

*This is a penultimate version of a paper accepted for publication in Philosophical Psychology. Please refer to the published version.*

## **The structure of color experience and the existence of surface colors**

Jan Degenaar & Erik Myin

Jan Degenaar, Faculty of Philosophy, University of Groningen, the Netherlands; Department of Philosophy, University of Antwerp, Belgium and Erik Myin (corresponding author), Department of Philosophy, University of Antwerp, Belgium, e-mail: erik.myin@ua.ac.be

Color experience is structured. Some ‘unique’ colors (red, green, yellow and blue) appear as ‘pure’, or containing no trace of any other color. Others can be considered as a mixture of these colors, or as ‘binary colors’. According to a widespread assumption, this unique/binary structure of color experience is to be explained in terms of neurophysiological structuring (e.g. by opponent processes) and it has no genuine explanatory basis in the physical stimulus. The argument from structure builds on these assumptions to argue that colors are not properties of surfaces and that color experiences are neural processes without environmental counterpart. We reconsider the argument both in terms of its logic and in the light of recent models in vision science which point at environment-involving patterns that may be at the basis of the unique/binary structure of color experience. We conclude that, in the light of internal and external problems which arise for it, the argument from structure should be rejected.

### **Keywords**

Argument from structure, color, explanation, opponent processes, structure of experience

### **1. Introduction**

Not all colors come equally. Cross-cultural studies of color naming showed that four ‘focal’ colors are widely singled out by color terms across languages (Berlin & Kay, 1969; Regier *et al.*, 2005): red, green, yellow and blue. There is a considerable psychophysical literature showing that normal perceivers can identify ‘unique’ shades of red, green, yellow and blue that appear not to be tinted by the other hues (Hurvich, 1981; Valberg, 2001). That is, red, green, yellow and blue lights can appear in a ‘pure’ or ‘unique’ form, that is as experienced without any trace of another hue. In contrast, ‘mixed’ or ‘binary’ hues, such as orange and indigo, appear as mixtures of two other (unique) hues. For example, magenta can be described as a slightly bluish red. In contrast, it does not seem natural to describe red as a purplish orange. Four focal colors correspond closely to the unique hues (Miyahara, 2003; Kuehni, 2005). Thus it seems natural to attribute a perceptually special status to particular shades of red, green, yellow, and blue, which we will call the ‘unique colors’.<sup>1</sup>

---

<sup>1</sup> For critique on the idea that there are four primitive unique hues, see Saunders and Van Brakel (1997). Our use of the term ‘unique/binary structure’ is only meant to capture the widely accepted (although not completely

The unique/binary structure of color experience has figured prominently in arguments to the effect that colors are not properties of surfaces, and that color experiences are neural processes lacking an environmental basis. We will focus on a particular argument, which we will call *the argument from structure*.<sup>2</sup> It has two premises (concerning structure) and two conclusions (concerning colors), and the argument for each conclusion can be said to form a strand. The first premise is this:

*Anti-realism about physical structure.* This is the idea that there is nothing analogous to the unique/binary structure to be found in the stimulus.

The idea here is that the unique/binary structure of colors has no genuine basis in the non-neural. The second premise is this:

*Realism about neural structure.* This is the idea that the structure of color is matched by a structure in the human visual system.

From the two premises two conclusions are then drawn. The first premise is used to argue for the first conclusion:

*Anti-realism about physical colors.* This amounts to the thesis that colors are not physical properties of the extracranial world.

From this conclusion, together with the second premise, a second conclusion is drawn:

*Realism about neural colors or color experiences.* The idea here is that colors or color experiences, which presumably lack a physical basis in the environment, are a property (Hardin, 1988/1993) or a construct (Palmer, 1999; Zeki, 1983) of the brain.

As Hardin puts it:

“so much about the colors – particularly the division between unique and binary hues – is manifestly bound up with the peculiar characteristics of the human visual system that it seems a vain pursuit indeed to search for a parallel set of structures in the general order of nature.” (Hardin, 1993, p. 67)

Building on a selection and interpretation of explanations given in the science of color vision, the argument from structure leads to the conclusions that “We are to be eliminativists with respect to color as a property of objects, but reductivists with respect to color experience” (Hardin, 1993, p. 112).

In this paper we critically examine this argument from structure, through scrutinizing both of its premises in relation to its two conclusions. In **section 2** we address the first strand of the argument, the claim that colors are not in the environment because the environmental stimulus lacks the required structure. We will address this strand as coupled to a requirement that any candidate environmental correlate of the relevant structural properties (unique versus

---

uncontroversial) special status of four focal colors. We certainly do not deny the existence of further focal colors that also have a somewhat special status, and we are not committed to the idea that binary colors are truly *composed of* the unique colors.

<sup>2</sup> ‘Arguments from structure’ can take somewhat different forms. An important feature of our analysis is that we show how ideas about the explanation of the structure of colors figure in the argument. We argue that assumptions regarding this explanation, although widely shared by proponents and opponents of the argument, are in fact mistaken.

binary) be explanatory with respect to the structure of experience. If this requirement is justified, it is not sufficient for refuting the argument from structure to point out that there is *some* possible description of the environment in terms of the unique/binary structure, for non-explanatory environmental properties are rejected. This explanatory requirement has not been met in recent defenses of color realism (Byrne & Hilbert, 2003; Churchland, 2007).

In **section 3** we examine the second strand in the argument from structure, in which it is concluded from the conclusion of the first strand, plus the idea that the unique/binary structure is to be explained fundamentally in terms of neural processes, that colors or color experiences inhere in the brain. We question both the evidence on which this reasoning builds, especially its interpretation of opponent processing models, as well as its logic, especially the problematic reliance on *a priori* excluding contingent explanatory possibilities. Alternative explanations have been unavailable for a long time. This may have raised the impression that both strands of the argument from structure have color science at their side.

Recent analyses, however, offer a different perspective. In **section 4**, we shall discuss the possible role of the environment in the explanation of the unique/binary structure of color experience. We shall focus in particular on an analysis of the stimulus as probed through our photoreceptors, which reveals systematic differences between unique and binary colors (Philipona & O'Regan, 2006). In this approach, environmental factors come to carry much of the explanatory weight regarding the structure in color experience that is carried by neurophysiological factors in opponent channel models.

**Section 5** discusses the consequences for the argument from structure as an argument for anti-realism about physical colors and neural reductionism about color experience. It shouldn't be taken for granted that key facts of vision science support the argument from structure. If the empirical approaches to which we draw attention are on the right track, the explanation of the structure of color essentially involves 'extradermal' factors. We claim that the conclusions of both strands of the argument become seriously discredited by these results. Contrary to the conclusion of the argument from structure, there very probably is both organismic and environmental reality to the structure of color.

## **2. The first strand: from lacking physical structure to anti-realism about physical colors**

Minimally, the unique/binary structure of color experience consists in the fact that some colors have a special place in human color vision, and that these colors are natural points of reference for describing color experiences. Stronger notions of the unique/binary distinction could invoke the idea that binary colors must be composed of unique colors, and that unique colors are not composed of any other colors. Such stronger notions may not be obvious to those familiar with mixing paint, especially in the case of green which after all lies between yellow and blue and can be seen as a mixture of these (for those familiar with mixing light, the 'uniqueness' of yellow in this respect might not be obvious). The argument from structure as we shall construe it need not appeal to such a strong notion of uniqueness. It needs only to refer to the less controversial fact that certain shades of red, green, yellow and blue appear to be special points of reference within color vision. The resulting structure of color experience is thought to be fundamental to our concept of colors (see and cf. Hardin 1993, p. 66).

According to the first premise of the argument from structure there is no physical structure in the environment analogous to the unique/binary structure in experience. From

this, it is concluded that there are no environmental colors. The absence of a physical structure intuitively comes down to the following. There are reflectances on the environmental side, and experiences on the experiential side. The issue is not that some (sets of) reflectances reliably cause types of experiences. Rather, the relevant point for the defender of the argument from structure is that the structure discernible in experience is not genuinely present in the environmental properties. In particular, there is no properly physical property of the particular reflectance patterns for unique colors, that distinguishes them from the reflectance patterns for binary colors.

According to the proponent of the argument from structure, no scrutinizing from a purely physical angle will deliver the required difference. Based on the idea that the unique/binary structure is fundamental to what colors are, an inference is drawn from the absence of appropriate environmental structure to the claim that there are no environmental colors. As Hardin puts it:

“If hues are physical complexes, those physical complexes must admit of a division into unique and binary complexes. No matter how gerrymandered the physical complex that is to be identical with the hues, it must have this fourfold structure, and (...) it must be possible to characterize that structure on the basis of physical predicates alone.”  
(Hardin, 1993, p. 66)

We think the argument from structure should be understood in an explanatory context. That is, we propose that the question of there being a ‘genuine basis’ or a ‘match’ for the structure of experience in the environmental realm is to be decided by whether or not reference to independent structure in the environmental realm carries explanatory weight regarding the structure of experience. ‘Independent’ here means that the environmental properties in the explanation are not defined in terms of experience itself, or in terms of higher level, in particular neural, properties. This is not to say that a genuine environmental structural basis cannot refer to any neural properties in its explanation of the structure of experience. It can do so, as long as the neural properties do not themselves explain the unique/binary structure of color experience – in section 4 we shall describe a model in which the environment plays such a role. Interestingly, the requirement of explanatory independence can be applied both to the physical (or environmental) as to the neural level. For a neural explanation of the structure of color experience is only independent of the environment if it stands apart from such structural properties as might be found in the environment. That is, if the neural structure itself can be explained in terms of the environmental structure, it loses its status as independent.

There are several reasons for interpreting the argument from structure in this explanatory based way. First, it provides a version of the argument that is harder to refute. Under the proposed interpretation, the first strand of the argument from structure, leading to the denial of the physicalist thesis that colors exist ‘out there’, cannot be simply refuted by pointing at some non-explanatory, dependent physical properties. At the same time, the explanatory requirement ensures that the second strand of the argument from structure yields a substantial claim, because the explanatory requirement applies to the neural level as well. Thus, by an explanatory reading, the argument against physical colors becomes more robust, and the assertion of the neural reality of color experiences becomes substantial. Second, this taking of the argument from structure is responsive to considerations given in favor of a ‘naturalization’ of ontology (Ladyman & Ross, 2007; Hurley, 2010, and of course Quine, 1960). For a question about existence of a genuine basis or match is decided by reference to

what plays a role in explanations. Third, the proposed interpretation seems to accommodate existing treatments of the argument from structure. For example, when Hardin writes that, in order to establish that hues are ‘physical complexes’, “it must be possible to characterize that structure on the basis of physical predicates alone” (Hardin, 1993, p. 66), this seems to fit very well with our requirement of independence for physical predicates.

The proposed interpretation of the argument from structure implies that two physicalist responses to the first strand of the argument from structure no longer can be mounted against it. To reiterate, in this strand, it is concluded that there are no physical colors, based on the premise that there is no physical property in the environment that corresponds to, or explains, the structure in experience. A first physicalist response is to concede that the unique/binary structure may be fundamental to our experience of colors, but to deny that this structure is fundamental to colors themselves. This opens up the possibility to accept the idea that the unique/binary structure lacks an environmental reality without accepting anti-realism regarding environmental colors (e.g. Churchland, 2007). If the environment does not contain the unique/binary structure, one might prefer to downplay the importance of this structure rather than the environmental reality of colors. This type of color physicalism is being put aside by the assumption that physical properties, in order to count as colors, must have a unique/binary structure which contributes to the explanation of the structure of experience.

A second physicalist way of answering is to accept that colors are structured into unique and binary hues, but to claim that a dependent physical structure is sufficient for the reality of surface colors. Byrne and Hilbert (2003), following this path, define the property of “hue quantity” associated with a certain reflectance profile (a related proposal can be found in Bradley & Tye, 2001). A hue quantity of a reflectance profile corresponds to how much of one of the hues red, green, blue and yellow is estimated to be present in the profile, when seen by some observers. Surfaces that are colored unique red are then the set of surfaces with reflectance profiles that have a maximal value on the red hue quantity dimension, and a minimal value on the other hue quantity dimensions.

The identification of hue quantities clearly derives from properties of observers. Hilbert and Byrne make clear that they think the relevant observer properties can be neural properties, related to the organization of the neurophysiology of color in terms of opponent channels (to which we shall turn in section 3.1). Thus they grant that the characterization of having a certain hue quantity may have recourse to exactly those high level neural properties of observers which are assumed to directly explain the structure of color experience in a neural model. They think that this does not imply that the property of being unique red itself is not physical. But such newfound object properties are not independently characterized, and therefore they remain explanatorily impotent. In Byrne and Hilbert’s hue quantities proposal it would be wrong to claim that the instantiation of hue quantities explains why human color experience has a unique/binary structure. That would turn matters on their head, as it is rather the case that the properties of experience explain the instantiation of hue quantities by surfaces.

The ‘hue quantity’ approach does grant that the explanatory burden for the binary structure of experience can be carried by neural structures, in particular opponent neurophysiology. As will become clear in the next section, we think that there are both logical and empirical reasons against relying on opponent neurophysiology to explain the unique/binary structure of experience. It is a live possibility that the environment *does* play a stronger explanatory role than is presumed by both anti-realists about physical color, like

Hardin, and realists about physical colors, like Churchland or Byrne and Hilbert. If an independent structure could be established, of course, this would directly refute the first strand in the argument from structure and open up the prospects of a realism about physical colors that does live up to the standards flowing from the explanatory form of the argument from structure. We will explore this niche of possibilities in section 4.

### **3. The second strand: from neural structure to realism about neural colors**

The second strand of the argument from structure takes two premises. The first one is the conclusion of the first strand, namely that environmental colors do not exist. From this, plus the additional premise that the unique/binary structure of colors is explained by the structure of neurophysiological processes, it is concluded that color experiences are neurophysiological properties without a physical basis in the environment.

Is the hypothesis forming the second premise of this strand justified? Can the unique/binary structure of color experience be explained by the peculiarities of our neurophysiology? Philosophers writing on the subject have expressed widespread confidence that *opponent processes* in the visual system can directly explain the structure of color experience (Hardin, 1988/1993; Clark, 1993; Churchland, 2005; 2007). However, as we point out below, the evidence base for this claim is less robust than these philosophers – unlike many color scientists – assume.

#### *3.1. The opponent processes hypothesis*

It is clear that the difference between unique and binary colors cannot be simply explained by the sensitivities of our photoreceptors: while there are presumably four unique colors, there are photoreceptors of three sensitivity types. Many vision scientists have therefore searched in ‘subsequent stages’ in the visual system for processes that could account for the unique/binary structure of color experience. The most popular neurophysiological account appeals to *opponent processes* in the brain (Hering, 1920; Hurvich, 1981; Hardin, 1993; Clark, 1993). The idea, derived from psychophysical studies, is that the visual system is organized in pairs of antagonistic physiological processes: one process corresponding to the black/white dimension; one for red and green; and one for yellow and blue. The processing of one color in an opponent process would then exclude the processing of the other, so that when the red/green opponent channel signals red, it cannot signal green. If the experience of unique colors correlates with the extremes of such antagonistic processes, these channels would be the neural process underlying the structure of color experience. While the experience of unique colors would involve activity in one color-sensitive channel only with the other channel at equilibrium, the experience of binary colors would have to result from the integration of signals from both channels. The opponent processes could then be interpreted as *color opponent* processes, which would be the physiological correlate of the unique/binary distinction (Hardin, 1988/1993; Clark, 1993; Churchland, 2005).

For example, in a philosophical discussion of color experience it has been suggested that the color opponent processes hypothesis generates predictions regarding negative after-images, and that these predictions are borne out (Churchland, 2005). The reasoning is as follows. Visual pathways are fatigued by elongated stimulation with the same color. As a

result, we become temporally less sensitive to this color. If the extremes of the opponent processes are signaling the unique colors, we may expect that a red/green channel that is fatigued with a red stimulus favors the signaling of green when a neutral stimulus is presented. Against a neutral background, the resulting after-image of unique red may then be expected to be unique green – assuming that the fatiguing takes place in the alleged color opponent stage in the visual system. But the trouble here is that for normal human perceivers the after-image of red is cyan and of green it is magenta (Wilson & Brocklebank, 1955; Pridmore, 2008). Cyan and magenta can indeed be classified as ‘green’ and ‘red’ respectively, but surely not as unique green and unique red. So after-images do not support the hypothesis that opponent processes match with the unique/binary structure of color experience.

Psychophysical experiments have uncontroversially supported the idea that there are chromatically opponent channels in two ‘cardinal directions’ (Krauskopf *et al.*, 1982), and in the extensive literature on color vision the existence of *physiological* opponent processes is widely accepted. The neural activity originating in the different photoreceptor types is recombined into opponent processes that inherit a wavelength-dependency from the photoreceptors at their basis (Valberg, 2001; Gegenfurtner, 2003; Wuerger *et al.*, 2005). For example, neurons with opponent properties have been found in the lateral geniculate nucleus in macaques, which have similar color vision as humans (Sandell *et al.*, 1979), (De Valois *et al.*, 1966; Derrington *et al.*, 1984). The activity of these neurons displays a wavelength-dependency combining the photoreceptor inputs into chromatically opponent processes.

Importantly, however, the extremes of the activation of these physiological opponent processes in the lateral geniculate nucleus do not generally coincide with the unique hues. For example, the yellow-blue dimension of color experience is not located at the equilibrium of any of these opponent processes (De Valois *et al.*, 1966; Derrington *et al.*, 1984). There is a clear mismatch between the physiological opponent processes in the lateral geniculate nucleus and the experience of the unique hues (Mollon & Jordan, 1997; Valberg, 2001; Wuerger *et al.*, 2005).

In response to this mismatch, some vision scientists have hypothesized a further stage in the visual system where neural activity does reflect the unique/binary structure of color experience (e.g. De Valois & De Valois, 1993; Wuerger *et al.*, 2005). While many neurons in the lateral geniculate nucleus are tuned in the two cardinal opponent directions, it has been found that neurons in visual areas in the cortex ‘differ widely in their chromatic preferences’ (Gegenfurtner, 2003). In the visual cortex of monkeys, for example, some cells respond to the wavelength of the stimuli, while the activity of other cells correlates more closely with surface colors as perceived by humans (Zeki, 1983). Instead of containing opponent channels, the macaque primary visual cortex may contain ‘hue maps’, in which perceptually more similar chromatic stimuli activate adjacent neural substrate (Xiao *et al.*, 2007). Also neuroimaging research of human visual cortex has revealed neural activity in visual cortex correlating with chromatic stimulation (e.g. Parkes *et al.*, 2009; Brouwer & Heeger, 2009). There are correlations between neural activity and various aspects of color, and these seem not particularly restricted to correlations to the unique colors.

While there is ample evidence of neural activity correlating to hues (Gegenfurtner, 2003), no opponent processes have been discovered to correlate precisely with the unique hues, and no special selectivity to the unique hues has been found. Thus the interpretation of opponent processes as neurophysiological correlates to the unique/binary distinction is certainly not to be taken for granted. According to present knowledge, the extremes of the

activation of known opponent channels do not match with the unique colors, and a well-supported neurophysiological account of the unique/binary distinction is presently unavailable.

### *3.2. Logical issues with neural correlations*

Suppose for the sake of argument that a neurophysiological correlate of the unique/binary difference will be found, within or without an opponent processing framework. Would that provide support for the hypothesis that the structure of color experience has a neurophysiological origin?

There is a fundamental difficulty with the proposed tight link between neurophysiology and the explanation of the structure of color experience. For if we find neural activity that mirrors the unique/binary structure of color experience, there remains the further question what gives rise to and therefore explains this structure. An unknown factor in the environment may structure neurophysiology, in which case the structure of colors may be thought of as deriving from the environment. As a result, neural correlates themselves have limited value for explaining the unique/binary structure of color experience as long as it remains obscure what structures the neural activity. If environmental factors, which are not themselves defined in terms of neurophysiology, are determining the structure of neurophysiology, then the explanation for the structure of color experience does not derive exclusively from the peculiarities of neurophysiology. The hypothesis that the structure of colors derives solely from the brain depends on excluding other potentially relevant factors.

As a result, it is hard to find sufficient empirical support for the second strand of the argument from structure. For a long time, a promising stimulus-involving account of the unique/binary structure of color experience has not been available. This might have reinforced the level of confidence in an exclusive neurophysiological account. But if an environment-involving explanation can be given this disqualifies the claim that the structure of colors derives exclusively from arbitrary peculiarities of neurophysiology.

## **4. An environmental contribution to the unique/binary structure**

Models in which the structure in the environment plays a crucial role in accounting for the unique hues can, for the purposes of this paper, be taken as being of two sorts. We shall speak of pure and mixed models. Though the two kinds of models are related in their orientation towards the environment, they differ in the degree in which they refer to properties of the visual system. Pure models do so only most minimally. Besides being concerned with *visible* light as it reflects off surfaces, such models don't rely on further properties such as the individual sensitivity spectra of the human cones, in their account of uniqueness. A mixed model more extensively refers to the human visual system, stopping short, however, of neural factors which themselves explain uniqueness. Such a model could refer, besides to environmental factors, to the specific sensitivities of the human photoreceptors. Importantly, both types of models meet the requirement of explanatory independence, as set out in section 2. This is obviously the case for pure models, as they relate uniqueness directly to environmental properties, such as surface reflectances. But it applies to mixed models as well, even if they rely in their explanation of uniqueness on the receptor sensitivities. The reason is



that the cone sensitivities do not independently explain the unique/binary structure of color experience. Indeed, the mismatch between the three photoreceptors and four unique hues is an important reason why theorists have proposed that the explanation for uniqueness lies at a post-receptor level.

The contours of a pure model can be found in Broackes (2011). Building on observations and suggestions of Shepard (1992), and Mollon (2006), Broackes points to the fact that variations in direct sunlight and skylight lie along a line between unique yellow and unique blue. Roughly, direct sunlight and skylight objectively fix unique yellow and unique blue and provide an easily available reference for calibrating color vision (Shepard 1992; Mollon 2006). Broackes indicates how an account of unique yellow and blue surfaces can be built from this. On this account, unique yellow surfaces are those that reflect most strongly under direct sunlight, and that gain most in luminance when the light changes from skylight to sunlight (Broackes, 2011, p. 622). Blue surfaces would then be those that, compared to other surfaces, are particularly strongly reflecting skylight, and that thus gain most in relative luminance when the light changes from sunlight to skylight.

Broackes suggests that these special characteristics of yellow and blue may form the basis, not just for these colors to form natural points of reference for color vision, but also for these colors to be experienced as particularly ‘pure’, and suitable primary elements for other colors. The idea here is that these colors, given that they match with and minimally darken generally available light, are likely to be taken as neutral. At the same time, these colors may be salient, or characterful, for yellow and blue objects are often more saturated than the sunlight or skylight. As Broackes puts it: “the colors of these illuminants give us *both* a kind of neutrality (in the relatively unsaturated cases) and untingedness and purity, *and* (in the more saturated cases of the same hue) a kind of *primary* characterfulness” (Broackes, 2011, p. 625). He proposes that the minimal way in which these colors darken the light under common conditions may give them a simplicity that makes them suitable elements from which other colors would seem to be composed. The account sketched by Broackes may then apply to a strong notion of uniqueness, according to which other colors appear composed of elementary unique colors. As he admits, more exact development of these ideas is desirable, and a lot more is needed to extend the account so as to encompass red and green.

A particularly interesting aspect of this account is that it is independent of the details of our photoreceptor sensitivities. For even with different cone sensitivities, blue surfaces will be sampled as reflecting more strongly under increasingly blue illumination (Broackes, 2011, p. 622). In such an account, the special status of yellow and blue would then come to the fore independently of the details of our sensory apparatus.

A mixed model is provided by Philipona and O’Regan, who do take into consideration the sensitivities of the human photoreceptor pigments (see Philipona & O’Regan, 2006, and Vazquez-Corral *et al.* for a further elaborated version). That is, Philipona and O’Regan studied how surfaces reflect light, but they were concerned specifically with how the light has effects on the three types of cones in the human eye. Thus, instead of the physicists’ ‘reflectance function’ of a certain surface, which shows what proportion of incoming light is reflected at any wavelength, they constructed a more biologically constrained reflectance measure, which takes into account only the light, both illuminating and reflected, as it affects the human photoreceptors.

Philipona and O’Regan plotted this biological reflectance measure for a wide variety of surfaces. They found that it was possible to characterize the biological reflectance function for

a particular surface by a 3x3 matrix. Such a matrix can be seen as specifying how a particular surface transforms any incoming light into a specific activation pattern in a receptor space. For most surfaces the receptor space is three-dimensional – which is what one would expect, given that there are three receptors whose responses can vary independently. Philipona and O'Regan came to the interesting observation, however, that some restricted classes of surfaces stood out, in that the activation patterns in receptor space were simpler for the lights they reflected. While the description of the illuminant requires three dimensions, some surfaces transform the light into a light that only needs one or two dimensions to be accurately described. In other words:

“certain surfaces have the mathematical property of being “singular”. What this means is that these surfaces take incoming light, which usually can vary in a 3-dimensional space defined by L, M and S cone activations, and transform it into light which varies only in either a 2- or in a 1-dimensional subspace of the LMS activation space. Because singular surfaces reduce variability from three dimensions to two or one dimension, they can be said to display a simpler behavior as concerns how they affect incoming light than the majority of surfaces.” (Vazques-Corral *et al.*, 2012)

The real interest of this finding was revealed when it turned out that the most strongly singular surfaces were quite exactly those surfaces picked out as ‘focal red’ and ‘focal yellow’ in research on focal colors, while ‘focal blue’ and ‘focal green’ corresponded to two other less strongly singular surfaces (see the plot on p. 335 of Philipona & O'Regan, 2006). Moreover, when surfaces with those focal hues were illuminated by the light source standardly taken as approximating daylight (D65), the light reflected off them showed accurate correspondence with light associated in psychophysical experiments with unique hues.

The Philipona and O'Regan approach shows how an explanation for the unique/binary structure of color experience is possible by appealing to nothing more than the patterns of sensory stimulation as available to the photoreceptors in the human eye. Strictly speaking, it is not necessary to presuppose opponent processes or other neurophysiological structures that match the structure of color experience. In the words of Philipona and O'Regan:

“these facts, not previously noted, have been obtained without appealing to neural mechanisms that underlie opponent channels in the visual system. They are merely a consequence of the asymmetries in surface reflecting properties induced by the limited way biological photoreceptors sample physical spectra. It could thus be argued that the reason the colors “red”, “yellow”, “green”, and “blue” are so often singled out among all colors as being worth giving a name, is that surfaces of these colors have the particularity that they alter incoming light in a simpler way than other surfaces (...).” (Philipona & O'Regan, 2006, p. 336)

It is tempting to make sense of these findings in a framework which emphasizes the temporal extendedness of color perception (Broackes, 1992; Noë, 2004; O'Regan, 2011). In this framework color perception is conceived of as becoming sensitive to color as a permanent property of the surfaces by means of sensitivity to the way the surface behaves over time under different illumination conditions. Such changes can be brought about by different means. Sometimes they might involve motion of the object, which may or may not be brought about by the perceiver. Alternatively, they might involve changes in the direct illumination, or in the indirect illumination, if surrounding objects move. Over time, one could become

sensitive to the underlying reflectance profile of a surface on the basis of minimal cues, so that one could detect the profile by just noticing the behavior of the surface in a minimal set of lighting conditions. Such a minimal set, moreover, is almost always available, as it exists when different nearby objects reflect differently on a surface (Ruppertsberg & Bloj, 2007).

In such a framework it is plausible that the singular surfaces of the Philipona and O'Regan analysis would be perceptually salient, for they would stand out as bringing about less change – and therefore as being simpler – whenever illumination conditions change. Such simple colors would be obvious points of reference for color vision and provide structure to color experience.

To sum up, the analysis of Philipona and O'Regan (2006) shows that there are systematic differences between focal red, green, yellow, and blue, compared to the other colors, and between the unique and binary hues. It can be derived from objective measures that surfaces with unique colors provide simpler patterns of sensory stimulation, compared to other colors. Rather than offering an *ad hoc* re-description of the stimulus, the analysis is based on independent evidence of surface properties and our retinal sensitivities, and it does not involve idiosyncrasies of human neurophysiology beyond the receptor sensitivities. The analysis strongly suggests that there is a crucial environmental contribution to the unique/binary distinction in color perception.

## **5. Consequences for the argument from structure**

According to an assumption that is widely shared in philosophical writings on color, the unique/binary structure of color experience is solely due to the structuring of experience by the peculiarities of neurophysiological processes, in particular the opponent channels. We have seen that, in contrast with this assumption, the explanation for the structure of color experience may crucially involve environmental factors. Below we discuss the consequences of this finding for the two strands of the argument from structure.

### *5.1. Consequences for the first strand*

The first strand of the argument from structure leads to the conclusion that strictly speaking physical colors do not exist. Does it follow from the alternative accounts considered that this conclusion is false, and that physical colors exist after all? It should be clear that only a pure model can establish what would count as a proof of the physical reality of surface colors according to the standards set by the argument from structure. For according to the argument from structure, such a proof would only be obtained if it would be established that there is some distal physical structure, specifiable completely independent of “the peculiar characteristics of the human visual system” (Hardin, 1993, p. 67), which is isomorphic to the structure in experience. A pure model comes as close to this as possible. Of course, even a pure account remains tacitly committed to a number of assumptions linked to human vision. For example, it is only concerned with *humanly visible* light, without even considering the realms of the ultraviolet or the infrared. But a demand for total independence is not reasonable, whenever the topic of theorizing is perception. Evidently, any perceivable property can only be perceived by a creature that is sensitive to it – so there will always be organismic dependence in an account of perception.

Therefore, on any reasonable assessment, it must be granted that, in providing properties of lights and surfaces that single out those that are seen as uniquely yellow and uniquely blue, in a way that is independent of but the most general characteristics of the visual system, pure models establish that, counter to the premise in the first strand of the argument from structure, there is an environmental basis for the unique/binary distinction.

A moment's reflection shows that not only pure, but also mixed models refute the premise of the first strand of the argument from structure. For, although mixed models appeal to more fine-grained aspects of sensory sensitivities, they are not thereby committed to a vicious kind of explanatory dependence as regards uniqueness. The characterization of the unique hues provided by Philipona and O'Regan takes into account how light affects our photopigments, but it will be difficult to resurrect in the light of it the idea that the explanation of uniqueness is not a matter of environmental physics but rather of neural processes. The reason is that the photopigment sensitivities do not themselves provide an explanation of uniqueness – unlike what color opponent physiological channels are hypothesized to do. The latter point is crucial in this context, because the exclusive appeal to neural structures such as color opponent channels is essential to the argument from structure. It is precisely the alleged neat mapping between the structure of neurophysiology and the structure of experience, together with the absence of such mapping between the structure of experience and the structure of the physical stimulus (if properly, independently characterized), that is supposed to lead to the conclusion that colors or color experiences reduce to neurophysiological properties. The appeal to the photoreceptor properties in the Philipona and O'Regan model does not assume such a direct mapping between these and neural properties.

Although the Philipona and O'Regan theory does not invalidate the opponent channel framework, it undeniably does shift the explanatory weight away from opponent channels and it provides a better explanation of the loci of the unique hues.<sup>3</sup> This is not to say that the Philipona and O'Regan approach is incompatible with post-receptor structuring. But in providing an alternative account of the loci of the unique hues, it does constrain the explanatory scope such structuring might have. If the model is correct, singling out a surface as unique or not unique does not necessarily require processes in neurophysiological opponent channels. For without taking into account any fact about opponent channels at all, it can be predicted by the model which surfaces will have a special status for perceivers like us. The Philipona and O'Regan approach thus allows for a characterization of the unique/binary distinction by means of a description which does not refer to color opponent channels.

## *5.2. Consequences for the second strand*

The second strand in the argument from structure requires that the explanation of the structure of experience be fundamentally or (in our terms) independently neural. It is quite clear that it is difficult to bring this strand to its conclusion if one of the models discussed in section 4 is true. For these models provide explanations of uniqueness which are in the required sense independently environment-involving. If true, they show that there *is* a genuine environmental

---

<sup>3</sup> The fact that it now appears that the structuring effects of opponent processes have been overestimated is fully consistent with the proposal that opponent processes optimize information transfer from the eye to the brain (Buchsbaum & Gottschalk, 1983).

explanatory basis for the structure of color experience. If so, the crucial premise that such environmental explanations do not exist should be rejected. The structure of color experience then does not provide support for the idea that color experiences are purely neural properties. In other words, the second strand of the argument from structure depends on the first strand. It is concluded that color experience belongs exclusively to the neural realm because of both the supposed absence of a genuine basis for structure in the environment and the supposed presence of such a genuine basis in the neural realm. Clearly, whenever there is an external structure, the conclusion of this line of argument no longer holds.

As we have seen, there are serious problems with the neural model that traditionally has been assumed to bring the explanatory goods, namely a physiological opponent process model. Still it remains possible that, even if pure or mixed models are correct regarding an environmental basis for uniqueness, some neural structures might be found which might have ‘internalized’ the structure of the environment in the organism’s sensory physiology.<sup>4</sup> It may then be argued that such neural structure does, after all, carry the explanatory load regarding color experience. By our own explanation bound criteria for ‘genuine existence’, this might seem to imply that color experiences are purely neural after all.

However, there are reasons to be skeptical about presupposing such a shift of the explanatory basis towards the brain. In the first place, it remains to be seen whether such ‘internalization’ of the unique/binary structure of hues is more than a possibility. Indeed, the more successful the pure or mixed environmental models for this structure, the less need there seems to be to reproduce the relevant structure within the brain. Powerful if controversial arguments have been given in favor of the ‘offloading’ of structure from the brain into the environment concerning other explananda in vision science such as visual stability or the filling-in of the blind spot (O’Regan 1992; 2011). Moreover, even if an internalization of structure were to be found, it would still itself have to be explained by the external structure. The latter is what is internalized. It comes first, in the temporal and in the explanatory order. This explanatory dependence remains even if, after the internalization has occurred, neural factors alone might suffice to bring about color experience (see also Hurley, 2010).

### *5.3. Conclusion*

The argument from structure should be rejected: the unique/binary structure of color experience does not support the claim that colors have no environmental basis at all, and neither does it legitimate the conclusion that color experiences only genuinely exist as neurophysiological phenomena. We highlighted analyses of structural aspects of color focused on the way the human eye is affected by the light reflected by surfaces, which indicate an objective explanatory basis for the unique/binary distinction without reference to peculiarities of neurophysiological processes beyond the basic sensitivities of our photoreceptors. From such analyses, it can be concluded that, while perceptual experience must of course depend on the perceiving organism, the explanation for the structure of color experience is at the same time crucially world-involving. The ontology of color and color experience should respect this wide explanatory basis.

---

<sup>4</sup> We thank a reviewer for raising this issue, and part of its formulation.

## Acknowledgements

We like to thank Kevin O'Regan, Bence Nanay, Victor Loughlin, and two anonymous reviewers for their excellent and very helpful remarks on earlier versions of this paper. For financial support we are grateful to the Research Foundation Flanders (project "Senses as Tools", G.0.321.09.N.10), the Research Council of the University of Antwerp (project "Imagery as Perceptual Activity"), and the Faculty of Philosophy of the University of Groningen.

## References

- Berlin, B.; Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley: University of California Press.
- Bradley, P.; Tye, M. (2001). Of colors, kestrels, caterpillars, and leaves. *The Journal of Philosophy*, 98(9), 469-487.
- Broackes, J. (1992). The autonomy of colour. In: Charles, D.; Lennon, K. (eds.) *Reduction, Explanation, and Realism*, 421-465. Oxford: Clarendon Press.
- Broackes, J. (2011). Where do the unique hues come from? *Review of Philosophy and Psychology* 2, 601-628.
- Brouwer, G.J.; Heeger, D.J. (2009). Decoding and reconstructing color from responses in human visual cortex. *The Journal of Neuroscience*, 29, 13992-14003.
- Buchsbaum, G.; Gottschalk, A. (1983). Trichromaticity, opponent colours coding and optimum colour information transmission in the retina. *Proc. R. Soc. Lond. B*, 220, 89-113.
- Byrne, A.; Hilbert, D.R. (2003). Color realism and color science. *Behavioral & Brain Sciences*, 26(1), 3-64.
- Churchland, P. (2005). Chimerical colors: some phenomenological predictions from cognitive neuroscience. *Philosophical Psychology*, 18(5), 527-560.
- Churchland, P. (2007). On the reality (and diversity) of objective colors. *Philosophy of Science*, 74(2), 119-149.
- Clark, A. (1993). *Sensory Qualities*. Oxford: Oxford University Press.
- De Valois, R.L.; Abramov, I.; Jacobs, G.H. (1966). Analysis of response patterns of LGN cells. *Journal of the Optical Society of America* 56 (7), 966-977.
- De Valois, R.L.; De Valois, K.K. (1993). A multi-stage color model. *Vision Research*, 33(8), 1053-1065.
- Derrington, A.M.; Krauskopf, J.; Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology*, 357, 241-265.
- Gegenfurtner, K.R. (2003). Cortical mechanisms of colour vision. *Nature Reviews Neuroscience* 4, 563-572.
- Hardin, C.L. (1988/1993). *Color for Philosophers: Unweaving the Rainbow*. Second edition (1993). Indianapolis: Hackett Publishing Company.
- Hering, E. (1920/1964). *Outlines of a Theory of the Light Sense*. Cambridge, MA: Harvard University Press.
- Hurley, S. (2010). Varieties of externalism. In: Menary, R. (ed.) *The Extended Mind*, 101-153. Cambridge, MA: MIT Press.
- Hurvich, L.M. (1981). *Color Vision*. Sunderland: Sinauer Associates Inc.
- Krauskopf, J.; Williams, D.R.; Heeley, D.W. (1982). Cardinal directions of color space. *Vision Research*, 22, 1123-1131.
- Kuehni, R.G. (2005). Focal color variability and unique hue stimulus variability. *Journal of Cognition and Culture*, 5, 409-426.

- Ladyman, J.; Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford: Oxford University Press.
- Miyahara, E. (2003). Focal colors and unique hues. *Perceptual and Motor Skills*, 97, 1038-1042.
- Mollon, J. (2006) Monge: The Verriest lecture, Lyon, July 2005. *Visual Neuroscience*, 23, 297-309.
- Mollon, J.D.; Jordan, G. (1997). On the nature of unique hues. In Dickinson, C., Murray, I., Carden, D. (eds.) *John Dalton's Colour Vision Legacy*, 381-392. Taylor & Francis.
- Noë, A. (2004). *Action in Perception*. Cambridge, MA: MIT Press.
- O'Regan, J.K. (1992). Solving the "real" mysteries of visual perception: the world as an outside memory. *Canadian Journal of Psychology*, 46(3), 461-488.
- O'Regan, J.K. (2011). *Why Red Doesn't Sound Like a Bell: Understanding the Feel of Consciousness* (New York: Oxford University Press).
- Palmer, Stephen E. (1999). *Vision Science: Photons to Phenomenology*. Cambridge, MA: MIT Press.
- Parkes, L.M.; Marsman, J.B.C.; Oxley, D.C.; Goulermas, J.Y.; Wuerger, S.M. (2009) Multivoxel fMRI analysis of color tuning in human primary visual cortex. *Journal of Vision*, 9 (1), 1-13.
- Philipona, D.L.; O'Regan, J.K. (2006). Color naming, unique hues, and hue cancellation predicted from singularities in reflection properties. *Visual Neuroscience*, 23, 331-339.
- Pridmore, R.W. (2008). Chromatic induction. *Color Research and Application*, 1, 77-80.
- Quine, W.V.O. (1960) *Word and Object*. Cambridge, MA: MIT Press.
- Regier, T.; Kay, P.; Cook, R.S. (2005). Focal colors are universal after all. *Proceedings of the National Academy of Sciences of the United States of America*, 23, 8386-8391.
- Ruppertsberg, A.I.; Bloj, M. (2007). Reflecting on a room of one reflectance. *Journal of Vision*, 7(13): 12, 1-13.
- Sandell, J.H.; Gross, C.G.; Bornstein, M.H. (1979). Color categories in Macaques. *Journal of Comparative Physiological Psychology*, 93(4), 626-635.
- Saunders, B.A.C.; Van Brakel, J. (1997). Are there nontrivial constraints on colour categorization? *Behavioral and Brain Sciences*, 20, 167-228.
- Sheppard, R.N. (1992). The perceptual organization of colors: an adaptation to regularities of the terrestrial world? In: Barkow, J.H.; Cosmides, L.; Tooby, J. (eds.) *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, 495-532. New York: Oxford University Press.
- Valberg, A. (2001). Unique hues: an old problem for a new generation. *Vision Research*, 13, 1645-1658.
- Vazquez-Corral, J.; O'Regan, J.K.; Vanrell, M.; Finlayson, G.D. (2012). A new spectrally sharpened sensor basis to predict color naming, unique hues, and hue cancellation. *Journal of Vision* 12 (6), 1-14.
- Wilson, M.H.; Brocklebank, R.W. (1955). Complementary hues of after-images. *Journal of the Optical Society of America* 45 (4), 293-299.
- Wuerger, S.M.; Atkinson, P.; Cropper, S. (2005). The cone inputs to the unique-hue mechanisms. *Vision Research*, 45, 3210-3223.
- Zeki, S. (1983). Colour coding in the cerebral cortex: the reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, 9, 741-765.