronounced β peak [Govardovskii et al., 2000], but the correlation in vividness remains even though the Schrödinger spectrum approximation does not hold exactly for these double-peaked relative quantum yield functions.

the boundary, which is linear with respect to the colour solid coordinates (x_i). For an observer/illumination combination that is described by n relative quantum yield functions ($f_1 \dots f_n$) the vividness of a reflectance r is:

$$V(r; f_1 \dots f_n) = \frac{\|x - 1/2\|}{\|b(x) - 1/2\|} \quad (4)$$

where $b(x)$ is the position of the boundary in direction of the vector $x - 1/2$ from the centre ($1/2$).

The numerator is the Euclidean distance from the centre of the solid. The division by $\|b(x) -$

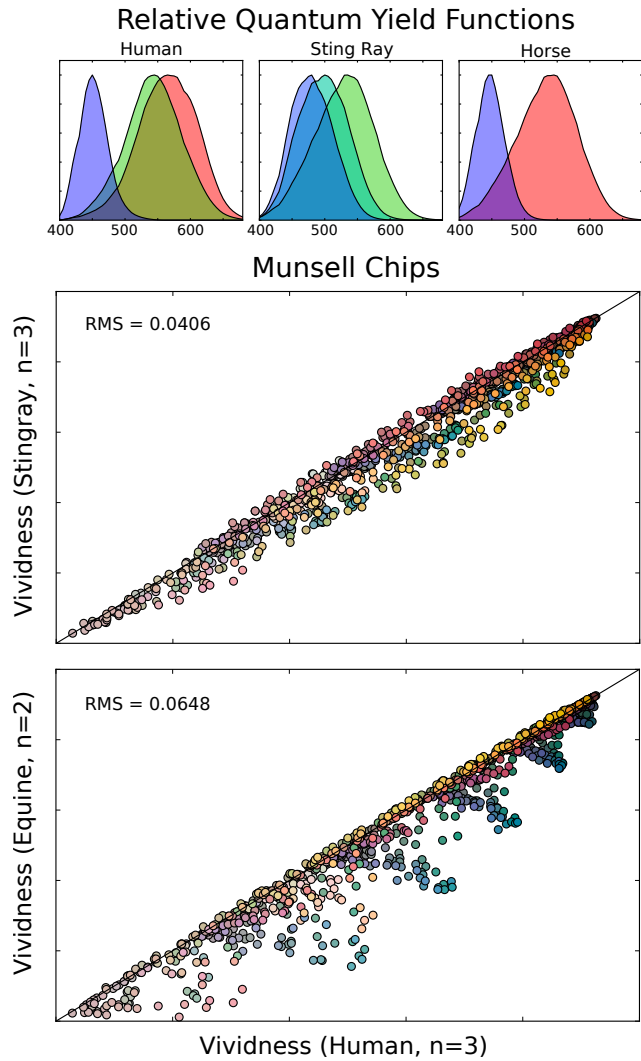


Figure 4: Demonstration of the correlation of vividness between three disparate species. The *top row* shows example relative quantum yield functions with D65 illumination Wyszecki and Stiles [2000] for trichromatic humans Stockman and Sharpe [2000] and sting rays, Hart et al. [2004] and dichromatic horses Carroll et al. [2001]. The scatter plots compare the vividness of the Munsell chips [Munsell et al., 1950, Parkkinen et al., 1989] – a collection of coloured stimuli designed to cover the human colour gamut uniformly – for the three species. The *middle* plot demonstrates *rule 1* – the approximate equality between vividness for the two species with the same number of cone classes (humans and sting rays). The *bottom* plot shows, in addition, the effects of *rule 2* – whilst the approximate equality holds for many cases, some chips which we see as blue or pink can be less vivid for a horse.

$1/2\|$ maintains an invariance between species. It has the effect of adjusting the numerator to reflect

more physical, rather than perceptual, properties. The latter being more difficult to assess for non-human species (and even for humans).

2.4 Properties of Vividness

The properties of colour solids we have discussed so far can be formally expressed with two rules, the first is motivated by our argument above and it is true insofar as it is approximate, and contingent on having “well behaved” relative quantum yield functions, the second is a geometric fact:

Rule 1: Approximate Equality

$$V(r; f_1 \dots f_n) \approx V(r; g_1 \dots g_n)$$

For two observers with equal numbers of spectral receptor classes (n), the vividness of a stimulus is approximately the same for both observers. This rule is motivated by the foregoing discussions, and is corroborated by figures 3 and 4. For more vivid colours the range of discrepancy decreases, and the approximation is better – a consequence of there being “fewer” metamers. This phenomenon can be seen in figure 3.

As this rule concerns the relative quantum yield functions f , which are obtained from both spectral sensitivities and the illuminant, rule 1 is both a statement about changes in photoreceptor sensitivities and in illumination.

Rule 2: Monotonicity with Dimensionality

$$V(r; f_1 \dots f_n) \leq V(r; f_1 \dots f_n, f_{n+1})$$

For a given illumination, an increase in the number of receptor types will result in an increase in vividness, so that colours are more vivid for species that have a large number of photoreceptor classes (e.g. birds) than for those with fewer (e.g. mammals). This effect can be observed in figure 4, and a proof is given in SI 1.

These rules³ allow one to describe the relationship between any two observers.

³For a fully formal approach we require another rule stating that V is not dependent on the order in which the relative quantum yield functions are specified.

As vividness increases the constraints on the variety of spectra that can realize the colour become increasingly restrictive, until, at the boundary of the solid there is a unique spectrum. As there are “more”⁴ spectra that correspond to each of them, less vivid colours are more prone to metamerism. This rule can be compared to the metamer mismatching transformation that occurs under variable illumination [Tokunaga and Logvinenko, 2010].

Colour *purity* [Wyszecki and Stiles, 2000] resembles vividness for the chromatic plane, and has fairly similar properties. This is because the chromaticity space is a scaled cross section of the colour solid at fixed luminance as it goes to zero (i.e. a slice of the solid very near black). Vividness also resembles other measures, such as *chroma* [Endler, 1990]. In addition to the difficulties we have already highlighted with chromaticity, purity is not defined for monochromats, is not mathematically well behaved for dichromats, and can be quite complex for tetrachromats and above (see supplementary material SI 3).

2.5 Mixing

Vividness has a fundamental relationship to physical order, as exemplified by the case of conservative mixing. That is to say, where two colours, x_a and x_b , are mixed in a ratio of $z_a : z_b$ and the resulting colour is:

$$x_{\text{mixed}} = \frac{z_a x_a + z_b x_b}{z_a + z_b} = k x_a + (1 - k) x_b$$

where $k = z_a / (z_a + z_b)$ and the various values of k correspond to various positions on the line segment between x_a and x_b . This holds regardless of the spectra that produce x_a and x_b .

The distance of points on this line segment are closer to (or the same distance from) the centre of the solid than the more distant of x_a and x_b . If the solid were a perfect sphere (i.e. if the distance to the boundary were fixed) we could conclude directly that vividness of the mixture was smaller

⁴Usually there is only one spectrum corresponding to a point on the boundary of the solid, but there might be more in cases where the surface is perfectly flat. Inside the boundary, or at points on the boundary without non-zero bounded curvature, there is an infinite number. Measuring the absolute sizes of the sets of spectra corresponding to a single colour is non-trivial because there is no ‘infinite-dimensional’ analogue to the Lebesgue measure. In this work we avoid this problem by always comparing between two different finite dimensional colour solids ($\subseteq [0, 1]^n$), rather than looking for a way of measuring spectrum-space ($[0, 1]^\infty$).

than that of the colours being mixed. As the solid is not a sphere, we must make a further observation: where the surface is (strictly) convex, a line segment connecting two points on the surface of the solid passes entirely within its volume (for a proof see supplementary material SI 1). The colour solid’s boundary colours have the distinguishing feature that they cannot be made by a mixing other object colours.

The rule applies to all observers, in spite of the fact that numerical values of V for a particular object colour may well be different.

Rule 3: Convexity of Mixing

$$V(kr_1 + (1 - k)r_2; f_1 \dots f_n) \leq \max\{V(r_1; f_1 \dots f_n), V(r_2, f_1 \dots f_n)\}$$

for all $k \in [0, 1]$. Mixing two colours results in a colour that is less vivid than the most vivid of the two (and perhaps less than both). Rule 3 can be written in a more general form, for a mixture of multiple colours, as:

$$V\left(\sum_i k_i r_i; x_1 \dots x_n\right) \leq \max\{V(r_i; x_1 \dots x_n)\}$$

with $k_i \in [0, 1]$ and $\sum_i k_i = 1$.

This expression describes a phenomenon familiar to anyone who has mixed paints, once a duller colour is mixed into a more vivid one there is no way to recover the original vividness except by adding an even more vivid paint.

This is not a property unique to vividness, but we must bear in mind the significance of this rule for our argument: Without this rule we would only have mathematical results about the geometry of colour solids, but with this rule, we can talk about the physical properties of vividly coloured materials, and thereby talk about their ecological function and how they are perceived (in a Gibsonian sense [?], at least).

3 Discussion

Object colour solids are useful representations of the colours of reflective surfaces, as opposed to lights. The mathematical properties of these spaces along with empirical evidence (Fig 4) leads to three main findings relevant to the evolution of colour in communication signals.

Firstly, we define a measure of colour, vividness, which corresponds to the distance of a colour from the (grey) centre of the solid, and show mathematically that a spectrum which is highly vivid (i.e. near the boundary of the colour solid) for one type of colour vision will be highly vivid for any type that has the same number or more spectral types of photoreceptor. We find empirically that the vividness of natural reflectance spectra is correlated for eyes with different sets of photoreceptors (Figs 4 and 3).

Secondly, colours on the boundary of the solid are attributable to a single reflectance spectrum, and the number of spectra that map to a given point in the colour solid increases towards the centre, so that many spectra can look ‘mid-grey’. Highly vivid colours are therefore more likely to be associated with a specific physical cause (i.e. material), and they will be less easily reproduced by alternative means than less vivid colours. Thirdly, mixing colours inevitably reduces the vividness of the more vivid colour, so that pure materials tend to be more vivid than mixtures. This is why vacuum-cleaner dust is greyer than home furnishings. For structural colours increasing regularity of nano-structure increases vividness, and mixing of pigments can only render them less vivid.

Due to entropy order does not arise by chance in nature, so vivid colours are indicative of some functional role. This need not be as a signal; the vivid colours of leaves and blood are due to high concentrations of light harvesting and oxygen transport pigments. This is not to say that and a dull coloured tissue cannot have a specific function, the implication only works one way – vividness requires some kind of order, but order does not necessarily result in vivid colour. Nonetheless, if an object has a functional role it is *a priori* worthy of attention, which is a requirement for any signal.

Historically, vivid coloration has been associated with the phenomenon of life, and this in turn has been related to thermodynamics. In *Tropical Nature* [Wallace, 1878] Wallace argues that colourfulness is a consequence of “vital energy”, and hence a natural attribute of living organisms. While Schrödinger proposed in his book *What is Life?* [Schrödinger, 1944], that life is characterised by order away from chemical equilibrium, that is by being non-entropic. Here we have seen why vivid colours are indeed likely to be associated with a system – such as life – that counters the effects of entropy.

3.0.1 Aposematism and Mimicry

The strong contrasts and vibrant colours of aposematic displays illustrate some of our findings. Vivid colours are salient, so they have the potential to promote both innate and learnt responses, and they will be seen consistently by a broad range of viewers, so they need not be predator-specific. For instance in tropical forests predators of insects include monochromatic strepsirrhine primates, dichromatic mammals and snakes, trichromatic primates and amphibians, and tetrachromatic birds and lizards [Kelber et al., 2003]. A consequence of rule 2 is that the achromatic extreme colours – black and white – will be effective for all receivers, those approximating the dichromatic Schrödinger spectra (which we would see as two series black, red, yellow, white and black, blue, cyan, white) would be effective for dichromats and above, the trichromatic series which adds purples and greens for trichromats and above, and so on. Similar principles apply in marine environments where predator colour vision ranges from monochromacy in cephalopods through di, tri and tetrachromacy in various fish to the multispectral system of stomatopods [Marshall et al., 2015].

A further benefit of vivid defensive signals arises because as vividness increases copying a colour entails more exact matching of the spectrum. A defended aposematic model may in principle outmanoeuvre a Batesian mimic by adopting more vivid colours outside the mimic’s physiological ‘gamut’ [Franks et al., 2009, Briscoe et al., 2010, Bybee et al., 2011] and conversely there might be a pressure for Müllerian mimics to share less vivid colours.

3.1 Modelling of Colour in Biological Signals

Investigation of colour signaling by animals and plants often starts by modelling photoreceptor responses to reflectance spectra. Receptor responses do not directly specify colour differences or colour appearance, which normally requires recourse to psychophysical models [Kelber et al., 2003, Kemp et al., 2015]. Models based on chromaticity assume that lightness (or luminance) is discounted, and although some add achromatic contrast as a separate parameter [Siddiqi et al., 2004, Olsson et al., 2017], all such models can lead to difficulties. For example they predict that dark colours are unrealistically distinct, and there may be an

implicit assumption that the strong achromatic contrasts, which are present in many signals are either irrelevant or have a qualitatively different function from chromatic components.

Object colour solids represent the full gamut of colours visible to an eye, and importantly offer a natural means of representing colour as a property of reflective materials rather than spectral lights. It is straightforward to define the locations of reflectance spectra in the object colour solid, and hence to explore broader questions about the gamuts of colour signals directed at various types of receivers (cf [Osorio and Vorobyev, 2008]). Vividness is a simple and well defined measure of colour, within a colour solid which can be related to the physical properties of coloured materials, and gives insight into how and why animals with diverse visual systems might evaluate colour. Of course it remains an empirical question whether vividness is a useful measure of colourfulness. One could test whether vividness should predict attention or salience better than colour saturation or purity, or to a scale based on colour distances measured in terms of just noticeable differences within the colour solid.

4 Conclusion

In *The Origin of Species* Darwin writes. “...the belief that organic beings have been created beautiful for the delight of man, ... has been pronounced as subversive of my whole theory, ...” [Darwin, 1859]. Contemporary literature rarely addresses this question directly, but offers a range of accounts of why certain colours or patterns should be attractive to animals [Zahavi, 1975, Grafen, 1990, Guilford and Dawkins, 1991, Anderson, 1994, Johnstone, 1995, Endler and Basolo, 1998]. Some refer to the nature of the sensation, and postulate general aesthetic principles, whereas others highlight the specific value of a stimulus. For example it is debated whether animals use particular colours in courtship displays because they resemble objects of value such as food items [Allen, 1879, Endler and Basolo, 1998] or because the colours are indicative of the quality of a potential mate [Hill, 1991]. Here we find that these distinctions may be blurred, because consideration of the physical causes of the coloration suggests that certain colours should *a priori* be significant regardless of specific associations with objects of relevance to the animal, or indeed its particular type of colour vision. There is no need

to invoke evaluative judgements [Prum, 2012] or co-evolutionary processes [Fisher, 1930] such as those associated with sexual selection. In addition the laws of thermodynamics mean that there is an immediate cost to producing pure materials and hence vividly coloured tissues, but this is not part of the main thrust of our argument, which is instead about a general psychological bias towards vivid colours.

Models of visual salience to humans typically include components that are akin to vividness [Niebur and Koch, 1996], and this kind of low level prediction can be compared and contrasted by higher level (evaluative) theories based on asking subjects about the aesthetic value of a colour. Palmer and Schloss’s [Palmer and Schloss, 2010] valence theory proposes that humans prefer colours associated with desirable objects, to those associated with decay, excrement and so forth. Given that decomposition and biological waste tend to be chemical mixtures, with low vividness, whereas pure materials tend to be more vivid it would be interesting to test whether vividness predicts colour preference as well as the valence. At a more practical level, we can offer some assurance to field biologists that it is reasonable to generalise from our own colour perception to that of other animals, despite their physiological differences [Bennett et al., 1994].

Data Availability

The code used to generate figures here have been made available, along with documentation and usage examples, as an open source project currently hosted at <https://github.com/lucas-wilkins/lemonsauce>. The numerical methods are also detailed in the supplementary information.

The Munsell chip data are available from the Spectral Color Research Group, University of Eastern Finland at <http://www.uef.fi/web/spectral/-spectral-database> [Parkkinen et al., 1989].

Acknowledgments

We would like to thank Mary Stoddart, Almut Kelber and Justin Marshall for their comments, and Justin Marshall also for the spectra used in figure 1.

References

- G. Allen. *The Colour Sense*. Wentworth Press, 1879.
- M. Anderson. *Sexual Selection*. Princeton University Press, 1994.
- M. Aronsson and G. Gamberale-Stille. Domestic chicks primarily attend to colour, not pattern, when learning an aposematic coloration. *Animal Behaviour*, 75(2):417–423, 2008.
- A. T. D. Bennett, I. C. Cuthill, and K. J. Norris. Sexual selection and the mismeasure of color. *The American Naturalist*, 144(5):848–860, 1994.
- A. D. Briscoe, S. M. Bybee, G. D. Bernard, F. Yuan, M. P. Sison-Mangus, R. D. Reed, A. D. Warren, J. Llorente-Bousquets, and C.-C. Chiao. Positive selection of a duplicated uv-sensitive visual pigment coincides with wing pigment evolution in heliconius butterflies. *Proceedings of the National Academy of Sciences*, 107(8):3628–3633, 2010.
- S. M. Bybee, F. Yuan, M. D. Ramstetter, J. Llorente-Bousquets, R. D. Reed, D. Osorio, and A. D. Briscoe. Uv photoreceptors and uv-yellow wing pigments in heliconius butterflies allow a color signal to serve both mimicry and intraspecific communication. *The American Naturalist*, 179(1):38–51, 2011.
- J. Carroll, C. J. Murphy, M. Neitz, J. N. Ver Hoeye, and J. Neitz. Photopigment basis for dichromatic color vision in the horse. *Journal of Vision*, 1(2):2–2, 2001.
- C. Darwin. *On the Origin of Species by Means of Natural Selection*. John Murray, 1859.
- J. A. Endler. On the measurement and classification of colour in studies of animal colour patterns. *Biological Journal of the Linnean Society*, 41(4):315–352, 1990.
- J. A. Endler and A. L. Basolo. Sensory ecology, receiver biases and sexual selection. *Trends in Ecology and Evolution*, 13(10):415–420, 1998.
- J. A. Endler and T. McLellan. The processes of evolution: toward a newer synthesis. *Annual Review of Ecology and Systematics*, 19:395–421, 1988.
- R. A. Fisher. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press, 1930.

- D. W. Franks, G. D. Ruxton, and T. N. Sherratt. Warning signals evolve to disengage Batesian mimics. *Evolution*, 63(1):256–267, 2009. ISSN 1558-5646.
- J. L. Gittleman and P. H. Harvey. Why are distasteful prey not cryptic? *Nature*, 286(5769):149–150, 1980.
- V. I. Govardovskii, N. Fyhrquist, T. Reuter, D. G. Kuzmin, and K. Donner. In search of the visual pigment template. *Visual neuroscience*, 17(04):509–528, 2000.
- A. Grafen. Biological signals as handicaps. *Journal of Theoretical Biology*, 144(4):517 – 546, 1990.
- T. Guilford and M. S. Dawkins. Receiver psychology and the evolution of animal signals. *Animal behaviour*, 42(1):1–14, 1991.
- W. Hamilton and M. Zuk. Heritable true fitness and bright birds: a role for parasites? *Science*, 218(4570):384–387, 1982. ISSN 0036-8075. doi: 10.1126/science.7123238. URL <http://science.sciencemag.org/content/218/4570/384>.
- N. S. Hart, T. J. Lisney, N. J. Marshall, and S. P. Collin. Multiple cone visual pigments and the potential for trichromatic colour vision in two species of elasmobranch. *Journal of Experimental Biology*, 207(26):4587–4594, 2004.
- G. E. Hill. Plumage coloration is a sexually selected indicator of male quality. *Nature*, 350(6316):337, 1991.
- R. A. Johnstone. Sexual selection, honest advertisement and the handicap principle: Reviewing the evidence. *Biological Reviews*, 70(1):1–65, 1995.
- A. Kelber, M. Vorobyev, and D. Osorio. Animal colour vision – behavioural tests and physiological concepts. *Biological Reviews*, 78(1):81–118, 2003.
- D. J. Kemp, M. E. Herberstein, L. J. Fleishman, J. A. Endler, A. T. Bennett, A. G. Dyer, N. S. Hart, J. Marshall, and M. J. Whiting. An integrative framework for the appraisal of coloration in nature. *The American Naturalist*, 185:705–724, 2015.
- J. Koenderink. *Color for the Sciences*. MIT Press, 2010. ISBN 9780262335171.
- M. Livingstone and D. Hubel. Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, 240(4853):740–749, 1988.
- A. D. Logvinenko. An object-color space. *Journal of Vision*, 9(11):5, 2009.
- R. Luther. Aus dem Gebiet der Farbreizmetrie. *Ztschr. Techn. Phys.*, 8:997 – 1012, 1927.
- L. T. Maloney. Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *JOSA A*, 3(10):1673–1683, 1986.
- J. Marshall, K. L. Carleton, and T. Cronin. Colour vision in marine organisms. *Current Opinion in Neurobiology*, 34:86–94, 2015.
- J. C. Maxwell. On the theory of compound colours, and the relations of the colours of the spectrum. *Philosophical Transactions of the Royal Society of London*, 150:57–84, 1860.
- A. H. Munsell et al. *Munsell book of color*. Munsell Color Co., 1950.
- E. Niebur and C. Koch. Control of selective visual attention: Modeling the ‘where’ pathway. *Neural Information Processing Systems*, 8:802–808, 1996.
- N. Nyberg. Zum Aufbau des Farbenkörpers im Raume aller Lichtempfindungen. *Ztschr. Phys.*, 28:406–419, 1928.
- V. A. Olson and I. P. Owens. Costly sexual signals: are carotenoids rare, risky or required? *Trends in Ecology and Evolution*, 13:510–514, 1998.
- P. Olsson, O. Lind, and A. Kelber. Chromatic and achromatic vision: parameter choice and limitations for reliable model predictions. *Behavioral Ecology*, 29:273–282, 2017.
- D. Osorio and T. R. Bossomaier. Human cone-pigment spectral sensitivities and the reflectances of natural surfaces. *Biological cybernetics*, 67(3):217–222, 1992.
- D. Osorio and A. D. Ham. Spectral reflectance and directional properties of structural coloration in bird plumage. *Journal of Experimental Biology*, 205(14):2017–2027, 2002.
- D. Osorio and M. Vorobyev. Photoreceptor spectral sensitivities in terrestrial animals: adaptations for luminance and colour vision. *Proceedings of*

- the Royal Society of London B: Biological Sciences*, 272(1574):1745–1752, 2005.
- D. Osorio and M. Vorobyev. A review of the evolution of animal colour vision and visual communication signals. *Vision Research*, 48:2042–2051, 2008.
- S. E. Palmer and K. B. Schloss. An ecological valence theory of human color preference. *Proceedings of the National Academy of Sciences*, 107(19):8877–8882, 2010.
- J. P. Parkkinen, J. Hallikainen, and T. Jaaskelainen. Characteristic spectra of munsell colors. *JOSA A*, 6(2):318–322, 1989.
- R. O. Prum. Aesthetic evolution by mate choice: Darwin’s really dangerous idea. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367(1600):2253–2265, 2012.
- J. P. Renoult, A. Kelber, and H. M. Schaefer. Colour spaces in ecology and evolutionary biology. *Biological Reviews*, 92:292–315, 2017.
- C. Rowe and T. Guilford. Hidden colour aversions in domestic chicks triggered by pyrazine odours of insect warning displays. *Nature*, 383(6600):520, 1996.
- E. Schrödinger. Theorie der Pigmente von grösster Leuchtkraft. *Ann. Phys.*, 4(62):603–622, 1920.
- E. Schrödinger. *What Is Life?* Cambridge University Press, 2012 reprint edition, 1944.
- A. Siddiqi, T. W. Cronin, E. R. Loew, M. Vorobyev, and K. Summers. Interspecific and intraspecific views of color signals in the strawberry poison frog *Dendrobates pumilio*. *Journal of Experimental Biology*, 207:2471–2485, 2004.
- A. Stockman and L. T. Sharpe. The spectral sensitivities of the middle-and long-wavelength-sensitive cones derived from measurements in observers of known genotype. *Vision research*, 40(13):1711–1737, 2000.
- R. Tokunaga and A. D. Logvinenko. Material and lighting dimensions of object colour. *Vision Research*, 50(17):1740 – 1747, 2010.
- M. Vorobyev. Coloured oil droplets enhance colour discrimination. *Proceedings of the Royal Society B: Biological Sciences*, 270:1255–1261, 2003.
- M. Vorobyev, A. Gumbert, J. Kunze, M. Giurfa, and R. Menzel. Flowers through insect eyes. *Israel Journal of Plant Sciences*, 45(2-3):93–101, 1997.
- A. R. Wallace. *Tropical nature, and other essays*. 1878.
- G. Wyszecki and W. S. Stiles. *Color Science: Concepts and Methods, Quantitative Data and Formulae*. Wiley-Interscience, 2000.
- A. Zahavi. Mate selection – a selection for a handicap. *Journal of theoretical Biology*, 53(1):205–214, 1975.

Supplementary Information

SI 1 Proofs

SI 1.1 Convexity and Symmetry

The convexity and symmetry of the colour solid is well established, but it is useful to quickly visit it, as many of the appendices use this fact.

The (weak) convexity of the colour solid means that if colours x and y belong to it, then so do all the colours lying the line between them, i.e. $kx + (1 - k)y$ for all $k \in [0, 1]$. Because x and y are assumed to be colours within the solid, there are corresponding reflectances r_x and r_y bounded between 0 and 1 at all wavelengths.

$$x = \int_{\Lambda} r_x(\lambda) f(\lambda) d\lambda, \quad y = \int_{\Lambda} r_y(\lambda) f(\lambda) d\lambda,$$

The reflectance $kr_x(\lambda) + (1 - k)r_y(\lambda)$ will result in the colour $kx + (1 - k)y$, and as this is a convex combination of the reflectance r_x and r_y this will always have values between r_x and r_y , and therefore between 0 and 1. Thus this colour will always belong to the colour solid.

A similar argument applies to the central symmetry of the colour solid. If x is a point within the solid, we can find at least one reflectance $r_x(\lambda)$ that will generate it (by definition). For every reflectance r_x bounded between 0 and 1 there is a corresponding reflectance $1 - r_x(\lambda)$ which is also bounded between 0 and 1. So, there will be a corresponding colour, $1 - x$, within the solid. We can also see from this that the centre of the solid $1/2$ will be mapped to itself through this operation, and thereby corresponds to the solid's center of inversion.

SI 1.2 Rules 2 and 3

The proofs of rules 2 and 3 can be demonstrated by planar geometry using only the facts that vividness is, by definition, the fractional distance from the centre of the solid to the boundary, and that the colour solid is convex.

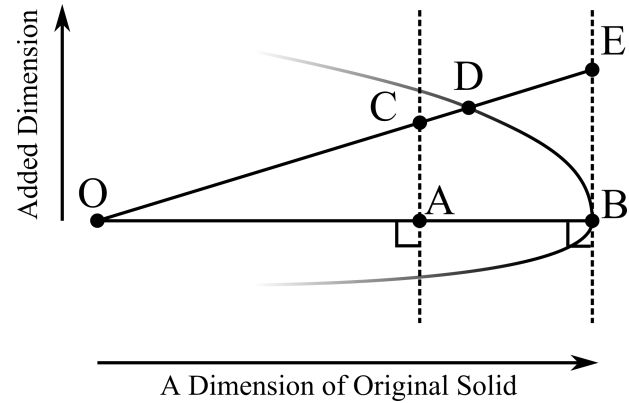
The first of these is a matter of definition, and the second follows from the space of reflectances being convex (essentially an infinite dimensional cube) and the integral $\int_{\Lambda} r(\lambda) f_i(\lambda) d\lambda$ acting as a linear projection of r .

SI 1.2.1 Monotonicity with Dimensionality

The proposition to be shown is:

$$V(r; f_1 \dots f_n) \leq V(r; f_1 \dots f_n, f_{n+1})$$

This states that the vividness in a colour solid either increases or stays the same when an extra dimension (spectral photoreceptor class) is added. This may be interpreted using the following diagram



Left to right is the component of the original solid parallel to the vector from the centre of the solid, O , to the colour A . The corresponding boundary colour is labeled B . The upwards direction shows the added dimension. The value in this dimension does not affect the value in the original dimension, so the position in this new space C is on a vertical line passing through A .

The vividness in the original solid is $\overline{OA}/\overline{OB}$ and becomes $\overline{OC}/\overline{OD}$.

We can then use a long established geometric fact (it is proposition 2 in Book VI of Euclid's Elements) that dividing two edges of a triangle by the same ratio gives points that lie on a line parallel with the third edge (and vice-versa). The line through B and parallel to the one passing through C and A is vertical by virtue of this proposition. This line contains point E , which is the position that D would need to be in for equal vividness. If D is closer to O than E then the vividness is greater.

As B is a boundary point, the convexity of the colour solid means that there are no points to be found further to the right on this diagram, demonstrating the proposition. In summary

$$\begin{aligned} V(r; s_1 \dots s_n, s_{n+1}) &= \overline{OC}/\overline{OD} \\ &\geq \overline{OC}/\overline{OE} \\ &= \overline{OA}/\overline{OB} \\ &= V(r; s_1 \dots s_n) \end{aligned}$$

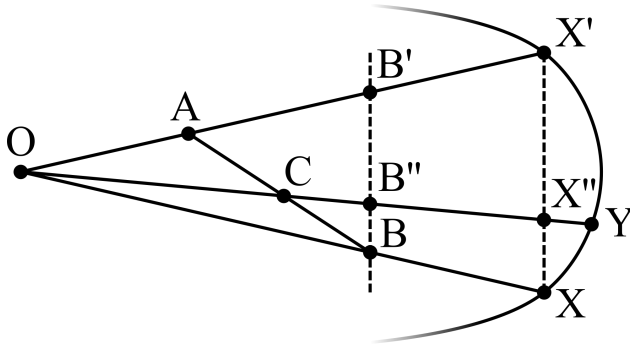
■

SI 1.2.2 Convexity of Mixing

The proposition to be demonstrated is that

$$V(kr_1 + (1 - k)r_2; f_1 \dots f_n) \leq \max\{V(r_1; f_1 \dots f_n), V(r_2; f_1 \dots f_n)\}$$

This can be interpreted in terms of the following diagram of a plane (usually *the* plane) containing the colours of r_1 and r_2 and the centre of the solid:



O is the centre of the solid. A and B correspond to the reflectance spectra r_1 and r_2 with the diagram drawn where B is the more vivid of the two. C is some convex combination of A and B , i.e. a point on the line between them. From the same rule used in the previous section (*Elements* VI, 2) we know that $\overline{OB}/\overline{OX} = \overline{OB'}/\overline{OX'} = \overline{OB''}/\overline{OX''}$. The ratio $\overline{OA}/\overline{OX'}$ is smaller than $\overline{OB'}/\overline{OX'}$ as we have chosen B to be the more vivid of the two.

The diagram shows that as \overline{OA} is (weakly) shorter than $\overline{OB'}$, then \overline{OC} is (weakly) shorter than $\overline{OB''}$. Furthermore, from the convexity of the colour solid it follows that $\overline{OX''}$ is weakly shorter than \overline{OY} . So, we can say that $\overline{OB''}/\overline{OX''} \geq \overline{OC}/\overline{OX''} \geq \overline{OC}/\overline{OY}$. Which is the proposition to be demonstrated as $\overline{OB''}/\overline{OX''}$ is $\max\{V(r_1; x_1 \dots x_n), V(r_2, x_1 \dots x_n)\}$ and $\overline{OC}/\overline{OY}$ is $V(kr_1 + (1 - k)r_2; x_1 \dots x_n)$.

SI 1.3 Adjustment of Minimal and Maximal Reflectance

Throughout this paper we have assumed that the minimal and maximal bounds on reflectance are 0 and 1. We can however work with any minimal or maximal bounds and still have a colour solid.

The colour solid is defined as the set of all colours formed by applying equation 3 to all reflectances r that lie between 0 and 1, we will show that if we replace the condition

$$0 \leq r(\lambda) \leq 1$$

with generalised bounds

$$l(\lambda) \leq r(\lambda) \leq u(\lambda)$$

then we can construct another solid, described by

$$x_i = \int_{\Lambda} r'(\lambda) f'_i(\lambda) d\lambda$$

where each x_i is scaled between 0 and 1 and where the function r' lies between 0 and 1.

Let $l(\lambda)$ be the lower bound of reflectance at each wavelength, and $u(\lambda)$ be the upper bound. Also, let $r'(\lambda)$ be a linearly scaled reflectance relative to those bounds, i.e when $r'(\lambda)$ is 0, the actual reflectance is the lowest possible $r(\lambda) = l(\lambda)$ and when $r'(\lambda)$ is 1 the actual reflectance is maximal, $r(\lambda) = u(\lambda)$. So that

$$r(\lambda) = l(\lambda) + [u(\lambda) - l(\lambda)]r'(\lambda)$$

and the corresponding quantum yield is

$$\int_{\Lambda} l(\lambda) f_i(\lambda) d\lambda + \int_{\Lambda} [u(\lambda) - l(\lambda)] r'(\lambda) f_i(\lambda) d\lambda$$

The first term is a constant offset, so may be ignored, and the second term can be seen as a rescaling of the relative yields, mathematically the same as a change in the illuminant. We can move the bounds into the relative yields and renormalise giving the relative yield functions for a new colour solid

$$f'_i(\lambda) = \frac{[u(\lambda) - l(\lambda)] f_i(\lambda)}{\int_{\Lambda} [u(\lambda) - l(\lambda)] f_i(\lambda) d\lambda}$$

or equally, working with the illuminant $l(\lambda)$ and spectral sensitivities $s_i(\lambda)$,

$$f'_i(\lambda) = \frac{[u(\lambda) - l(\lambda)] l(\lambda) s_i(\lambda)}{\int_{\Lambda} [u(\lambda) - l(\lambda)] l(\lambda) s_i(\lambda) d\lambda}$$

SI 2 Numerical Solutions

■ SI 2.1 Vividness

The most difficult part of calculating vividness is finding the distance from the centre of the solid to the boundary in the direction of the chosen colour. There are two ways that this might be easily computed: One way is to calculate the geometry of the solid's boundary, and then search the geometry for the point where a line extending from the centre through the colour crosses the boundary. This process can be optimised somewhat, but involves a lot of precalculation. Or, instead, one could use a linear programming method. This does not require precalculation, and is the method we recommend.

SI 2.1.1 Calculating Vividness Using Linear Programming

Finding the spectrum that lies on the boundary in a given direction from the centre can be understood as a linear programming problem. Linear programming is the name given to the optimisation of problems that have a linear objective function and linear constraints. The canonical form used for finding numerical solutions⁵ of such problems is:

$$\begin{aligned} \text{Minimise:} \quad & c \cdot x \\ \text{such that:} \quad & Ax \leq b \\ & Dx = e \end{aligned}$$

Where x is the vector to be determined, b and c are vector, and A matrix, parameters. Our task is to put the problem of finding the distance to the boundary into a form that can be understood by standard linear programming routines, such as the simplex method.

In this problem, we find the boundary colour by first finding the spectrum which lies on the boundary of solid. We must work with spectra, not colours, as the colour solid is determined by a constraint on the form of the spectrum.

Inequality Constraints (form of $Ax \leq b$)

If x is a vector representing a reflection spectrum, then we have constraints of the form

$$0 \leq x_i \leq 1$$

which we can write in block matrix form as:

$$\begin{pmatrix} I \\ -I \end{pmatrix} x \leq \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Equality Constraints (form of $Dx = e$)

There are also equality constraints. The colour of the boundary spectrum must lie on a line passing through the reference colour and the centre. The colours (s) lying on this line are given by:

$$s = k(s_{\text{ref}} - 1/2) + 1/2$$

where s_{ref} is the reference colour and k is a variable giving the position along the line. $1/2$ is used here to represent the vector $(1/2, 1/2 \dots)$.

⁵The equality constraints ($Dx = e$) are usually explicit in numerical procedures, but in mathematical treatments they are usually understood as a pair of inequality constraints: $Dx \leq e$ and $-Dx \leq -e$.

This has the form $s = kp + q$, where $p = s_{\text{ref}} - 1/2$ and $q = (1/2, 1/2 \dots)$ are vectors. To get this in the form of $Dx = e$ that is needed to specify the linear programming problem, we need to do a little algebraic manipulation. If we can write the constraint of the colour to a line in the standard form: $Mc = v$, then it is simple to see that when F is a matrix containing the discretised relative quantum yield functions (see equation eqn:discrete), the colour s is given by Fx and we have $MFx = v$, meaning that $D = MF$ and $e = v$.

There are multiple choices of M and v which constrain the colour in exactly the same way. One way of obtaining an M and v is by picking one dimension for which p is non-zero (the 'pivot'), using this to solve for k and substituting the solution back. The result of this procedure is an n -by- $n - 1$ matrix and an $n - 1$ vector. The matrix has can be understood as having ones on the leading diagonal and $-p_i/p_{\text{pivot}}$ in the pivot column, with the pivot row removed. For example, for a tetrachromat pivoting with $i = 2$ we have, the 4-by-3 matrix:

$$M = \begin{pmatrix} 1 & 0 & -p_0/p_{\text{pivot}} & 0 \\ 0 & 1 & -p_1/p_{\text{pivot}} & 0 \\ 0 & 0 & -p_3/p_{\text{pivot}} & 1 \end{pmatrix} \leftarrow \text{row removed}$$

The corresponding vector has values of the form $q_i - (q_{\text{pivot}}/p_{\text{pivot}})p_i$ and, like M , has the pivot row removed.

$$v = \begin{pmatrix} q_0 - (q_{\text{pivot}}/p_{\text{pivot}})p_0 \\ q_1 - (q_{\text{pivot}}/p_{\text{pivot}})p_1 \\ q_3 - (q_{\text{pivot}}/p_{\text{pivot}})p_3 \end{pmatrix} \leftarrow \text{row removed}$$

Objective Function (form of $c \cdot x$)

Lastly, we need the objective function which is to be maximised. This measures how far we are from the centre of the solid. A dot product using colour is the choice because it is signed (unlike the Euclidean distance), and is linear. The function

$$(s - 1/2) \cdot (s_{\text{ref}} - 1/2)$$

is the distance from the centre, projected in the direction of c_{ref} . To get this in the form $c \cdot x$ we rewrite it as:

$$s \cdot (s_{\text{ref}} - 1/2) - 1/2 \cdot (s_{\text{ref}} - 1/2)$$

The second term is a constant. This means that its value does not effect the minimisation and we can ignore it, giving:

$$s \cdot (s_{\text{ref}} - 1/2)$$

After substituting Fx for s , and rearranging using the commutativity of the dot product we find that

$$-c = (s_{\text{ref}} - 1/2)F$$

The minus sign is present because the problem is specified as a minimisation, not a maximisation.

SI 2.2 Colour Solid Geometry

The geometry of the colour solids have been calculated here using a convex pruning method.

Through convexity it is possible to demonstrate that the boundary of the colour solid is formed by reflectances that contain only zeros and ones. Every spectrum can be formed though a convex combination of zero-one spectra, thus any colour must be a convex combination of the colours formed by zero-one spectra.

To calculate the solid numerically, we use discretised relative quantum yield functions, $g_i(j)$, for a set of wavelengths $\lambda_0, \lambda_1, \dots, \lambda_j \dots \lambda_n$.

$$g_i(j) = \int_{\lambda_j}^{\lambda_{j+1}} f_i(\lambda) d\lambda \quad (5)$$

where i indexes the photoreceptor class, and j indexes the wavelength band. The algorithm works by incrementally enumerating all the reflectance spectra that may lie on the surface of the solid, ignoring all those for which this is clearly impossible.

Denote the set of colours defined by non-zero discretised reflectance only in the region $[\lambda_0, \lambda_k]$ by \mathcal{S}_k . This will be uniquely defined by a finite set of points lying on the convex hull of that set, \mathcal{H}_k . The sets \mathcal{S}_k and \mathcal{S}_{k-1} are related to each other by:

$$\mathcal{S}_k = \mathcal{S}_{k-1} \cup \{s + (g_1(k), g_2(k) \dots) | s \in \mathcal{S}_{k-1}\}$$

The first and second terms of the right hand sum correspond to the addition of reflectances in region $[\lambda_{k-1}, \lambda_k]$ with values 0 and 1 respectively. Thus \mathcal{S}_n contains all colours corresponding to discretised reflectances with zero or one everywhere, and has 2^n elements. For resolutions as low as $n = 10$ the full enumeration of this can be problematic. Even more so when it comes to obtaining the convex hull, the algorithms for which which are super-linear in both the number of points and the dimension. To avoid having such large sets we prune away fruitless candidates as we go along.

For calculating the bounding points of the \mathcal{S}_k , only the values \mathcal{H}_{k-1} matter. This is because the

addition of a vector, v , to any point in \mathcal{S}_k will result in a point that is a convex combination of points in $\{s + v | s \in \mathcal{H}_k\}$. If we let \mathcal{C} denote the operation which obtains the points defining the convex hull, then it follows that:

$$\mathcal{H}_k = \mathcal{C}(\mathcal{H}_k \cup \{s + (g_1(k), g_2(k) \dots) | s \in \mathcal{H}_{k-1}\})$$

Which, when applied iteratively, results in a computationally feasible number of points. The speed of this procedure may be increased by including the results of a low resolution version of the calculation (even just the white point). Call these known points \mathcal{K} , and then instead iterate using:

$$\mathcal{H}'_k = \mathcal{C}(\mathcal{K} \cup \mathcal{H}'_k \cup \{s + (g_1(k), g_2(k) \dots) | s \in \mathcal{H}'_{k-1}\}) \setminus \mathcal{K}$$

with:

$$\mathcal{H}_n = \mathcal{C}(\mathcal{H}'_n \cup \mathcal{K})$$

The last application of \mathcal{C} may be omitted if we know that $\mathcal{K} \subseteq \mathcal{H}_n$ – a condition which is desirable for the efficiency of the pruning.

The computational complexity of this algorithm depends on the number of photoreceptor classes, and becomes difficult for observers with more than 5 photoreceptors.

SI 2.3 Using Linear Programming

A linear programming method like the one described in the vividness calculation can be used to obtain a collection of points describing the geometry of the colour solid. To do this, one samples directions around the centre, and calculates the corresponding boundary spectra using linear programming. This method has a disadvantage in comparison with the method below in that it can only resolve fine detail such as points or sharp edges by using a very high density of points.

SI 2.4 Metameric Colours

This algorithm can then be used to calculate the colours for one observer that correspond to the colours for another. Call the discretised relative quantum yield functions for the two observers $a_i(j)$ and $b_i(j)$ respectively, and let $c_i(j) = (a_0 \dots a_{m_a}, b_0 \dots b_{m_b})$ where m_a and m_b is the number of photoreceptor classes for a and b . An $m_a + m_b$ dimensional colour solid can then be constructed for c_i . The set of observer b colours corresponding to an observer a colour can then be calculated as the m_b dimensional cross-section of the $m_a + m_b$ dimensional solid for which the observer a coordinates are constant.

As this involves calculating a $m_a + m_b$ dimensional colour solid, it is only ‘fast’ for $m_a + m_b < 6$. For other cases, a dynamic programming approach is prudent.

SI 3 Purity in Multiple Dimensions

If colour purity – the saturation relative to the maximally saturated colour of a given hue – is generalised to arbitrary dimension one obtains a quantity very similar to vividness. Much as with the extreme spectra of the object colour solid, the border of an arbitrary chromaticity diagram is given by the linear combinations of a few spectra, and as in the colour solid, when the dimension increases, the new boundary will include spectra from the lower dimensional boundary. These spectra are monochromatic, and play the same role as the extreme spectra. But the kind of results that purity yields cannot so easily be employed in developing the kinds of arguments that we developed in the main text, even in approximation.

One difficulty lies in the spectra of the purest colours being monochromatic (or combinations of a few monochromatic spectra). Monochromatic lights are particularly unnatural; an approximately monochromatic *reflectance* has the property of reflecting less and less light as it becomes more pure – reflecting none at the point of perfect monochromaticity.

Another difficulty is that observers with different numbers of photoreceptor classes (different dimensionality) have boundary sets that differ significantly in quality.

For monochromats there is no useful definition of saturation, and therefore no saturation based vividness-like measure exists.

In all other cases, the neatest formulation of the boundary is as the convex hull of all the monochromatic lights. If the monochromatic lights form a convex figure, the convex parts will not be part of the boundary. This means that for dichromats the chromaticity space is bounded by monochromatic two lights, one positioned at the wavelength with the highest ratio in sensitivities and one at the lowest λ_{low} and λ_{high} :

$$\{\delta(\lambda - \lambda_{\text{low}}), \delta(\lambda - \lambda_{\text{high}})\}$$

Unlike the dichromat Schrödinger spectra, these are completely dependent on the spectral sensitivities of the photoreceptors.

For trichromats the boundary often contains a good portion of the monochromatic locus and there is often only one linear connecting sections spanning from a short wavelength to a long one: the ‘purple line’:

$$\mathcal{S}^{\text{mono}} \cup \mathcal{S}^{\text{purple}}$$

The dimensionality of the set of monochromatic lights, and of the purple line is the same.

With tetrachromacy and above, the chromatic boundary becomes dominated by non-spectral (i.e. not monochromatic) spectra, most of the boundary is formed by a $n - 2$ dimensional surface filling in between the the 1 dimensional locus of monochromatic spectra. The surface is formed by linear combinations of $n - 2$ -tuples of monochromatic lights, and typically there will be a small number of dominant monochromatic light that are part of the combinations covering large portions of the surface – in other words, often, a bit of the monochromatic locus “sticks out”. Determining which bits stick out is not particularly simple, and whilst there are rules of thumb that help, it makes framing the arguments we have made here much more complicated. This is aside from the fact that the object colour solid is simply the most appropriate model for reflectance.