

MORE THAN MERE COLOURING:

The Role of Spectral Information in Human Vision

Abstract

A common view in both philosophy and the vision sciences is that, in human vision, wavelength information is primarily ‘for’ colouring, for seeing surfaces and various media as having colours. In this article we examine this assumption of ‘colour-for-colouring’. To motivate the need for an alternative theory, we begin with three major puzzles from neurophysiology, puzzles that are not explained by the standard theory. We then ask about the role of wavelength information in vision *writ large*. How might wavelength information be used by any monochromat or dichromat and, finally, by a trichromatic primate with object vision? We suggest that there is no single ‘advantage’ to trichromaticity but a multiplicity, only one of which is the ability to see surfaces, etc. as having categorical colours. Instead, the human trichromatic retina exemplifies a scheme for a general encoding of wavelength information given the constraints imposed by high spatial-resolution object vision. Chromatic vision, like its’ partner, luminance vision, is primarily *for seeing*. Viewed this way, the ‘puzzles’ presented at the outset make perfect sense.

PART I: REFRAMING THE PROBLEM

I. INTRODUCTION

What is the function of human colour vision? For most vision researchers, both philosophers and scientists, the answer is obvious, nay, a tautology: Colour vision is ‘for’ seeing the colours, whether or not colour is a genuine property of the external world. We have colour vision to see the sky as blue, apples as red, and Heineken beer bottles as translucent green. With this simple answer in hand, the main evolutionary questions about colour follow: what selective pressures during the evolution of human vision lead to the emergence of human colour vision? That is, what in general were the colours ‘good for’? And why do we see the very colours that we do? That is, why did *these* colours, the reds, blues, yellows, and greens of normal human colour vision, provide a selective advantage?

Answers to the first question, what one might call the ‘Good-For’ Question, commonly refer to the foraging practices of our ancestors. Seeing fruit as red and its background foliage as green made it easier to see the fruit amongst the leaves. Red fruit ‘pops out’ from amidst a green background (Mollon 1989; Regan, Julliot et al. 1998; Regan, Julliot et al. 2001). Or: Adding a third cone to the mammalian dichromatic retina made it possible to see finer gradations of the colour green, and hence to select the tender new leaves hidden among the \ less nutritious older foliage (Lucas, Dominy et al. 2003; Dominy and Lucas 2004). Or: Trichromaticity provided the wherewithal to differentiate various shades of red hence to choose the ripest fruits with the greatest sugar content (Sumner and Mollon 2000; Smith, Buchanan-Smith et al. 2003; Riba-Hernandez, Stoner et al. 2005). While variations on the foraging hypothesis predominate, there are few other hypotheses about the selective advantage of trichromatic vision all of which are ‘single-nose’ explanations. For

example, one recent theory cites the trichromat's ability to make fine-grained distinctions in facial colouring and its role in the perception of the general health, fertility, and the emotional states of smooth-faced con-specifics (Changizi, Zhang et al. 2006).¹

Importantly, answers to the 'Good For' question are *substantive* evolutionary hypotheses. In some sense of 'could', we could have been a species with only luminance vision, creatures who saw the world 'in black and white'; we could have been achromats with only one kind of daylight photoreceptor (or, what amounts to the same thing, a species with more than one type of cone but without the ability to compare their signals). Instead, the human retina contains three different kinds of cones, each of which responds to light across different but overlapping portions of the spectrum. Through a comparison of these three signals, our visual system can determine the predominant wavelength of light stimuli in the natural environment with reasonable accuracy—a capacity without which we could not see a coloured world. So, there is question to be answered here. Why are *we* trichromats when the vast majority of mammals are merely dichromats? And why do other mammals have dichromatic 'colour' (to be generous) rather than the monochromat's 'shades of gray'? Why *colour* at all?

Answers to the second question, 'why these colours and not some others?' are usually an extension of the author's first answer. For example, Sumner and Mollon (2000), both advocates of a fruit-foraging hypothesis, suggest that 'of all (the) possible combinations of cone sensitivities, the spectral positions of the actual primate pigments are optimal for fruit foraging' (p. 163). So we see the colours that we do—our three cones are tuned the way they are—because this tuning was particularly well-suited or at least *better* suited than the other options embodied throughout our evolutionary history, relative to the task for which the colours evolved. Trichromaticity makes possible the perception of just those colours that are *telling*, between tender shoots and tough old roughage, between young, fertile women with a healthy glow and the old, pallid and barren, etc. Again, such views about the spectral tuning of our three cones are put forward as empirical hypotheses.²

To sum up, the dialectic of the debate about human colour vision usually goes like this. Human colour vision is for seeing the colours. (Call this assumption 'Colour-for-Colouring'.) Hence we need to ask, first, why seeing the colours was beneficial—the 'Good For' Question. And, second, why do we see the very colours that we do—why are we trichromats of *this* kind?

¹ Clyde Hardin offers a very different kind of answer. In *The Virtues of Illusion* (Hardin, C. L. (1992). "Virtues of Illusion." *Philosophical Studies* 68(3): 371-382.), in a section entitled "How to Engineer A Wavelength Detector", Hardin portrays human colour vision as a solution to an 'informational bottleneck'. The human retina has 120 million photoreceptors but the optic tract has only 1.2 million fibres, or 'output' lines. Somehow information about wavelength must be conveyed to various cortical and subcortical regions using this limited hardware and yet do so without compromising our system for high spatial resolution luminance vision. The answer: Encode only wavelength *differences* and piggy-back this system onto the existent opponent system for luminance contrast processing. As the reader will see, much of this paper is spent attempting to question how Hardin sets up the question.

² As the reader will have surmised, this way of putting the question makes some very basic assumptions about the relation between the tunings of the photoreceptors and the nature of our colour categories. The strongest assumption is that the tunings, which are the result of the expression of specific genes, determine the colour categories. Hence there is no cultural or linguistic influence on the categories of human colour perception (at least insofar as we all have the same genes for the three cone types). The relation is one of biological determination, from genes to tunings to colour categories. Most theorists have a weaker view about the relation between photoreceptor tuning and colour categories, some variation on the view that our cones (and hence our

In this paper, we want to back-up to the first move in this chain of inquiry and deny the assumption that human cortical colour vision is ‘for’ colouring. Instead, our ‘colour’ system is better understood as two *chromatic contrast* systems, each of which encodes additional information about the retinal image and which, individually, increase our capacity for wavelength discrimination. Like all other mammals, we have a luminance system to accomplish many of the general tasks of mammalian vision—to guide our actions, to discern objects, properties and events. Like all other diurnal mammals, we also have a dichromatic (two cone) system for luminance information (which *sums* the cone signals) plus a very rough capacity for wavelength discrimination (through the *comparison* of the two cone outputs). This first, rudimentary chromatic system encodes coarse-grained spectral contrast at low spatial resolution, information suitable for a variety of tasks such as basic scene segmentation—i.e. any visual task that can be solved with this sort of coarse chromatic and spatial information. With the exception of the other Old World primates, however, our species *Homo sapiens* is unique among diurnal mammals. We have an additional newer, chromatic contrast system with a higher spatial resolution and finer wavelength discrimination. It is this system, the Red-Green (or M-L) system that allows us to use chromatic information for many of the tasks of object perception, the very tasks commonly attributed to luminance vision. The addition of this second chromatic system was a landmark in Catarrhine visual evolution: the newer Red-Green system added speed, accuracy, and an economy of resources to dichromatic vision. Thus, we have two intertwined systems of vision, a luminance system and a chromatic one, with this latter system divided into two parts.

Later in this paper we will say more about how the various tasks of human vision are parceled out amongst these various systems. But the basic upshot is this: Both chromatic and luminance contrast information are used for the multitude of tasks that comprise human vision. These two types of information may be used in parallel (separate mechanisms), or pooled together (via a common computation) or alone (e.g. via a luminance mechanism for global motion for which there is no chromatic counterpart). But always, information is used opportunistically as a function of task and the most reliable signal. In short, chromatic vision, like luminance vision, is for *seeing*.³ Spelling out this theory of Spectral Vision is the primary task of this paper.

But first a word of warning: Quine was right. You cannot dislodge a well-entrenched view, a ‘near tautology’ in this case, by philosophical hand-waving or magical incantation. (‘OUT, OUT, damned colours!’). Sometimes the only way to pry loose a seemingly ‘unmovable’ view is to *shift the landscape around it*. This explains both the format and length of this paper. We start with three puzzles taken from the neurophysiology of colour vision to illustrate why Colour-for-Colouring is far more puzzling than any tautology has the right to be. Then we present an alternative, the Spectral View, and how it resolves these puzzles. Like all landscaping projects, however, the process can appear chaotic and more than a bit messy. We offer, by way of hope, the landscapers’ mantra: ‘It’s gonna look GREAT. We *promise*.’

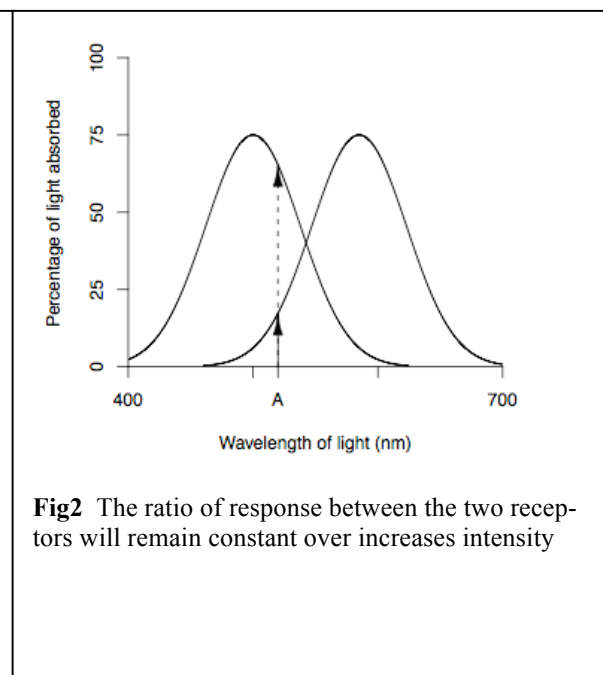
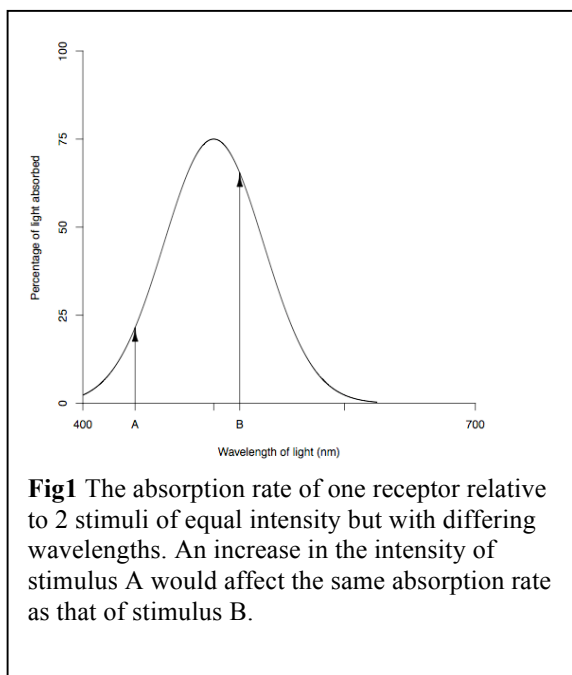
³ The view we are presenting—and its near neighbors—is not original to us, or not in its skeletal form. Probably the first person to have expressed something like this view is Peter Gouras. There have been many researchers who, since the late 1980’s, devoted their research careers to providing an alternative to the “Colour-for-Colouring” view, and their work has been invaluable in providing an empirical leg to stand on, for the view put forward here. Among those researchers are Fred Kingdom, Kathy Mullen, Karl Gegenfurtner, Peter Schiller,

II. THREE PUZZLES

A. Why is Trichromatic Vision an Anomaly in Diurnal Mammals? Of the 4600 current species of mammals, almost all are dichromats or functional dichromats. ‘Robust’ trichromats—species that have three kinds of cones and enough of those cones to compare their signals, are rare. Catarrhrhine primates—roughly, primates with nostrils that face downwards as opposed to sideways—include humans and all of the apes (chimpanzees, gibbons, orangutans, and gorillas) and the Old World monkeys (among others, the Macaque, the Bonobo, baboons, mandrills, and vervet monkeys). Until recently, the only known other robust trichromatic mammal was one species of platyrrhine primate, the Howler Monkey.⁴ So robust trichromaticity is very rare in the mammalian world—less than 50 species out of 4600.

Until about twenty years ago, this fact, that we are ‘one of the few’ was viewed as clear evidence of the superiority of the human visual system. The standard story went something like this.

It is common knowledge that every monochromat, an individual with only one kind of photoreceptor, is colour blind: a monochromat cannot discriminate between two lights on the basis of wavelength alone. Figure 1 illustrates the rate of photon absorption as function of stimulus wavelength, with each stimulus presented at a constant intensity. Increase (or decrease) the intensity of any stimulus enough, and the maximum (minimum) absorption can be affected. Thus the graded response of photoreceptor does not *indicate* wavelength: its response *conflates* intensity and wavelength. In essence, this is The Principle of Univariance. You cannot tell the intensity and wavelength of the stimulus from the rate of photon absorption alone. However, if we add another cone to the system, the standard story goes, an organism gains wavelength discrimination (Figure 2). Given two differently tuned cones,



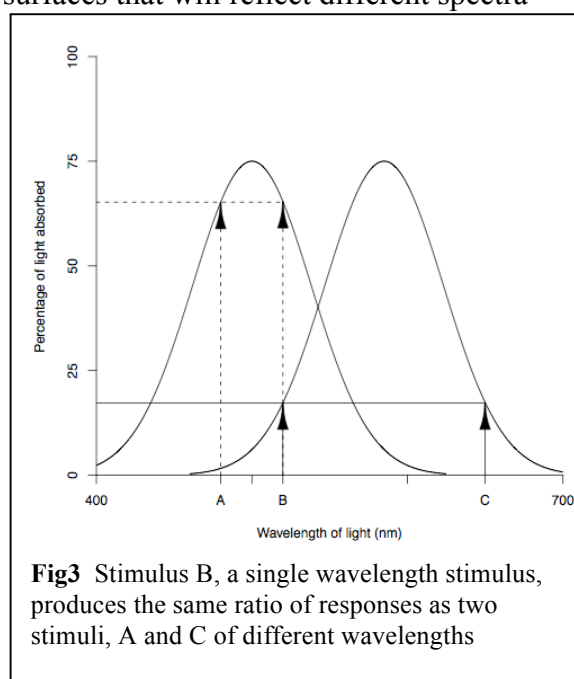
⁴ Recently, four species of Australian marsupial have been discovered to be trichromatic although it is not known whether they possess true trichromatic wavelength discrimination (Arrese et al. 2002 and 2005.) For the true aficionado of Oz fauna, these are the quokka, quenda, the honey possum and the fat-tailed dunnart but alas

their responses to a single wavelength stimulus will occur in a fixed ratio regardless of light intensity.

Unfortunately, dichromatic systems have only *rough* wavelength discrimination. Any dichromat can be “fooled” by a light stimulus of two or more wavelengths. For any given ratio of cone responses caused by a light of a single wavelength, there will always be a stimulus composed of two or more wavelengths of light that will produce exactly the same response (Figure 3). Thus, a dichromat has only very poor colour vision. By adding more receptors with overlapping responses, more accurate wavelength discrimination is possible. No matter how many receptors are added, however, the same principle holds: As trichromats, we can be fooled by a stimulus composed of three lights of different wavelengths; a tetrachromat can be fooled by a stimulus of four lights; and so on. In the laboratory, we can easily produce colour metamers for normal humans, surfaces that will reflect different spectra of light but which are indistinguishable to the human trichromatic observer.

Laboratories aside, trichromats are rarely ‘tricked’ under natural viewing conditions. Nature contains few colour metamers for human trichromats. So, practically speaking, human trichromats have very good wavelength discrimination. The world is simply less accommodating to the dichromat. Natural dyes and pigments—the primary cause of colour in natural objects—absorb and reflect light across the continuous spectrum of visible light producing spectrally complex reflections. For dichromatic systems, this complexity results in bountiful metamers. So, according to the common way of thinking, the benefit of two cones over one is a very rough capacity for wavelength discrimination: i.e. *very bad colour vision*.

(Indeed, it is only in the last 20 years or so that dichromatic species have been acknowledged as having colour vision at all.) Given these facts, it is easy to see why *Homo sapiens* have been portrayed as at the apex of mammalian colour vision: for wavelength discrimination, three cones really are better than two.



The twists and turns of the evolution of vision suggest a more complicated story however. Colour vision did not evolve as one imagines, as a progression from one cone, to two cones, and finally to three cones for the lucky few species. Approximately 400 million years, the common ancestor of all terrestrial vertebrates (and one marine species, the lungfish) had four different cone types as well as a set of coloured oil droplets (Vorobyev 2003; Jacobs and Rowe 2004; Vorobyev 2004). Coloured oil droplets act as colour filters, allowing only a limited range of wavelengths to interact with the photopigments (Vorobyev 2003; Hart, Bailes et al. 2008). And because they reduce the amount of useable light, oil droplets also reduce luminance sensitivity. Thus a coloured oil droplet, added to a cone, produces a new functional type of cone, an addition that can potentially change the kind of wavelength processing that is possible in the retina. Now, given the genetics of photoreceptors alone, we do not know how the common ancestor of terrestrial vertebrates *used* the four types of cones and various oil droplets—whether its retina contained the requisite wiring to compare any of these cone outputs whether in a given retina the same cone type existed with and without oil

droplets. Hence our common ancestor's capacity for wavelength discrimination is not known. But potentially, our common ancestor had a rich capacity for colour vision—the wavelength discrimination of an anatomical tetrachromat with whatever increased capacities for color vision/wavelength discrimination that oil droplets provided. (For example, four different cone types and three different oil droplets represents the potential for 16 different kinds of receptors—16 receptors with different sensitivities to wavelength—assuming that each oil droplet could be paired with each type of cone.)

In the 300 to 400 million years since, from this rich beginning, the evolution of colour vision in mammals has been a history of cone *loss* (Jacobs and Rowe, 2004). Beginning with Eutherian mammal, small nocturnal species, the conditions of night vision favored highly sensitive photoreceptors, the rods, the only type of photoreceptor for night vision in mammals. (In a dark-adapted retina, rods respond to the absorption of a single photon.) Even with such sensitive receptors, however, most systems of night vision also have neural convergence, the pooling of signals from multiple receptors, to increase their signal-to-noise ratio. To put this the other way around, cones, which are relatively insensitive to light, are ill-suited to night vision; so too is the capacity for wavelength discrimination which requires a *comparison* of receptor signals not a pooling. So nocturnal behavior would have provided a strong evolutionary pressure towards a rod-based, convergent system as opposed to the relatively insensitive, divergent arrangement of a cone-based retina (Jacobs 2009) .

The mystery posed by current mammalian species is why, given the re-emergence of diurnal species, the mammalian retina never regained its rich capacity for colour vision. The evolutionary history of mammalian colour vision is from four or more types of cones and numerous oil droplets to a rod-rich, dichromatic (two cone) retina. But this riches-to-rags story has not been reversed. In our lineage, that of catarrhine primates, trichromaticity has re-evolved through the reduplication of the genes for the LWS cone opsin (the progenitor of our L cone) and their subsequent divergence into genes for the L and M cones. As a result, in addition to a short wavelength cone, each catarrhine species has two cones derived from the LWS cone with distinct but near tunings. The L and M cones of humans are the result of this same reduplication and divergence. But our story is the exception that presumably proves the rule (whatever that rule might be!).

So this is a puzzle. When our ancestors became nocturnal, why were mammals left with two cones not one (e.g. how did dichromatic vision help nocturnal vision)? And when diurnal mammals reappeared, why were catarrhine primates the only mammals (apart from the Howler monkey) that re-evolved into robust trichromats? If dichromatic vision is the poor man's version of colour, one would think that *real* colour vision—trichromatic colour vision—would have made a stronger re-emergence over the last 300-400 million years.

B. Why does the Colour System Occupy such a Large and Central Part in human Vision? Although the human retina has approximately 900 million rods and only 100 million cones, or a ratio of 9 to 1 rods to cones, the ratio changes dramatically at the 'output' end of the retina, the retinal ganglion cells. Approximately 50% of all ganglion cells are wavelength sensitive cells (Zrenner, Nelson et al. 1983). In the cortex, all of the early visual areas from primary visual cortex, V1, through the extrastriate areas, V2, V3, VP, and V4, have a robust response to colour contrast. Here again, approximately 50% of the neurons in these areas are wavelength sensitive (Dow and Gouras 1973; Gouras 1974; Thorell, De Valois et al. 1984; Kiper, Fenstemaker et al. 1997; Johnson, Hawken et al. 2001). Whatever the exact ratio of luminance to chromatic cells it is clear that a non-trivial percentage of our cortical resources

are involved in some kind of colour processing. But why? Is discerning colour so difficult that it requires 50% of our visual resources at the level of V1? Even assuming that wavelength discrimination is a very difficult problem, we are still left with the question of why seeing the colours was so important in the first place, why the expenditure of resources would be worth the cost? It is easy to imagine, for example, why better spatial acuity would have proven advantageous. (Who would want to meet a large carnivore without his or her glasses?) But the central deficit of the human dichromat is wavelength discrimination, e.g. the inability to discriminate pale lavender from celadon green.

As we said above, the standard answers given are in terms of food foraging, yet recent attempts at investigating these claims have had mixed results. For one, our ancestors were omnivores not fructivores, so fruit-foraging was only one of the many food-related tasks of our ancestors. Second, recent research has suggested that the dichromat is equally good at selecting ripe fruit based on dichromatic information (Osorio and Vorobyev 2008)—and that, in any event, the ripeness of fruit can be discerned just as easily on the basis of surface lightness, at least for fruit within arm's reach (Hiramatsu, Melin et al. 2008). So, the advantage of trichromaticity in terms of food foraging is not well established. Third, given the complexity of primate vision, foraging is only one of many tasks of primate vision writ. Indeed, no one has suggested *any* task that would account for the huge population of wavelength-sensitive cells in human visual cortex in—that is, established an advantage of trichromaticity large enough to explain the chromatic sensitivity of 50% of V1 cells.

In addition to the sheer numbers of chromatic cells, the colour system occupies pride of place in our visual system—the fovea of the retina, the small, central, pitted area of the retina densely-packed with receptors. Again, all things being equal, this arrangement ought not to occur in nature given its cost to spatial processing. If a system has one type of cone, then every area of visual space must be monitored by that receptor. But if a system has two types of cones and the goal is to compare their signals for the purpose of wavelength discrimination, then every area of visual space must be 'looked at' by two receptors, one of each type. This means that a dichromatic system must have twice as many cones to monitor the same area of visual space as its monochromatic counterpart. Assuming that there is a fixed capacity for photoreceptors in the retina (i.e. you can only jam so many receptors into a given space), a dichromatic system will have half the spatial resolution of a monochromatic one. This explains the common wisdom about the cost of colour: additional cones can be added only at the expense of spatial resolution. Against this assumption, then, it is odd that the fovea, the very area of primate retina that is used *for high spatial resolution vision*, contains all three types of cones in human vision. So, all things being equal⁵, this is a surprising organization.

In fact, on the assumption that human colour vision is for seeing the colours, there are many other ways that the retina could have been organized. For example, one route around the incompatibility of colour processing and spatial resolution would be to have a small, peripheral region of the retina, dedicated to either colour or spatial resolution. A nice example of this kind of trade-off is found in raptors: a dedicated fovea for high-resolution depth perception. This circular fovea of tightly packed receptors looks forward at the midline of visual space, ahead of the raptor. By combining the signals from the left and right eye, it provides stereoscopic depth perception for when raptors drop upon prey from high altitudes.

⁵ As will become apparent by the end of this paper, the phrase 'all things being equal' is crucial—because as it

(A good idea during high velocity descent if you consider the alternatives.) Along similar lines, we might have had small peripheral areas of the retina devoted to colour detection and identification. A low resolution ‘colour fovea’ could be shifted from region to region of visual space when colour information is needed, while the monochromatic central fovea could encode the high resolution spatial information required for form, shape, depth, and texture processing. After all, why would an organism need fine-grained spatial information for *colour* vision, if the colours, in fuzzy spatial outline, could be ‘snapped to the grip’, the boundaries provided by luminance vision? So, again, the puzzle is this: what was so beneficial about colour vision that something of utmost importance to primate vision—spatial resolution—was sacrificed for it?

C. Why Are the Blue Cones so Rare? As human trichromats we have three different kinds of cones: the S cones are maximally sensitive to short or ‘blue’ wavelengths; the M cones prefer medium or ‘green’ wavelengths, and; L cones respond preferentially to long or ‘red’ wavelengths. Hence, their colloquial names, the blue, green and red cones (Figure 4). Now, *prima facie*, if you were designing a system for wavelength detection from scratch, you would expect the three cones to overlap equally, such that the comparisons between their responses would be equally accurate for wavelength discrimination. This is clearly not the case, as the L and M cones almost entirely overlap in their response while the S cone covers much of the short wavelength range of human colour vision without any overlap from the M cones. Moreover, S cones comprise only a small fraction of our colour system, only 5 percent of all cones. The vast majority of cones are M and L cones, with the ratio of M to L cones varying from 1:1 to 1:4 depending upon the individual. Finally, there are no S cones within the central two degrees of the fovea, the area of the fovea with maximal spatial resolution, i.e. no trichromatic colour vision in the central two degrees of human vision

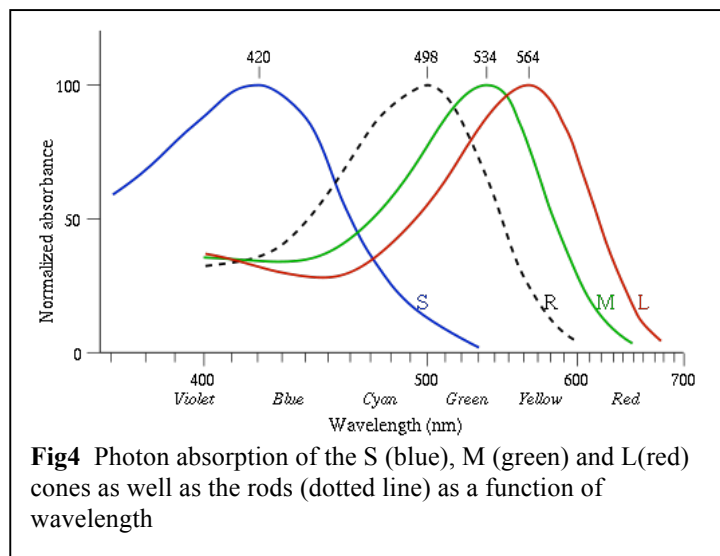


Fig4 Photon absorption of the S (blue), M (green) and L (red) cones as well as the rods (dotted line) as a function of wavelength

The standard explanation of the unequal role played by blue cones draws upon the phenomenon of chromatic aberration. As sunlight passes through the lens of the eye, its degree of refraction is wavelength dependent: Short wavelengths bend more than longer ones. Thus when any image of the distal world, by passing through the lens, forms a multiplicity of images on the retina, each defined by wavelength and displaced relative to the others. (Think here of a very bad three-colour reproduction in a newspaper, those blurred photographs in which the RGB images are improperly registered.) In a retina with a single kind of cone, this problem is minimized by the fact that every cone is maximally responsive to the same range of wavelengths. An M cone responds preferentially to the ‘green’ image, for example. Adding new types of cones, say S cones, creates a problem. An S cone will respond maximally to a different image, the spatially displaced ‘blue’ image. So by adding an S cone to the M cone system, this new input only serves to magnify the spatial blur of chromatic aberration.

This explanation makes enormous sense of why blue cones would not be used for high acuity spatial processing. But it does not make sense of the actual problem at hand—namely why

blue cones have a minimal presence in human colour vision. If the trichromatic system is for colouring—for *determining colours*—the spatial blur of the colour image ought not to matter. As we know, the luminance system provides the visual system with the high acuity spatial information about objects in the distal world. It provides the *lines* (and texture) of the visual world, metaphorically speaking. But with the ‘lines’ in hand, as we said above, the visual system needs only the rough co-ordinates of where the colours are to go, in order to ‘colour within the lines’. At the initial stages of human visual processing, a good bit of ‘blur’ in the colour borders of the perceived world ought not to make any difference to our conscious perceptions of a coloured world, given the spatial information provided by luminance processing.

To summarize, colour systems, both our own and those of other species, raise a number of questions about the standard explanations of human trichromaticity. Although our robust trichromatic visual system is vastly superior to a weak dichromatic one for wavelength discrimination, robust trichromatic mammals are the exception. Although a trichromatic system is generally thought to sacrifice spatial acuity to the service of wavelength discrimination, our three types of cones occupy pride of place in the fovea. And although our trichromatic system is supposed to be for ‘colouring’ and not spatial processing, the blue cone system still has a minimal role in human colour vision. Why would this be?

III. RECASTING THE QUESTION

Up until very recently, most research on vision, whether on a natural or artificial systems, has assumed that *the primary stimulus of any visual system is light intensity*. When light interacts with the various media and surfaces of the world, the intensity and direction of light is profoundly affected. Thus the clues in virtue of which we navigate through/discern the properties of the distal world are those to be found in the luminance image, in effect the intensity of light that falls upon the retina at each point, as filtered by the receptors. The computational components of vision, whatever those happen to be, thus start with information about the luminance image and presumably yield systematic information about, say, global expansion motion within the image, stereoscopic depth relations, and so on. (Note that this general view is neutral with respect to the issue of modularity even though it is difficult to describe it in a theory neutral way.) Within this larger framework, the computation of colour is generally understood in a similar way, as one of the many components of human visual processing. Naturally colour vision requires more than luminance information. On the input end things, there must also be chromatic encoding of the wavelength properties of the image. Still, it is common to regard human colour vision as one of the sub-processes of human vision, an evolutionary add-on the survival value of which we can ask about separately. More pointedly, most research on the evolution of colour vision or the role of colour processing in vision more generally is driven by a seemingly innocuous question: What does vision do with *colour* information that it could not do with luminance information—and what mechanisms subserve these abilities?

When one stops to think about it, however, strictly speaking that is *not* the right question, the question ‘posed’ in the evolution of any visual system. First, as we all know, a light wave has more than one dimension. In addition to amplitude (or intensity), a light wave has wavelength and polarity as its other ‘intrinsic’ properties, plus its direction of propagation and velocity. So when sunlight is reflected, refracted, absorbed, filtered and scattered by the various media and surfaces within an environment, each of these dimensions is affected in profoundly complex yet *law-like* ways. The laws of optics, and the laws of physics below the level of optics apply to all three dimensions with different effects upon different dimensions

Thus each dimension of light carries the potential for information about its causal history, the media through which it was transmitted, the surfaces from which it was reflected, by which it was absorbed, etc. The multiple dimensions of light embody both dependent and independent sources of information about the distal world.

A less well-known fact about vision concerns the receivers of organic systems. Only one kind of light transducer has evolved on Earth: the organic molecules, known as photopigments, that change their shape through the absorption of electro-magnetic energy. Importantly, each photopigment has a selective response to *all three* of the intrinsic dimensions of light, to the amplitude, wavelength, and the polarity of light waves—a preferred wavelength of light, a preferred polarity (relative to the spatial orientation of the photopigment), as well as having increased absorption given an increase in light intensity. Thus, the graded response of a photoreceptor conflates wavelength, intensity and polarity. Photoreceptor response does not signal any one dimension of the stimulus alone.

Given this shared starting point, the evolution of any visual system serves to answer this question: *Which* aspects of the light stimulus will prove both accessible and useful relative to a specific light environment and the behaviors the species? It is this ‘choice point’ and the extraordinary diversity of possible ‘responses’ which account for the astounding diversity in retinal construction and visual systems. More relevantly to our purposes, this explains why the question ‘what could a visual system do with *colour* that it could not do with luminance information?’ is not the right question. The most general question about the evolution of ‘colour’ vision is not about ‘colour’ per se, but about *wavelength*. If stimulus wavelength could be disambiguated from intensity and polarity, might this be useful to a given species?

In the mammalian case, recent evolution has been given a simpler task. The random orientation of photopigments in the receptor membrane of mammals ensures that polarity is largely eliminated as a causal factor in photon absorption. Thus, the mammalian photoreceptor conflates only two dimensions of light, intensity and wavelength. So mammalian photoreceptor provides a graded response to the stimulus as a function of both variables.

A second problem with the standard question is a common one in reverse engineering. When we ask ‘what could a visual system do with colour *that it could not do with intensity information?*’ we make the tacit assumption that evolution selects traits that are *necessary* for survival. When a sensory system guides the real-time behaviour of an organism, *any* information that makes its visual computations faster, cheaper or more reliable is ripe for selection. The question is always one of *significant benefit*. We therefore want to ask about the *benefits* that would occur if chromatic information were added to a luminance system.⁶

Taken together the above two points suggest that we ought to be asking a significantly different question. The standard question ‘what could you do with *colour* that you *could not do* with luminance information?’ is a question about a property of the world, colour, not

⁶ One bit of terminology that may be confusing to the reader concerns the use of ‘wavelength’ and ‘intensity’ versus ‘chromatic’ and ‘luminance’. You are not alone. Briefly, wavelength and intensity are properties of the light stimulus while ‘chromatic’ and ‘luminance’ are terms that refer to types of visual cells or systems. A luminance system compares the outputs of one kind of receptor. Luminance is a measure of light relative to the absorption properties of one type of receptor (or set of receptors). A chromatic cell compares the outputs of two different types of photoreceptors, e.g. an M and L cone. For a more detailed (and coherent) explanation, see

about one dimension of the stimulus, wavelength. It also confuses the evolutionary prerogative, traits which are of *survival advantage*, with a question about *computational necessity*. Recasting the question, we should ask: ‘What survival advantage would encoding spectral/wavelength information bring to human vision writ large?’

PART II: THE COSTS AND BENEFITS OF SPECTRAL VISION

V. SPECTRAL INFORMATION AND OBJECT VISION

Before exploring how spectral information could be used, let us say a few words about what spectral information *is*. Insofar as neurophysiologists of vision agree upon any First Principle of Vision, it would be this: *The proximal stimulus of all vision is contrast*, not light per se. To see anything at all, be it only the detection of a large looming shadow, there must be contrast in the light stimulus. So, for example, if a luminance system is presented with an image of a black square against an identical black background, this will not be of much help to that system. A luminance system needs a dark figure against a lighter background or, conversely, a light figure against a darker ground. It needs *positive or negative contrast*, the difference in luminance between two spatial regions. The same rule holds for any dimension of the light stimulus. Although, light at each point in the image has a complex wavelength composition, this information is not useful unless there is a *difference* in the spectral power distribution (wavelength contribution) from point to point in the image. There must be some spatial border, demarcated by differences in spectral composition—a difference in colour between, say, green here, and yellow just there. In spectral vision, then, the proximal stimulus is *spectral contrast* and the most important information in the retinal image is the location, amount, and sign (positive or negative) of that contrast.

Above we said that, until a decade or so ago, most vision researchers paid little attention to the potential uses of wavelength information. This was due in part to a basic assumption about the relation between wavelength and intensity information in the world (see below). But for contemporary researchers, the predominance of luminance function was confirmed by the seminal experiments of Livingstone and Hubel in the late 1980’s on primate colour vision (Hubel and Livingstone 1987; Livingstone and Hubel 1987; Livingstone and Hubel 1988), whose experiments seemed to show that, overwhelmingly, human visual processing was performed by a variety of luminance systems.⁷ From the first person perspective, colour is a

⁷ Based upon single-cell recordings in Macaque, Livingstone and Hubel concluded that there were two subdivisions of the central pathway from retina to cortex in primates, the geniculo-striate pathway. The M or magnocellular pathway was associated with low spatial resolution, high temporal resolution (‘fast’ or ‘zippy’) luminance contrast signals; the P or parvocellular pathway was correlated with high spatial resolution, low temporal resolution, luminance and colour encoding. It was Livingstone and Hubel’s psychophysical experiments that seemed to have the most influence in entrenching the ‘colour-for-colouring’ view in the minds of vision researchers. Livingstone and Hubel presented subjects with a number of different scenes, each chosen to depict one property that human vision was known to process—e.g. depth from occlusion, complex shape, linkage between spatially disparate parts of an object, and so on. Each image was presented in two ways. In the first image, all spectral contrast was removed, i.e. subjects were shown a standard black and white photograph. In the other, the scene was rendered in just two colours each with same intensity, thereby removing luminance contrast as a stimulus. Lo and behold, the black and white photographs were easy to see while the two-colour images were almost always difficult to interpret. The perception of apparent motion, complex form, and depth as revealed by many different kinds of cues was degraded or absent with the bi-colour stimuli. Hence Livingstone and Hubel’s conclusion that colour information, and its physiological substrate, the P channel, played a minimal role in human vision.

central perceptual property. But nonetheless it is a proper subsystem of human vision largely concerned with a single task—seeing the colours.

Livingstone and Hubel’s research had a profound effect upon vision science, sparking a new wave of interest in both in the neurophysiology and psychophysics of colour perception. Yet like all ‘big’ scientific results, Livingstone and Hubel’s conclusions were controversial. How could it be that the Magnocellular Pathway, a pure luminance pathway that comprised only 10% of the neural output of the retina, nonetheless appeared to do 90% of the work? This seemed, to many vision scientists, very *odd*.

Still, even scientists who were wary of the Livingstone and Hubel’s experimental results agreed that chromatic processing—colour—was unlikely to make a significant contribution to general vision. The problem here, as ever, was the ‘Good For’ question. Colour might be lovely for viewing sunsets, scenery, and paintings; it might even be *helpful* for picking berries faster than your competitors (e.g. bears). But apart from those two tired answers, the potential uses of colour have always seemed strangely impoverished. What exactly is the function of a colour system? More specifically, even prior to Livingstone and Hubel’s research, vision science has long accepted that wavelength information did not provide any additional information about the image over and above what was already contained in its intensity counterpart. As nature would have it, *any two objects that differ in colour also differ in lightness/darkness*. For example, a red strawberry differs in colour from its background of green foliage, a fact that is visually apparent to every normal trichromat. But a ripe red strawberry is also *darker than* its background of green leaves (a counterintuitive fact given that the red strawberries seem to jump out, from among the leaves, at the viewer.) More generally, wherever there is spectral contrast, there is also intensity contrast. Hence if you had only a luminance system, the general encoding of luminance edges would include all those edges also defined by chromatic contrast. Adding a chromatic system serves to encode only redundant information.

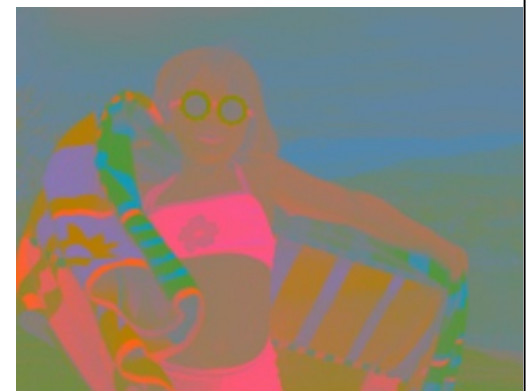
If one thinks along evolutionary lines, however, the argument looks less than convincing.



a. Full colour image



b. Intensity image.



c. Isoluminant image

Fig 5 Illustration of the separation of a colour image (a) into its luminance (b) and spectral (c) components (Original image courtesy of Kathleen Akins; isoluminant and isochromatic photographs courtesy of Karl

1. In the natural world, where there is wavelength contrast, there is intensity contrast.
2. Hence spectral/wavelength contrast information is *redundant*.
3. Hence spectral contrast is *not useful* to mammalian vision.

Most readers will be aware of the dubious leap from (2) to (3). Redundancy does not always amount to the absence of utility. (Fortunately, safety engineers of nuclear power facilities are fairly clear on this point.) But put that aside for a moment and look at the central empirical issue, namely the *relation between wavelength contrast and intensity contrast in an image*.

We can think of any retinal image as having two distinct components, an intensity component (or the intensity of light at each point in the image) and a spectral or wavelength component (or the wavelength composition of light at each point). Figure 5 illustrates these two components. Figure 5(a) is the original colour image, a standard colour photograph that contains both kinds of information. In the middle is the Intensity Image, which illustrates the intensity of the original photograph. At bottom in Figure 5(c), the ‘spectral image’ or Isoluminant Image is the original colour image with the intensity rendered constant across the image as a whole. Note that although the Isoluminant Image appears ‘fuzzy’, this is an illusion caused by our own visual system. Objectively, both the Intensity image and the Isoluminant Image have the same spatial resolution. Like the Intensity Image, the Isoluminant Image also contains contrast, here spectral contrast, the difference in predominant wavelengths between two points. Looking at various parts of the towel, for example, it contains many areas of *high* spectral contrast, e.g. where the *purple* and *orange* stripes meet. An edge delineated by *low* spectral contrast, for example, is the barely visible line between the distant mountains (saturated *blue*) and the hills of the foreground (de-saturated *green*).

As one can see from Figure 5, the Isoluminant and Intensity images bring out quite different features of the original image. For example, in the Intensity Image, it is not always apparent where the girl ends and towel begins given that her outline is partially obscured by shadows—e.g. under her outstretched arm. But in the Isoluminant Image, the distinction is clearly visible in virtue of the contrast between her skin tone and the brightly coloured towel. On the other hand, without the shadows in the Isoluminant Image it is impossible to see the fine details of body shape, e.g. the three-dimensional contours of her outstretched arm.

There is a general lesson about shadowing and spectral vision lurking here. First, the assumption that ‘where there is wavelength contrast, there is intensity contrast’ is true. Natural dyes and pigments produce an indefinite variety of absorption spectra (hence surface colours) such that objects that differ from their background in predominant wavelength also differ in intensity. The converse of this mantra is *not* true however. Where there is intensity contrast, there need not be spectral contrast. In the natural world, where directional lighting creates shadows, there are numerous intensity boundaries that crosscut object edges. Shadows can—and do—fall almost anywhere. So an intensity border in the image has many potential sources in the world. It could result from the edge of an object, from the variations in an object’s surface colour (the leopard’s spots), from a shadow cast by a nearby object or a shadow created by the object’s own shape or surface texture. In an Isoluminant Image, however, there are no shadows. Thus a spectral edge is caused by an object boundary or by surface colouring (e.g. the spots on a leopard or the edge of the leopard itself) while an intensity edge indicates either an edge of an object or a shadow.

This fact is extremely important for object vision. As researchers in artificial vision know, shadows are a mixed blessing. On the positive side, shadows are a potential source of vast amounts of visual information—the shape of objects (from shading of various kinds), their surface texture (from the patterned shadowing that texture creates), the location of objects relative to each other (given the shadows they cast on each other). On the negative side

strongly directional light can obscure the crucial edges and features of objects in even very simple scenes. (Figure 6, a photograph taken in bright morning light, provides an excellent example of problem.) Without background knowledge, shadows make scene segmentation both expensive and time consuming. The standard wisdom about shadows and image processing is just this: you can't live with them and you can't live without them. Shadows are irreplaceable sources of information but only if they are recognized *as* shadows—i.e. distinguished from the object boundaries and differences in surface colour of objects.



Fig 6 Illustration of the problems facing any luminance system for scenes with heavy shadowing even in bright light conditions (Photograph reproduced courtesy of Gena Hahn.)

Enter spectral contrast in addition to intensity contrast. Given spectral contrast, low-level vision could easily instantiate the following simple rules:

1. To find a shadow within an image, look for edges defined solely by intensity contrast.
2. To find an object boundary or a patch of surface colour, look for a 'combined' edge of both spectral and intensity contrast.
3. To find a patch of surface colour (as opposed to an object boundary), look for a 'combined' edge defined by intensity and wavelength contrast—but one that does not show evidence of shadowing.



Fig7 Same image as Fig7 but partially 'coloured' to reproduce the original scene

Of course, if you want to separate the spots on a leopard from dappled shadowing and leopard edges, you may be out of luck.⁸ Unless the leopard moves, Rules (1) – (3) probably won't work. However, Rules (1) – (3) will explain *why* leopard spots are such effective camouflage for dichromatic or trichromatic perceivers (an epiphany that might strike you just as your time on earth comes to an end).

The above (three) rules can be applied to

images without any additional knowledge of the properties of distal objects, hence at a very low level of visual processing. This is why any trichromatic or dichromatic subject/predator, observing the scene in Figure 7 *in situ*, would have seen the two figures quickly and

⁸ The leopard's spots are good example of how camouflage takes advantage of these rules. A leopard's spots, composed of multiple shades of brown, from light to dark, appear to be *shadowed circles*, thus creating

accurately. Both figures, clad in red and green, were lying on a background of royal blue sheets, a scene with high spectral contrast despite the shadowing. With both sources of contrast, the figures jump out (Figure 7). Importantly, there is no reason to think that, *pace* Mollon (Mollon 1989) that our ability to quickly take in the scene and see the two figures clearly results from the representation of the reds, greens, and blues of the scene. Rather, it is the addition of further *visual contrast*—here, chromatic contrast—that does the hard lifting. (For the photographer, using black and white film, the challenge is to see this scene as it *would appear* without the spectral information.)



Fig 8 The R-G filtering of the image in Figure 6(c)
(Image provided by and with the permission of Karl Gegenfurtner.)

Some further images should make these points more plausible and also extend the argument, here images that represent cone signals as opposed to the intensity and wavelength information present in the photographs. All of the images that follow began as full colour photographs. From each photograph, two more images are derived, a Luminance Image and an R/G Chromatic Image (Figure 8). The R/G image represents the relative photon catch of the L and M cones (see the section ‘Spectral Encoding in the Monochromat’ for an explanation of how cones act as spectral filters). Magenta areas are those in which the L cone absorption is greater than the M cones absorption. Green areas represent the opposite

state of affairs, where M cones have greater absorption than L cones. Brightness represents the extent to which absorption is greater in one of the other cone. Thus, magenta and green form a continuum, with bright magenta at one end, bright green at the other, and grey at the mid-point between them representing equal absorption. In Figures 9 and 10, these Red/Green images have been re-rendered using a greyscale instead of the magenta/green scheme in Fig. 8.

Figure 9 offers a good illustration of how a quick comparison of chromatic and luminance contrast can sharply differentiate objects from shadows. The original image, Figure 9(a), depicts five brightly coloured hats in row. The Luminance Image in Figure 9(b) presents the total photon catch of the three cones, the information available to the luminance system. Figure 9(c) illustrates the luminance *edges* as discerned by an edge detection program. Given the luminance edges alone, the boundaries of the hat brims are conflated with the shadows created by the hats. Figure 9(e) represents the spectral contrast edges derived from the R/G chromatic image in Figure 9(d). Here, the outlines of the hats alone are visible. By comparing the luminance and chromatic images we can also see clearly that the wall is textured, not mottled in colour. (Interestingly, the wrinkles in the caps do show up in the R/G spectral images—a good reason to suspect that there is spectral inter-reflectance between the caps.)

The second set of images (Figure 10) depicts a more complex natural scene. Here the original image has both fine and gross spatial detail plus a good range of both luminance and spectral contrasts. In this instance, there is clearly a symbiotic relation between the chromatic and luminance images vis-à-vis edge information. Because the pink flowers of the cacti have a primarily red/green spectral contrast with their background the R/G chromatic image

showcases the flower edges. The luminance edges in Figure 10(f) highlight the white spikes



Fig9a. Full colour image.



Fig9b Luminance image



Fig9c Luminance edges

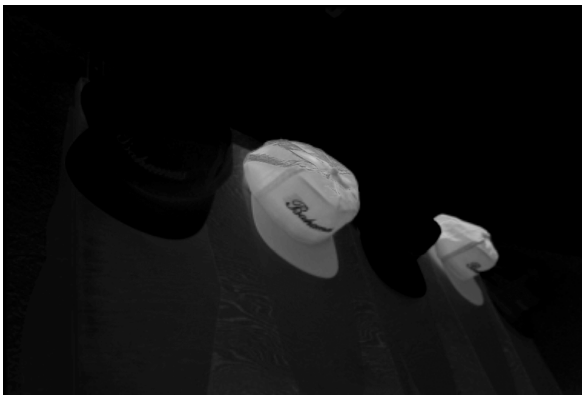


Fig9d Red/Green spectral image

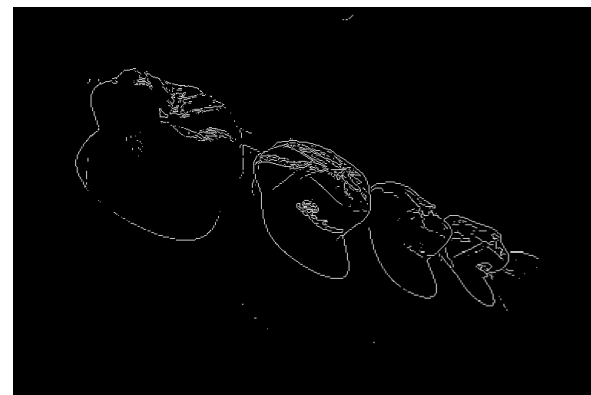


Fig9e Red/Green spectral edges

Original colour photograph, luminance image and R-G image supplied by Karl Gegenfurtner; images with contrast edges supplied by permission of Bistra Dilkina and Kathleen Akins.



Fig10a Luminance image

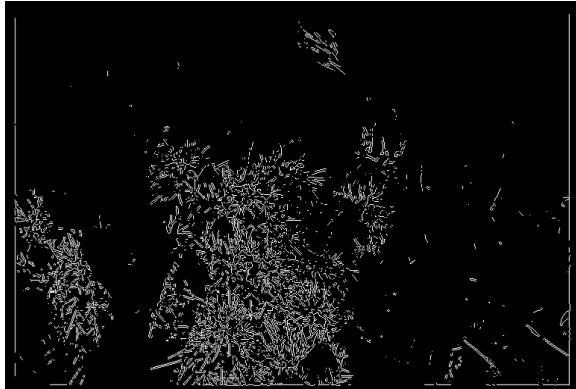


Fig10b Luminance edges



Fig10c False colour R/G image

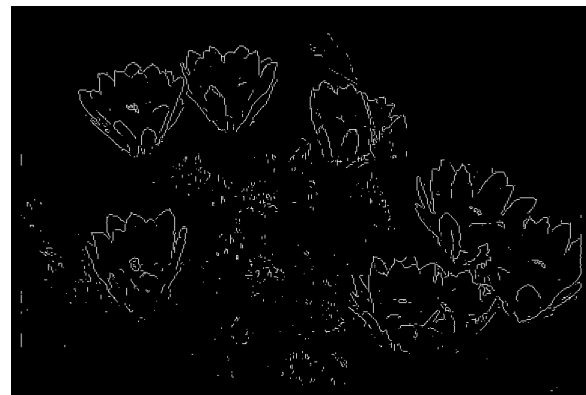


Fig10d R/G image edges



Fig10e Original colour image (e)

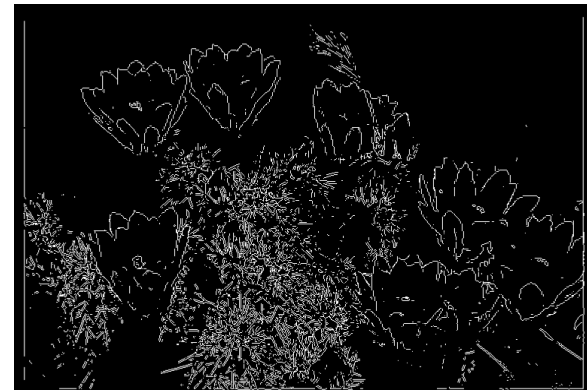


Fig10f Combined edges

Original image, luminance image, and R/G image courtesy of K. Gegenfurtner. Three images with calculated edges courtesy of Bistra Dilkina and Kathleen Akins. Note the difference between the edges computed the R/G image and those from the luminance image above. Clearly, both systems, chromatic and luminance, provide quite different information about the original colour image.

of the cacti and the blades of grass. The luminance and R/G chromatic edges, *taken together*, delineate most of the scene. The central image-processing problem for this image, then, is not the disambiguation of objects from shadow. It is how to discern any object edges at all in those areas with low intensity contrast. The combined resources of the two systems yield easily distinguishable object boundaries. Chromatic contrast complements luminance

contrast and vice versa. It is probably now clear why adding a chromatic system has utility. First, it is true that any object that differs in colour from its background will also differ in lightness. But it is not true that spectral contrast and luminance contrast are *the same* in the visual image. The two will often differ in their *spatial location* as well as in their *strength* and *sign* (positive or negative). Indeed, recent scene analysis of 700 images, of both natural and artificial scenes has shown that in human vision R/G chromatic information and luminance information provide an independent sources of information about the distal world (Hansen and Gegenfurtner 2009).

Second, as briefly mentioned above, redundancy is a very useful property. Redundancy increases computational accuracy and efficiency in any computational system with a high level of noise—and luminance vision is a prime example of a noisy system. Where chromatic and luminance contrast align—at object borders and around patches of surface color—summation of these edges will yield higher *visual* contrast. (If *some* contrast is useful, *more* can be much better.) Alternatively, this parallel source of information can be used to construct a ‘fail safe’ system, an old engineering trick for reducing risk. Construct two parallel but independent mechanisms to compute the same function (perhaps using different algorithms). Insofar as the two results agree, the world is probably as represented. But if the answers do not agree, more resources must be thrown at the problem: There is only one way that the world could be. In sum, *chromatic contrast is beneficial because it provides an additional, independent source of visual information, one that can be used in parallel and complementary ways or combined with luminance signals for a more reliable signal.*

One obvious question probably remains for the (still conscious) reader: What explains Hubel and Livingstone’s results? Perhaps most importantly, Livingstone and Hubel’s experiments asked the wrong question. They asked whether each system, of and by itself, was *sufficient* for the perception of various visual properties, i.e. a question about independent function. When two systems normally work together, however, testing them in isolation reveals little about their individual contributions to normal function. This principle is abundantly clear when we think about motor systems. Take a person with two legs who walks with a normal smooth gait. Now, remove the left leg and observe the result. With the right leg alone, the subject hops up and down, frequently overbalancing. But based on this observation, we do not infer that either the right leg or the left contribute ‘hopping’ to the smooth gait of the two-legged human. Neither leg ‘hops’. Here, the moral is readily apparent: Separate tests of a system’s interdependent components need not reveal their individual functions under normal conditions. This is a moral that applies equally well to integrated sensory systems, but a point that was obvious only in retrospect.

⁹ The psychophysical experiments also suffered from a very practical problem: the isoluminant images were too dim (cf. Footnote 6). Because a chromatic system *compares* signals, chromatic systems require greater energy to function; they have higher thresholds. So a fair test of any chromatic system must use stimuli that are proportionately brighter. Once the intensity of chromatic stimuli is increased, overall perception of object shape, depth relations and surface textures improves Kingdom, F. A. A. (2003). "Color brings relief to human vision." *Nat Neurosci* 6(6): 641-644, den Ouden, H. E. M., R. van Ee, et al. (2005). "Colour helps to solve the binocular matching problem." *The Journal of Physiology* 567(Pt 2): 665-671, Kingdom, F. A., K. Wong, et al. (2006). "Colour contrast influences perceived shape in combined shading and texture patterns." *Spat Vis* 19(2-4): 147-159, Kingdom, F. A. A. and R. Kasrai (2006). "Colour unmask dark targets in complex displays." *Vision Res* 46(6-7): 814-822, Kingdom, F. A. A., K. Wong, et al. (2006). "Colour contrast influences perceived shape in combined shading and texture patterns." *Spatial vision* 19(2-4): 147-159, Gheorghiu, E. and F. A. Kingdom (2007). "Chromatic tuning of contour-shape mechanisms revealed through the shape-frequency and shape-amplitude after-effects." *Vision Res* 47(14): 1935-1949, Shimono, K., S. Shioiri, et al. (2009).

V. ENCODING THE SPECTRAL DIMENSION OF LIGHT.

A. 'General' versus 'Specific' Encoding

As we said in the introduction, a central thesis of this paper is that the human colour system represents a general encoding of wavelength information. Here, 'general' contrasts with 'specific'. One common (and true) saying among neuroscientists is that the more simple the behavioral repertoire of the organism the more *specific* the information initially encoded and vice versa—the more complicated the behavioral repertoire of the creature, the more *general* the information encoded (Lennie, Trevarthen et al. 1990). To see the difference between a general and specific encoding, begin with the frog's four classes of retinal ganglion cells (ganglion cells are the 'output' cells from the retina)(Letlvin, Maturana et al. 1959). One kind of cell is sensitive to borders between light and dark areas, another to moving edges, another to reductions in overall illumination, and the last to small dark moving convex edges. The information processed by each ganglion cell of the frog's retina is considered 'specific' because of the discrepancy between the complex properties of the retinal image and the limited number of luminance patterns to which the ganglion cells react. For example, sit a frog in front of the Mona Lisa. The image falling on its retinae will form a complex pattern of wavelength and intensity contrasts—roughly the same pattern that would fall on the back of your eyes if you were looking at the Mona Lisa. But because the ganglion cells of the frog respond to only four distinct patterns of illumination, few ganglion cells will respond—probably only the 'border detectors' will fire, responding to the frame of the picture. The frog's ganglion cells miss or *filter out* most of the characteristics of the proximal stimulus. (This is also why amphibians do not edit aesthetics journals.)

For the frog, of course, such highly filtered information is behaviorally appropriate. As Letlvin, Maturana et al. (1959) first claimed, in the natural environment of the frog, light patterns on the frog's retina are correlated with particular states/events of the world. A pattern of dimming illumination is statistically correlated with the movement of a large-ish opaque object, often a predator; small convex moving edges correlate with flying insects; vertical borders between light and dark areas are usually caused by the edges of objects such as tree trunks, and; moving edges indicate either moving objects beyond the frog or the frog's own movement. Thus, one finds a more or less direct wiring from these classes of cells to pathways that control the frog's behavior in relevant ways. The ganglion cells that react to looming shadows stimulate evasive behavior; those that are sensitive to small dark moving spots cause tongue-swiping behavior and so on.

In a mammalian or primate visual system, the initial encoding of the retinal image must subserve an indefinite number of complex behaviors. As *Homo sapiens*, we can pick our way through a tangled forest floor, recognize an old acquaintance across a concert hall, pick green peas, catch (or dodge) a speeding missile or secure a squirming baby into a pair of overalls. Such diverse behaviors, each guided largely by visual information, have wide-ranging and sometimes competing, informational requirements. The initial retinal encoding—the 'front end' on this large and complex system—must be compatible with the informational demands of each of the visual tasks that lie downstream.

More concretely, in a general encoding the receiver must satisfy the twin constraints of *range* and *resolution* relative to the appropriate dimension of the stimuli. It must encode the stimuli *across* a wide a range as possible and it must be able to make fine discriminations *within* that

range.¹⁰ For example, a diurnal visual system must be able to accommodate the large range of luminance conditions from dawn to dusk (or, in a nocturnal animal, from darkest night to the illumination of a full moon). At the same time, within that wide range, it must be able to differentiate between very small differences in intensity in to see anything at all. A general encoding thus finesses both range and resolution constraints.

Despite the above talk of a continuum of encoding from most specific to highly general, it is important to realize that a completely general encoding (whatever that would mean) is not an ideal of any biological system. Every sensory system must satisfy a temporally shifting body of often competing pressures—constraints imposed by that species' environment, its skeletal and musculature system, its behavioral repertoire, the relative importance of processing tasks, and so on. Inevitably, compromise occurs, and generality is lost. Yet this is a loss that makes no difference given that generality has only pragmatic value.

Consider the problem of implementing high acuity vision in mammals. What is needed to see the fine spatial detail and what are the limitations on this ability? At bottom, spatial resolution is limited by the size of each receptor's visual field—how much of the visual world each cell monitors. To gain high spatial resolution across a mammal's entire field of vision requires a retina that is tightly packed with receptors. In practice, this solution is unworkable. For one, if the entire retina were as densely packed as the fovea, the retina would send an astronomical number of axons to visual cortex—approximately one billion inputs. To house all that hardware, your head would need to be, well, *large* (and perhaps come with its own set of retractable wheels for mobility). It would also be an open question whether all that input would serve any purpose, whether it would increase accuracy or reliability to any significant extent. Yet another problem with such a retina would be its very low sensitivity. Very small receptive fields (like butterfly nets) catch fewer photons. Hence the evolved solution: There is a small, central, dense area of photoreceptors, the fovea, with a large peripheral area that increases in sensitivity but decreases in spatial acuity the further out from the fovea. In the turn, the fovea can be moved from spot to spot as needed. Thus a compromise is struck by a division of duties, one available to any mammal with moveable eyes (i.e. not rabbits).

This example nicely illustrates why a 'completely' general encoding is *not* an ideal of biological sensory systems. Any system with finite resources faces competing constraints—whether it's the size of your head and strength of your neck or the number of receptors that can be fitted into a finite space—and such constraints themselves lead to elegant and efficient neural organization. For these reasons, 'tailored' sensory encodings are the norm in biological systems (cf.(Akins 1996)). Indeed, how such constraints and compromises have shaped our own system of spectral encoding will be an important part of the story below.

B. Spectral Encoding and The Monochromat. As we said above, strictly speaking, monochromats are colour blind. Because the response of a single cone conflates amplitude, wavelength and polarity information, a monochromat cannot discriminate between visual stimuli on the basis of wavelength alone. *Nonetheless, a monochromat makes use of spectral information even if wavelength is not explicitly encoded.* Given the bell-shaped response curve of photoreceptors to light across a range of wavelengths, each type of photoreceptor acts as a *colour /wavelength filter*. That is, the photoreceptor responds to light only within a limited range of wavelengths and within that range, it prefers—preferentially absorbs—some

wavelengths to others. Thus every photopigment ‘filters out’ those wavelengths to which it either has no response or responds minimally. E.g. An S cone responds maximally to light at 420 nm while the same wavelength produces only a minimal response in an M cone (see Figure 4).

Consider the use of colour filters in black and white photography, a practice that often strikes the non-photographer as nonsensical. If one uses modern ‘colour corrected’ black and white film, variations in the gray scale of the image signal intensity differences in the reflected visual scene. When a color filter is attached to the camera lens, however, the luminance values of selected wavelengths are changed. Figure 11a, for example, is a photograph taken without a colour filter. When a filter is added to the camera, one that blocks ‘red’ or long wavelength light but which transmits short wavelength or ‘blue’ light (Figure 11b), some of the fruits and vegetables turn ‘black’. These are the fruits and vegetables that reflect light predominantly from the red/orange range of the spectrum. Here, the blue filter makes the red fruit visually salient not by ‘colouring’ it red but through luminance contrast enhancement (making the objects of interest visible against a background).

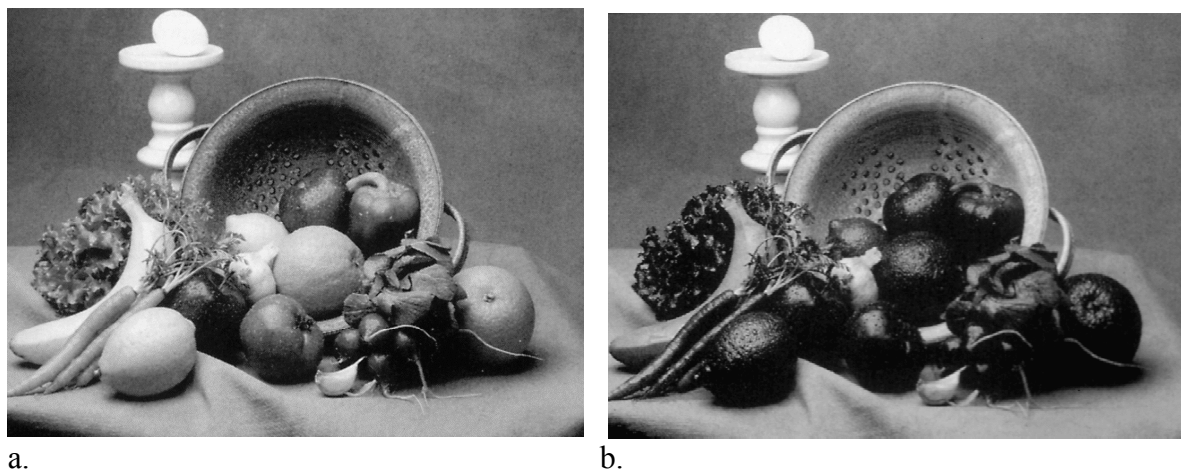


Fig 11a Black and white photograph without colour filter **Fig 11b** Identical scene photographed with a blue filter on camera

If we think of each photoreceptor as a *de facto* spectral filter, it seems obvious that ‘tuning’ photoreceptors could be hugely advantageous in the processing of natural scenes, e.g. for any animal that forages for red berries hidden in green foliage. One of the first researchers to both see and to test this hypothesis was Lythgoe (1979) who suggested that both the rods and cones of certain fish function as spectral filters that highlight the contours of objects in aquatic environments. Although we rarely think about it, water *is* coloured and the colour of a particular body of water depends upon both the depth of the water and the type of organic particles in suspension. Depth matters because when sunlight enters water some light is transmitted and some is absorbed, a process is wavelength selective. Clear ocean water with little organic matter selectively absorbs red and violet light, wavelengths from either end of the spectrum. By about 45 meters of depth, only the intermediate ‘blue’ wavelengths are left in ocean waters. In contrast, marshes and swamp contain decomposing plants, tannins, lignins and yellow-green plankton, all of which absorb the short ‘blue’ wavelengths. By a depth of only 3 meters the water is a dark red-orange. These facts explain the difference

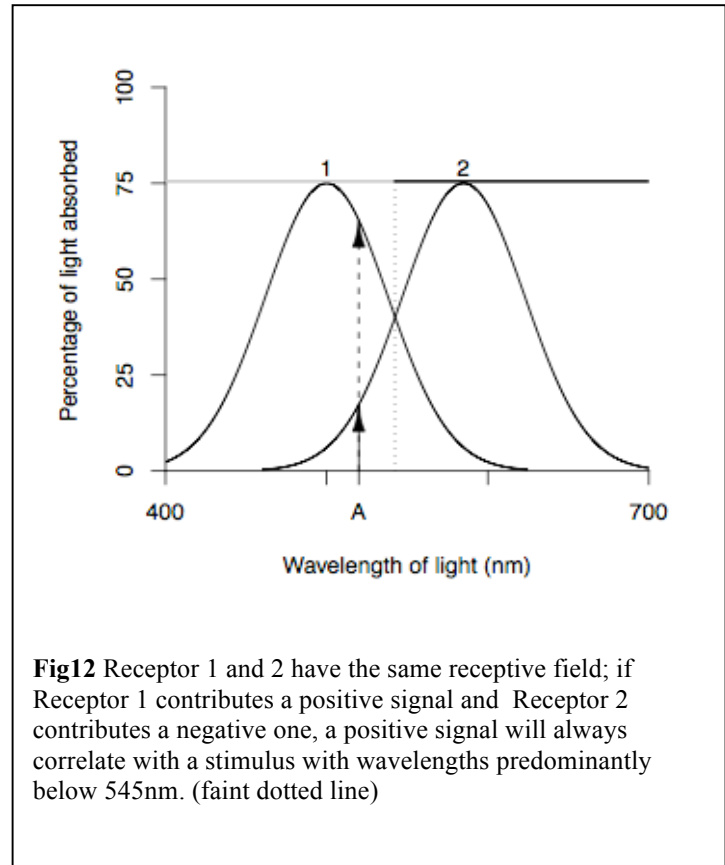
between the shimmering blue colour of the Mediterranean and the dark brown, brackish waters of a swamp. In effect, then, water forms a coloured backdrop for the inhabitants of any underwater world and this backdrop can be used to the advantage by both predators and prey. If the cone photopigment of the predator is ‘matched’ to the colour of the water, it will be maximally sensitive to the background. Any dark object will be highlighted against the brighter background. Conversely, if the photo pigment is ‘offset’ from the dominant wavelength of the ambient light, it will maximally *insensitive* to background — i.e. be perceived as even darker than it is. An offset pigment will create a dark background against which to spot bright prey. Thus a surface-feeding fish, looking up at dark prey against a bright sunlit background, will often have photopigments that are matched to the water colour in the lower half of the retina (the lower half looks upwards). Conversely, the upper retina of a bottom-feeder often contains offset photopigments, thus highlighting the lighter prey against the murky bottom. For example, the skipjack tuna, which spots its prey from below, has only one photo pigment that is matched to the background light. The walleye, bluegill and piranha all inhabit dark, particle-laden (hence, red-shifted) waters. Near infra-red light is common in this ‘black’ water habitat especially during the dusk and dawn hours when surface feeding occurs and as one guess, all of these species are dichromats with one cone type ‘matched’ to the near-infrared. Finally, the pacific salmon, which migrates from fresh water streams to the ocean, has rods that change in spectral sensitivity during migration. This shift maintains the ‘match’ with the background as the salmon migrate between water of different colours (Munz and McFarland 1975; Lythgoe 1979; Levine and MacNichol 1982).

The above examples of spectral filtering in fish show how, by highlighting the difference between figure and ground, spectral information can be used for predation. However, because *all* photoreceptors have bell-shaped tuning curves, *all* photoreceptors are spectral filters. Hence every visual system—whether it belongs to fish or fowl or something on the hoof—runs up against the same problem: Are (my) photoreceptors, individually and taken together, likely to enhance, obscure, or be merely neutral vis-à-vis the visibility of (my) objects of interest? The survival of every predominantly visual species is strongly tied to photopigment sensitivity. Moreover, because the tuning of each receptor can shift from generation to generation as a result of genetic drift, *spectral tuning represents a species’ specific adaptation for the enhancement of visual contrast*. In our own case, for example, the tuning curves of all three cones, taken together, mimic the spectral power distribution of daytime sunlight at the earth’s surface. In this way, our daylight system maximizes luminance contrast. And in exactly the same way, the spectral sensitivity of the human rod system is matched to the SPD of moonlight so as to maximize night vision.

B. Spectral Encoding and The Dichromat. Let us move on, then, to the encoding of spectral information in a system with two cones. In the first puzzle of colour vision above, the paucity of mammalian trichromats, the conclusion seemed to be that dichromats gained, through the addition of a second cone, only *bad* colour vision. A different way to understand the same facts, however, is that a second cone makes possible the comparison of *ranges* of wavelengths, a ‘low’ end and a ‘high’ end of the spectrum (Figure 12). Take two differently tuned receptors looking at the same circular region of the visual field. Let Receptor 1 contribute a positive signal and Receptor 2 a negative signal to a single ganglion cell. Given this antagonistic relation, a positive cumulative response will signal that the light contains wavelengths predominantly in the lower range of the spectrum; a negative response will signal light predominantly in the upper range. A ‘null’ response, when the responses of the receptors cancel each other out, will signal an approximately equal distribution of light across

the spectrum.¹¹ In other words, this kind of cell would signal ‘positive’ versus ‘negative’ *spectral contrast*.

In primitive visual systems, antagonistic pairs of cones like those in Fig. 12 are used for simple motor tasks. Take as an example the crustacean, *Daphnia pulex* (Baylor and Smith 1953; Smith and Baylor 1953; Menzel 1979). In blue light, *Daphnia* become highly active, lean forward, and move steadily at right angles to the angle of illumination; in red light, *Daphnia* remain upright and move slowly in parallel with the angle of illumination. One hypothesis is that this is related to foraging behavior. In phytoplankton-rich layers of water, the chlorophyll of the plankton absorbs blue light and hence produces red ambient light (the so-called ‘red tide’ with which human shellfish lovers are familiar); in phytoplankton-poor layers, the water is more blue. Thus, in foraging, when *Daphnia* reach a red layer of water, they slow down to eat, following the horizontal layer of phytoplankton. When the water becomes blue, *Daphnia* begin to move *through* the layer, looking vertically for more phytoplankton rich red water. In this way, *Daphnia* exhibit a typical pattern of wavelength-specific behavior, in which blue and red light produce positive and negative phototaxis.¹²



As the reader will be quick to point out, phototaxis is not vision. To actually see—to have vision—is to delineate the furniture of the distal world. In *Daphnia*, both cones monitor the same region of visual space and together discern the predominant colour of light within that common field. So a ganglion cell of this kind, with a spatially homogenous visual field, does *not* encode wavelength contrast across a *spatial boundary*. Object vision (as opposed to phototaxis) requires a way to discern where and to what extent spectral contrast exists in the image—i.e. spectral contrast across a spatial border.

All things being equal, one would expect that spectral contrast encoding would follow in the footsteps of luminance contrast encoding, via the spatially-opponent center-surround cell. In the luminance system, each center-surround ganglion cell ‘watches’ a particular circular area of visual space, an area that is sub-divided into a ‘center’ (a smaller circular area in the middle) and a ‘surround’ (the outer concentric region). There are two basic classes of center-

¹¹ More correctly, such an arrangement captures spectral contrast as filtered by the two receptors involved.

¹² Note that although no one knows, for *Daphnia*, exactly how this behaviour is affected, there is no suggestion here that copepods see colours. Rather, what is postulated is only a simple mechanism that compares the outputs of two differently tuned light receptors and that then produces a positive or a negative signal, each wired

surround cells, ‘ON’ and ‘OFF.’ When a bright light is shone in the center region of an ON-center cell, it fires in response to the illumination; when light is shone on the surround, the cell is inhibited (visual cells fire spontaneously, the only neurons to do so). An ‘OFF-center’ cell has just the opposite response (i.e., it is excited by light in its surround but inhibited by light in the center region). A center-surround cell signals either positive or negative luminance contrast between two regions of visual space. Call this a center-surround *luminance* cell.

Suppose, then, that one took the existing hardware in the retina, a standard center-surround cell, and hooked up the center and surround to *different* types of cones, say the center area to a ‘red’ cone and the surround area to a ‘blue’ cone. This cell would be a *chromatic* cell with a spatially opponent organization (by definition). However, it would *not* be a cell that signaled wavelength contrast across a spatial border—i.e. ‘predominantly red light’ versus ‘predominantly blue light’. Rather, this cell would have a complex informational content, a multiplexed signal of wavelength and intensity contrast. Given a moving colour border passed through its center, it would respond by reversing its response (either it will start firing or it will stop). This is why such cells are called chromatic: They respond to a border defined by wavelength alone. But chromatic cells also react a bright white light targeted upon their centers, to a luminance edge, and to a wide variety of ‘contrast stimuli’ composed of various wavelengths and intensities. So the tactic of ‘up-cycling’ a luminance center-surround cell into a chromatic center-surround cell does not work—at least not if one needs a ‘colour detector’. Each chromatic cell *reacts to* a specific wavelength difference—but it does not *signal* that wavelength difference given its complex response properties.¹³

Below, we will return to the question of whether this ‘failure’ of the chromatic center-surround cells matters in the grand scheme of human vision. Here, simply note that, a center-surround chromatic cell has an organization that piggy-backs nicely upon a cell with a simpler organization, the luminance ganglion cell. It is easy to see how this more complex version, the chromatic cell, might have evolved. Moreover a chromatic cell does respond to wavelength differences among many other types of visual contrast. It discriminates between wavelengths. And, of course, evolution latches onto what works, not what, from the point of view of the universe, might have been an ideal solution. This is a (Rolling Stone-eque) theme to which we will return to in a moment. For now, however, note that the obvious solution to wavelength encoding—a center-surround chromatic cell—must fail as a *general solution* to the problem of wavelength encoding.

The fly in the ointment—nay, the elephant in the ointment—is chromatic aberration. Consider a chromatic center-surround cell with a ‘blue’ surround and a ‘red’ center. Given the two different tunings of the cones, the blue cones prefer ‘blue’ light, while the red cones prefer ‘red’. Recall from above that chromatic aberration, the refraction of light as it passes through the lens, is a wavelength sensitive process: long wavelength light is ‘bent’ less than short wavelength light. So when a stimulus composed of both red and blue light passes through the lens, two separate images are created: a ‘red’ and a ‘blue’ image each of which occupy slightly different spatial locations on the retina. Now take a chromatic cell, with an S cone center and an L cone surround. When such a cell passes over a red/blue border, the center and surround do not agree upon where exactly the colour border occurs. There are two separate images and the L and S cells respond to different ones. For this particular colour

¹³ That would require *two* spectrally opponent cells, an L/S cell and a S/L cell, wired together in an opponent fashion. Their combined output would indeed produce a ‘genuine’ center-surround cell for spectral contrast

opponent cell, then, a center-surround configuration makes little sense. It cannot do just what it ought, namely locate with precision a border of spectral contrast. Importantly, this is a problem for *any* scheme of spectral contrast encoding across a spatial border.

Alas, there is worse to come. One of the central benefits of a dichromatic system is that it allows the visual system to significantly extend the species' range of visible light. Given the spectral power distribution of sunlight at the surface of the earth, there is a very broad range of light with useful energy. The sharp tuning curve (i.e. small range) of an individual cone, however, limits each cone's response to a reasonably small portion of this spectrum.¹⁴ Thus a second cone significantly extends *what is visible to the species*. (To get an intuitive sense of what the monochromat is missing, buy a pair of very dark sunglasses and then return to reading this paper. The issue of what the monochromat is missing in rather drier terms of information, see Akins (Akins Forthcoming), forthcoming.) To optimize the visible range, the ideal configuration is to have two cones that overlap very little or not at all.

On the other hand, there are also good reasons why the two cones of a dichromatic system should be close together. As we saw above, given chromatic aberration, the greater the overlap between two types of cones, the less the chromatic aberration; the less the aberration, the greater the spatial acuity. So whenever high acuity is of value, extensive overlap is important. Second, as the reader has probably realized, there is no such thing as "the" luminance system or "the" chromatic system. There can be many different luminance systems, single cones systems that operate over different spectral ranges. Similarly, there are indefinite possibilities for chromatic systems each of which makes a specific type of wavelength discrimination. If you are a panda and you eat only bamboo, then you need to select the tender green shoots. You will be well served by two very similar middle-range cones for greater discrimination among the greens. Here, optimal contrast enhancement favors cones with very similar peak sensitivities.

The upshot of the above considerations is that there are a variety of pressures upon cone sensitivity. Some pressures seek to drive overlap to minimum; others pressures have the opposite effect. The dichromat faces a dilemma: which way to jump?

As one would guess, this dilemma has been resolved on a case-by-case basis. On one end of the continuum, some dichromatic rodents have two, spectrally independent cones—one medium/long wavelength cone plus a non-overlapping short (near UV) wavelength cone. Presumably, in a species with nocturnal behaviour, spectral range under scotopic (low light) conditions is far more important than spectral contrast during daylight vision. Hence, spectral contrast is given up entirely. Occupying the middle ground, there are the dichromatic diurnal primates. All such primates have one short wavelength cone that peaks in sensitivity around either 430 or 435 nm. plus one medium/long wavelength cone, two cones the overlap only slightly in sensitivity. Thus the retina of dichromatic diurnal primate encodes spectral contrast at low spatial resolution (through the overlap of 'blue' and 'green/red' cones) plus a good spectral range of useable sunlight through the combined ranges of both cones. (Note that there is no mammal that occupies the other end of the continuum, which has a visual system with high spatial resolution spectral information but with little spectral range.)

¹⁴Why isn't it possible, then, for the response of a given cone to cover the entire range of the spectrum—for a single cone to have a very broad response range? The problem here a broadly tuned cone is a very noisy one. Hence, whether or not such photoreceptor pigments are possible, it is not a viable option to have cones with

In the case of Old World Primates, nature seems to have stumbled upon a particularly Good Trick, as Daniel Dennett would say. It has added *two additional* cones instead of one. Two of the cones overlap extensively (the M and L cones) and hence limit chromatic aberration. The third cone (the S cone) is the outlier for the purpose of spectral coverage (Figure 4). Thus, one gains two distinct but interacting systems for spectral contrast: a high acuity L and M cone system ($L - M$) and a low spatial acuity S cone system ($S - (M + L)$). Taken together, they form a single system of spectral contrast with an excellent range. We *Homo sapiens* have the best of both worlds: We can have a broad range of visible light and good spectral contrast/wavelength discrimination as well.

C. Spectral Encoding and the Human Trichromat. We can now summarize the physiological details of the human trichromatic system and show, finally, how and why the human colour system exemplifies a general encoding of spectral contrast information. In the (phylogenetically) older S cone system, the ganglion cells are spectrally antagonistic: they pit the S cone signals against a combined M and L cone signal, i.e. an $S - (M + L)$ signal. In virtue of the ‘spatial smearing’ of short wavelength light, however, many S ganglion cells have a *homogeneous receptive field without a center-surround configuration*. They also have relatively large receptive fields (i.e. they ‘watch’ a comparatively large part of the visual world). Thus, like the retinal cells of *Daphnia*, they signal ‘predominantly blue light’ versus ‘predominantly yellow light’ (hence their common name ‘blue/yellow’ ganglion cells) within a single area of retinal space.

Although no blue-yellow cell by itself encodes spectral contrast across a spatial border, nonetheless spectral contrast can be inferred from the population of blue-yellow cells as a whole. To do so, a visual system requires a way to keep track of the spatial location, within the retina, of each blue-yellow ganglion cell. This task is accomplished, in all mammals, by means of retinotopic maps in visual cortex. (This is always what *Daphnia* lack, a detailed spatial map.) With this orderly arrangement, the system can compare the signals of nearby cells and discern where blue-yellow borders occur. When one population of cells signal ‘predominantly blue light’ and a contiguous population says ‘predominantly yellow light’, there is a blue-yellow border between them. Of course, a large drawback of this form of spectral encoding is that it has very low spatial resolution. Remember that a luminance ganglion cells compares the photon catch between the center and surround regions, a spatial border *within* that cell’s receptive field. The blue-yellow system differentiates between the signals of two separate populations of ganglion cells, hence between visual fields as a whole. Moreover, S-cones are very scarce (recall that only 5% of all cones are S-cones) which results in much larger visual fields of the blue/yellow ganglion cells (relative to luminance cells in general). For all these reasons, this older chromatic system encodes only low spatial resolution blue-yellow (spectral) contrast.

The phylogenetically newer, and predominant, M and L cone system makes for a much more interesting story. With the problem of spectral range finessed by the B/Y system, this allows for the possibility of a chromatic system in which the tuning curves overlap extensively, the L-M system. This close overlap of the M and L cones minimizes chromatic aberration and that, in turn, makes it possible the use of the center-surround organization typical of luminance cells. This gives the M-L system better spatial resolution than the B/Y system.

What about the ‘problem’ of wavelength encoding? Recall that a chromatic center-surround cell does not signal wavelength contrast *per se*. But as the Rolling Stones said (as opposed to

Quine or Sellars), this may not matter. What saves the situation, and makes the R/G system an integral component of human vision, is the ‘fit’ between the luminance and R/G chromatic systems. As we saw above, the two systems, taken together, respond to far more contrast in the retinal image than either system alone. Although neither system encodes just one dimension of the stimulus—wavelength and intensity are not encoded by the chromatic and luminance systems respectively—the signals of the two systems are independent, one of the other. It is this independence that poses the potential for a huge informational advantage of the two-component system. Independence allows a system to compare the signals of each system, and to use that comparison as a reliable guide to real difference in the world. In particular, chromatic cells respond to wavelength differences, a feature that a luminance cell simply does not have. By comparison and in combination, the retinal R/G chromatic and luminance signals can be used to discern, where relevant, wavelength and intensity properties of the retinal image—and with that information of the probable causes of the patterned response of photon absorption by the receptors. Chromatic cells give us what we need.

Importantly, the addition of the M cones to the fovea did not result in *any* loss of spatial resolution, contrary to the accepted *a priori* reasoning about the costs of a third cone. Although we do not know how the trick is done, the visual system can infer (at least in the crucial central 2 degrees of fovea) the following counterfactual condition—what an L cone’s response *would have been* had it occupied the location of the M cone in question (Williams, Sekiguchi et al. 1991). *For the purposes of computation, then, an L cone and M cone responses seem to be interchangeable for luminance computations.* This is why we can afford to have a trichromatic fovea: it doesn’t cost anything to add a third cone, counter to all expectations. In the central 2 degrees of the fovea, there are *only* M and L cones (no S cones) and this particular duality of cones preserves the spatial resolution of a single cone luminance system—a very good trick indeed.

Finally, the particular wiring of the M and L chromatic system has additional benefits. For many years it was assumed that every chromatic cell must be wired selectively such that the surround received input from one kind of cell (e.g. L cells) while the center region received input from the other type (e.g. M cells). This was based on our knowledge of the blue-yellow system which *is* selectively wired. We now suspect, however, that chromatic opponency in L-M cells depends upon random M and L inputs to the center and surround. (Dacey, Packer et al. 2000; Dacey 2000; Dacey, Diller et al. 2000). In the central two degrees of the fovea, each ganglion cell center is ‘fed’ by just one cone, either an M or L cone for the greatest possible spatial resolution. However, the surround area receives a random combination of L and M cone inputs. The result is that the ‘pure’ center (fed by either an M or an L cone) will always differ in wavelength sensitivity from any ‘combined’ M and L surround. In effect, numerous new types of spectral filters are created by the random input to the surround. Out beyond the central 2 degrees of the fovea, the centers of the L-M cells receive input from more than one receptor type. Such cells will continue to be chromatic cells as long as the combined spectral sensitivity of the center differs from that of the surround. At the periphery, the receptive field of the L/M ganglion cells is so large, that random inputs no longer create an effective difference in wavelength sensitivity. Eventually such cells become luminance cells. But within the fovea and parafovea, L-M cells form a population that is responsive to very fine-grained wavelength differences (Mullen and Kingdom 1996; Mullen and Kingdom 2002; Crook, Manookin et al. 2011). Pass a coloured border of any kind (within the spectral range of M and L cells) across the retina, and *some cell or other* will

reverse its response. The net effect is a system for fine-grained *wavelength discrimination*.¹⁵

The two systems taken together, L-M and (M+L)-S, represent about as general an encoding of spectral contrast as nature is likely to produce. The S system has a very wide spectral range but provides only coarse-grained chromatic contrast at low spatial resolution. The L-M system has a limited spectral range. But in compensation, it has very fine-grained wavelength contrast, reasonably good spatial resolution and, moreover, does not interfere with the high spatial acuity of luminance vision.¹⁶ Considered on a continuum from highly specific to very general encodings, spectral information in the human retina can be regarded as ‘reasonably general’: It provides spectral contrast information over a wide range of wavelengths, and it has the capacity for fine-grained wavelength discrimination—both spatially fine-grained and spectrally fine-grained—within a selected range of visible light.

VI. THREE PUZZLES REVISITED

It is now time to return to the three physiological puzzles with which we began, to see how the view of Spectral Vision resolves them.

Why is Trichromatic Vision an Anomaly in Diurnal Mammals? On the traditional view of Colour-for-Colouring, there is a standard explanation of why dichromatic systems are better than monochromatic ones, and why trichromatic systems are better still: First there was crude wavelength discrimination, and then there was better wavelength discrimination, but ‘good enough’ to yield useful colour perception.

On the view of Spectral Vision, most diurnal mammals have two cones because of the multiple benefits of dichromaticity. A second cone serves to: (a) increase the range of light that can be used by the visual system; (b) add a second ‘filter’ that can put to a novel use (e.g. looking up at prey as opposed to down at the ocean bottom) and (c) make available low spatial resolution chromatic contrast. The advantages of chromatic contrast are themselves multiple. At the most basic level, chromatic contrast represents a new source of information about the retinal image, visual contrast information that the luminance system does not encode.¹⁷ And although the spatial resolution of the luminance system far exceeds that of the S - (M+L) system, the low acuity chromatic signals can still be used in tandem with either the L-M or luminance systems for a *limited* set of tasks. This *can* work, we now realize, because the human vision processes spatial information at a number of different spatial scales. Thus a border, discerned by the blue-yellow system at low spatial resolution can delineate an object border that, on the basis of the luminance signal alone, was ambiguous (i.e. not clearly a shadow border or an object border). This is recent realization and vision science is just beginning to untangle the Blue-Yellow channel’s contribution to specific visual tasks ((cf.

¹⁵ Just outside of the fovea, in the parafoveal region that surrounds it, *both* the center and the surround areas of these cells are innervated by a random, aggregate M and L signal. Thus, beyond the fovea, the M/L cells are not spectrally opponent; rather they encode merely luminance contrast. The parafovea is thus rendered ‘spectral contrast blind.’ This explains why only the fovea has ‘colour vision’—not because there are no cones outside of the fovea but because the ganglion cells do not have spectral contrast.

¹⁶ Lest anyone ask whether such a ‘general’ system for spectral contrast encoding is really general enough to deserve the name, the luminance system displays much the same kind of compromises between spatial acuity, luminance range and luminance contrast sensitivity but the argument for this view would be quite long.

¹⁷ We say ‘may not’ because the scene analysis done by Hanson and Gegenfurtner was done for the human visual system—i.e. the luminance function for daylight vision and the M-L color axis. The assumption is that

(Mullen, Beaudot et al. 2000; Michna, Yoshizawa et al. 2007; Leh, Ptito et al. 2010).¹⁸

These multiple distinct advantages, taken together, seem adequate to explain the universal nature of the dichromaticity in diurnal mammals. But they also explain why mammalian dichromatic systems do not form a homogeneous set. What drove the evolutionary process were the advantages of spectral information to different species given the very specific constraints imposed by each species' behavior and its environment. As a result of this diversity, each species came to have a somewhat different spectral tuning than its two-coned fellows (Jacobs 1993).¹⁹

Finally, it is clear why so few diurnal mammals are robust trichromats. The trichromaticity of Old World primates depended upon very specific conditions for it to arise and be useful. According to the theory we have advanced, this new M cone was useful because it produced a new chromatic system, the L-M system, through the random contribution of M and L signals to a center-surround ganglion cell. With a high overlap in the tunings of the M and L cones, the effects of chromatic aberration were minimized, thus creating a chromatic system with higher spatial resolution. Perhaps the most novel and significant event in the evolution of the L-M system, however, was the New Trick. Via neural processes that as yet remain unknown, the visual system of Old World Primates can treat the input from the M and L photoreceptors interchangeably. There was thus everything to be gained and little to be lost by the addition of the M cone to Old World Primate vision. The floodgates were opened. With two different systems for contrast encoding, the signals from the L-M chromatic system and the L+M luminance system could be directly compared or combined over a wide range of spatial resolutions. A new partnership was born.²⁰

B. Why does the Colour System Occupy such a Large and Central Part in human Vision? Chromatic vision is a large part of the human system because it is both a parallel and complementary system to luminance vision. When chromatic information in the image is sparse, noisy, ambiguous, or absent, the visual system will use luminance information alone; when luminance information is sparse, noisy, ambiguous, or absent, the visual system will opt for chromatic processing. More commonly, though, our system will use both chromatic and luminance processes *in tandem* to solve the multiple, complex problems of human visual processing. Because the spatial resolution of the L-M system is quite good, output from the

¹⁸ While this sounds counter-intuitive, we you can see this for yourself by taking a colour photograph that depicts (unexpected) objects in the far distance, but which you can still see clearly in the colour photograph. Now render the photograph in black and white and ask someone else (who has not seen the photograph in colour) what is depicted. Often one's general sense of what the photograph contains will be compromised. This is because the chromatic system to distinguish the rough boundaries of the salient objects from shadow edges, and *that* information can then be brought into line with the sharper boundaries of the luminance image.

²⁰ In a recent experiment by Jacobs, G. H., G. A. Williams, et al. (2007). "Emergence of novel color vision in mice engineered to express a human cone photopigment." *Science* **315**(5819): 1723-1725. a third cone was added to the mouse retina. Recordings from single cells in visual cortex revealed that the outputs of the third cone had been integrated into retinal processing, producing a new chromatic response. Note that this, in itself, shows only that the random introduction and placement of the new will result in the inclusive wiring of this new cone into the normal organization of the retina. But the real question is whether the addition would turn out to be, overall, a benefit or an impediment or of no consequence. That is, is this the sort of genetic 'improvement' that would survive in the long run? Without a mechanism that allows us to use either M or L cone signals for the luminance processing without loss of acuity, trichromaticity in the mouse ought to have a cost —the loss of spatial resolution.

L-M system is more usefully integrated with luminance output—it can be used to directly to verify or falsify or ‘shore up’ the conclusions of the luminance system over a range of spatial resolutions. And because any integrated system of luminance/chromatic vision would require the *general* encoding of *both* intensity and wavelength information, chromatic processing has developed into an important component of Old World primate. This is why wavelength sensitive cells occupy about half of our visual ‘hardware’ at the level of the retina, LGN, and primary visual cortex.

The spectral system occupies pride of place in our visual system—the fovea—because evolution appears to have overcome several barriers to spatial resolution in a chromatic system. First, the large overlap between M and L receptor sensitivities decreased spatial blur in the image. Second, the S cones were excluded from the central 2 degrees of the fovea, where spatial resolution matters most, thus avoiding further chromatic aberration. And finally the New Trick preserved the spatial resolution of the luminance system. With these barriers removed, the human spectral system has proven apt for object vision, for vision that requires good spatial resolution.

C. Why Are the Blue Cones so Rare. On the traditional story, we are trichromats because it significantly increased our ability to discriminate between ‘near’ colours. The lack of S cones was attributed to the problem of chromatic aberration. But if the S cone input was used for colour alone, why would spatial resolution matter? To put this another way: if the advantage of trichromaticity is to see more colours, why has one third of the system been virtually sub-planted by the other two thirds?

As is now obvious, the answer lies in a re-interpretation of ‘colour’ function. There is one overwhelming reason for a diurnal mammal to have a second cone, namely to bring into alignment the mammal’s ability to respond to light with the light that is in fact most readily available within the environment. The advent of chromatic contrast information is another reason, but the utility of chromatic contrast is dependent upon the type of visual system within which it must be integrated. Old World Primates, it turns out, currently occupy the pinnacle of spatial resolution within a framework of general encoding. (Other species, such as the raptors, have higher visual acuity, but their visual systems are not general. High acuity information is used to control downward descent.). Thus the input from S cones, subject to high chromatic aberration, is an uneasy fit. In other words, S cones serve a necessary function—the widening of spectral range—but they cannot play a central role in Old World primate vision given its focus on high spatial resolution object vision. Instead, cones are sparse, limited to the fovea outside of the central two degrees, and their integration into cortical function is limited. Indeed, the output of the S cone system continues to serve the phylogenetically older visual functions. That is, the responses of Blue-Yellow cells are often highly specific, they have substantive projections to sub-cortical regions, presumably for motor guidance such as eye and head orientation (cf. (Leh, Ptito et al. 2010)). In other words, S cones are not ‘equal partners’ because the cortical visual function of Old World Primates moved on, to different pastures and, or perhaps we should say, to greener ones.

VII. CONCLUSION

Let us return to the two standard evolutionary questions posed at the outset of this paper, questions that follow from the Colour-for-Colouring view. The first question, the ‘Good For’ question asks about the utility of human colour vision: what is colour vision ‘good for’?

The second, narrower question concerned the colour categories: Why do we have precisely the (trichromatic) colour categories that we do?

At first glance, The Spectral Theory may not appear to answer the ‘Good For’ question insofar as this is a question about *colour*. We do see the colours, the reds, greens, and blues of human trichromatic vision. Yet the Spectral Theory has not given any explanation, either functional or evolutionary, of *that* specific visual capacity. For on this view, seeing surfaces, lights and transparent media as coloured *are* specific visual functions, much like, say, the perception of linear or trajectory motion. Like the perception of projective motion, colour vision proper is understood as one process (or perhaps many), that itself requires quite high level information about object properties—e.g. object shape, depth relations, direction of the illuminant and its colour, and inter-reflection from nearby objects. And this process (or multiple processes) is (are) almost certain to use information from all three channels of human vision, from the older B-Y, R-G, and luminance channels. In this article, we have outlined the Spectral Theory more generally, and have said nothing about the complex mechanisms that might underlie that capacity our capacity to see colours as properties of the distal world. Nor have we said anything about the functions that colour perception might subserve, given the ubiquitous presence of chromatic information in human visual processing writ large. Instead, the Spectral Theory recasts questions about the advantages of representing the colours in terms of the advantages of encoding spectral information. It asks: What are/were the selective advantages that chromatic systems conferred upon human vision throughout its long evolution? In answer to *this* question, there was no single answer. Rather, we have suggested that there will be as many answers to this question as there are distinct uses for chromatic information.

That said, the very multiplicity of answers to the ‘Good For’ question about chromatic processing in general undercuts many of the ‘single capacity’ theories of colour vision currently on offer. For example, we can agree with Sumner and Mollon (2000) that the red-green perceptual pop-out may have been a contributing factor to the evolution of human vision (assuming that our ancestors were largely frugivores). But we would disagree about the *mechanisms* underlying that capacity. Perhaps red/green pop-out stems from a L-M chromatic *contrast* mechanism for scene segmentation. One need not represent the reds and greens to see fruit among the leaves, to suddenly see the fruit as spectrally different or as salient. (That ripe strawberries are darker than their background foliage is a fact that strikes most people as counter-intuitive. We tend to think that pop-out and brightness must go together, not merely pop-out and chromatic contrast.) From the first person perspective, of course, we also see the fruit as red and the background foliage as green. But the explanation of pop-out need not be an explanation in terms of *colour*. Likewise, the abilities to differentiate between ripe and unripe fruits, tender and chewy foliage, robust and fallow complexions may not require the colours *per se*, just chromatic contrast processing. So ‘single use’ theories face a hurdle. They must show that the capacity to represent the colours—a far more complex capacity than chromatic contrast processing albeit one that is necessary for seeing the colours—was beneficial for some function, one that could not have been more easily performed by a chromatic and luminance contrast systems.

More generally, the Spectral Theory does not deny the importance of environmental colour to the evolution of cone systems. If every photopigment acts as a colour filter, and if those ‘colour filters’ are applied to the retinal image both individually and in combination, then what we can see is dependent upon the selection of appropriate colour filters (i.e. the appropriate photopigments). Woe to the photographer who misjudges the colour of the light

or the wavelength contrast between the photographic subject and its background before he or she applies a colour filter to the camera lens. Vision requires visual contrast and that in turn requires the correct filter relative to the colour of the subject and background. Undoubtedly, colour within the environment, both the statistical distribution of colour throughout the environment and the specific colours that are salient to human behavior, is a large factor in any stable configuration of cone sensitivities accrued over time. What the Spectral Theory questions is whether a visual system needs to represent colours *qua* the reds, blues and greens in order to make use of wavelength as a rich source of information. To this we answer our answer: Not in general.

The Spectral Theory has more direct consequences for our second question, at least depending upon the background suppositions of the reader. If you think that the human colour categories are *partly if not wholly* a result of the spectral tuning characteristics of the S, M and L cones, then the Spectral Theory provides an answer to the question, ‘why this precise organization of a trichromatic vision and not some other?’ In answer to *this* question we claim that there was no single behavioral or environmental fact that shaped the tuning of the three daylight receptors. The tuning of the cones, in their present configuration, depend upon the precise colour difference between ripe and unripe fruit, between tender and tough foliage. Instead our three cones evolved into a stable configuration that works well for human daylight vision in general—i.e. for an integrated system composed of two chromatic and (probably multiple) luminance channels, all of which are used for vision *writ large*. A shift in cone sensitivity might well affect the colours that we see. But the consequences of a new tuning will reverberate through the visual system as a whole, from the initial signal characteristics of luminance and chromatic ganglion cells, on through the cortical functioning of, e.g. motion, depth, and shape processing in all of their variations. It is these kinds of changes, the global effects on function brought about by a change in receptor tuning, against which success will be measured. On balance, does the new tuning promote or detract from human vision? As Sumner and Mollon (2000) say, it may be true that our M and L cones are optimally tuned for seeing red fruit against green leaves. But the M and L cones were not tuned—optimally or sub-optimally—for this goal, or indeed for any other specific visual task. This is a central lesson of the Spectral Theory.

REFERENCES

- Akins, K. (1996). "Of Sensory Systems and 'Aboutness' of Mental States." Journal of Philosophy **93**(7): 337-372.
- Akins, K. (Forthcoming). Black and White and Colour. Phenomenology and the Neurophilosophy of Consciousness R. Brown, Springer Verlag.
- Baylor, E. and F. A. A. K. Smith (1953). "The Orientation of Cladocera to Polarized Light." The American Naturalist **87**(833): 97-101.
- Changizi, M. A., Q. Zhang, et al. (2006). "Bare skin, blood and the evolution of primate colour vision." Biol Lett **2**(2): 217-221.
- Crook, J. D., M. B. Manookin, et al. (2011). "Horizontal cell feedback without cone type-selective inhibition mediates "red-green" color opponency in midget ganglion cells of the primate retina." Journal of Neuroscience **31**(5): 1762-1772.
- Dacey, D., O. S. Packer, et al. (2000). "Center surround receptive field structure of cone bipolar cells in primate retina." Vision Res **40**(14): 1801-1811.
- Dacey, D. M. (2000). "Parallel pathways for spectral coding in primate retina." Annu Rev Neurosci **23**: 743-775.
- Dacey, D. M., L. C. Diller, et al. (2000). "Physiology of L- and M-cone inputs to H1 horizontal cells in the primate retina." J Opt Soc Am A Opt Image Sci Vis **17**(3): 589-596.
- den Ouden, H. E. M., R. van Ee, et al. (2005). "Colour helps to solve the binocular matching problem." The Journal of Physiology **567**(Pt 2): 665-671.
- Dominy, N. J. and P. W. Lucas (2004). "Significance of color, calories, and climate to the visual ecology of catarrhines." Am J Primatol **62**(3): 189-207.
- Dow, B. M. and P. Gouras (1973). "Color and spatial specificity of single units in Rhesus monkey foveal striate cortex." J Neurophysiol **36**(1): 79-100.
- Gheorghiu, E. and F. A. Kingdom (2007). "Chromatic tuning of contour-shape mechanisms revealed through the shape-frequency and shape-amplitude after-effects." Vision Res **47**(14): 1935-1949.
- Gouras, P. (1974). "Opponent-colour cells in different layers of foveal striate cortex." J Physiol **238**(3): 583-602.
- Hansen, T. and K. R. Gegenfurtner (2009). "Independence of color and luminance edges in natural scenes." Vis. Neurosci. **26**(1): 35-49.
- Hardin, C. L. (1992). "Virtues of Ilusion." Philosophical Studies **68**(3): 371-382.
- Hart, N. S., H. J. Bailes, et al. (2008). "Visual ecology of the Australian lungfish (Neoceratodus forsteri)." BMC Ecol **8**: 21.
- Hiramatsu, C., A. D. Melin, et al. (2008). "Importance of Achromatic Contrast in Short-Range Fruit Foraging of Primates." PLoS ONE **3**(10): e3356.
- Hubel, D. H. and M. S. Livingstone (1987). "Segregation of form, color, and stereopsis in primate area 18." J Neurosci **7**(11): 3378-3415.
- Jacobs, G. H. (1993). "The distribution and nature of colour vision among the mammals." Biol Rev Camb Philos Soc **68**(3): 413-471.
- Jacobs, G. H. (2009). "Evolution of colour vision in mammals." Philos Trans R Soc Lond, B, Biol Sci **364**(1531): 2957-2967.
- Jacobs, G. H. and M. P. Rowe (2004). "Evolution of vertebrate colour vision." Clinical & experimental optometry : journal of the Australian Optometrical Association **87**(4-5): 206-216.
- Jacobs, G. H., G. A. Williams, et al. (2007). "Emergence of novel color vision in mice engineered to express a human cone photopigment." Science **315**(5819): 1723-1725.

- Johnson, E. N., M. J. Hawken, et al. (2001). "The spatial transformation of color in the primary visual cortex of the macaque monkey." Nat Neurosci **4**(4): 409-416.
- Kingdom, F. A., K. Wong, et al. (2006). "Colour contrast influences perceived shape in combined shading and texture patterns." Spat Vis **19**(2-4): 147-159.
- Kingdom, F. A. A. (2003). "Color brings relief to human vision." Nat Neurosci **6**(6): 641-644.
- Kingdom, F. A. A. and R. Kasrai (2006). "Colour unmasks dark targets in complex displays." Vision Res **46**(6-7): 814-822.
- Kingdom, F. A. A., K. Wong, et al. (2006). "Colour contrast influences perceived shape in combined shading and texture patterns." Spatial vision **19**(2-4): 147-159.
- Kiper, D. C., S. B. Fenstemaker, et al. (1997). "Chromatic properties of neurons in macaque area V2." Vis Neurosci **14**(6): 1061-1072.
- Leh, S. E., A. Ptito, et al. (2010). "Blindsight mediated by an S-cone-independent collicular pathway: an fMRI study in hemispherectomized subjects." Journal of cognitive neuroscience **22**(4): 670-682.
- Lennie, P., C. Trevarthen, et al. (1990). Parallel Processing of Visual Information Visual Perception. L. Spillmann and J. S. Werner. San Diego, Academic Press.
- Lettvin, J. Y., H. R. Maturana, et al. (1959). "What the frog's eye tells the frog's brain." Proceedings of the institute of radio engineers **47**: 1940-1951.
- Levine, J. S. and E. F. MacNichol, Jr. (1982). "Color vision in fishes." Scientific American **246**(2): 140-149.
- Livingstone, M. and D. Hubel (1988). "Segregation of form, color, movement, and depth: anatomy, physiology, and perception." Science **240**(4853): 740-749.
- Livingstone, M. S. and D. H. Hubel (1987). "Psychophysical evidence for separate channels for the perception of form, color, movement, and depth." J Neurosci **7**(11): 3416-3468.
- Lucas, P. W., N. J. Dominy, et al. (2003). "Evolution and function of routine trichromatic vision in primates." Evolution **57**(11): 2636-2643.
- Lythgoe, J. N. (1979). The Ecology of Vision. Oxford, Oxford University Press.
- Menzel, R. (1979). Spectral sensitivity and color vision in invertebrates. Handbook of Sensory Physiology, Vol VII/6A: Comparative Physiology and Evolution of Vision in Vertebrates. H. Autrum. Berlin, Springer Verlag. **VII/6A**: 503-580.
- Michna, M. L., T. Yoshizawa, et al. (2007). "S-cone contributions to linear and non-linear motion processing." Vision Research **47**(8): 1042-1054.
- Mollon, J. D. (1989). "'Tho' she kneel'd in that place where they grew..." The uses and origins of primate colour vision." J Exp Biol **146**: 21-38.
- Mullen, K. T., W. H. Beaudot, et al. (2000). "Contour integration in color vision: a common process for the blue-yellow, red-green and luminance mechanisms?" Vision Res **40**(6): 639-655.
- Mullen, K. T. and F. A. Kingdom (1996). "Losses in peripheral colour sensitivity predicted from "hit and miss" post-receptoral cone connections." Vision Research **36**(13): 1995-2000.
- Mullen, K. T. and F. A. A. Kingdom (2002). "Differential distributions of red-green and blue-yellow cone opponency across the visual field." Vis Neurosci **19**(1): 109-118.
- Munz, F. W. and F. N. McFarland (1975). Evolutionary adaptations of fishes to the photopic environment. Handbook of Sensory Physiology: The Visual System in Vertebrates. F. Crescitelli. Berlin, Springer Verlag. **VII/5**.
- Osorio, D. and M. Vorobyev (2008). "A review of the evolution of animal colour vision and visual communication signals." Vision Res **48**(20): 2042-2051.

- Regan, B. C., C. Julliot, et al. (1998). "Frugivory and colour vision in *Alouatta seniculus*, a trichromatic platyrrhine monkey." Vision Res **38**(21): 3321-3327.
- Regan, B. C., C. Julliot, et al. (2001). "Fruits, foliage and the evolution of primate colour vision." Philos Trans R Soc Lond B Biol Sci **356**(1407): 229-283.
- Riba-Hernandez, P., K. E. Stoner, et al. (2005). "Sugar concentration of fruits and their detection via color in the Central American spider monkey (*Ateles geoffroyi*). " Am J Primatol **67**(4): 411-423.
- Shimono, K., S. Shioiri, et al. (2009). "Psychophysical evidence for a purely binocular color system." Vision Research **49**(2): 202-210.
- Smith, A. C., H. M. Buchanan-Smith, et al. (2003). "The effect of colour vision status on the detection and selection of fruits by tamarins (*Saguinus* spp.)." J Exp Biol **206**(Pt 18): 3159-3165.
- Smith, F. E. and E. R. Z. G. Baylor (1953). "Color Responses in the Cladocera and Their Ecological Singificance." The American Naturalist **vol. 87**(No. 832): 49-55.
- Sumner, P. and J. D. Mollon (2000). "Chromaticity as a signal of ripeness in fruits taken by primates." J Exp Biol **203**(Pt 13): 1987-2000.
- Thorell, L. G., R. L. De Valois, et al. (1984). "Spatial mapping of monkey V1 cells with pure color and luminance stimuli." Vision Res **24**(7): 751-769.
- Vorobyev, M. (2003). "Coloured oil droplets enhance colour discrimination." Proc Biol Sci **270**(1521): 1255-1261.
- Vorobyev, M. (2004). "Ecology and evolution of primate colour vision." Clinical & experimental optometry : journal of the Australian Optometrical Association **87**(4-5): 230-238.
- Williams, D. R., N. Sekiguchi, et al. (1991). The cost of trichromaticity for spatial vision. Pigments in Perception. B. B. Lee and A. Valberg. New York, Plenum: 11-22.
- Zrenner, E., R. Nelson, et al. (1983). "Intracellular recordings from a biplexiform ganglion cell in macaque retina, stained with horseradish peroxidase." Brain Res **262**(2): 181-185.