# Revelation and normativity in visual experience

(Zoltán Jakab, NSERC postdoctoral fellow, Rutgers University)

## 1. Revelation in color perception

One variant of philosophical realism has it that our visual experience reveals to us the very essence of color. In order to know immediately, and exactly, what properties the colors are, all one needs is to see them. Colors are out there, pervading objects' surfaces and filling transparent volumes. Want to know what property redness is? Look at a drop of blood, a glass of red wine, and perhaps a few more red things. That will acquaint you, immediately and exhaustively, with the essence of redness. So much so that you can't even expect any other way of grasping the essence of redness (*mutatis mutandis* for other colors). (Johnston, 1997, p. 138; Campbell, 1997, pp. 178-179; McLaughlin, 2003, p. 97; Russell, 1912, p. 47; Strawson, 1989, p. 224; Stroud, 2000). It is of course possible to learn empirical facts about color via scientific inquiry, but such facts alone will never teach one the essential nature of color, nor is the knowledge of such facts needed to grasp the essential nature of color via perception. This seems to imply that whatever color science might teach us is for the most part irrelevant (McLaughlin, 2002, pp. 97-98). To summarize: color vision is (1) *revelatory* with respect to color, and (2) nothing else is revelatory – *nothing else could possibly be revelatory* – about color.

The first part of this notion, (1), derives from a common-sense intuition that some philosophers wish to preserve in their theory of color. Very roughly, the intuition is that colors are the way they look to us to be, or that *what it is like to see the colors* derives straightforwardly from what the colors themselves are like (Tye, 2000, chapters 3 and 7). The second part, (2), probably derives form the observation that whatever color vision reveals to us cannot be communicated in language, simply because what it is like to see the colors cannot be communicated. Either one can see the colors or one will, seemingly, never learn what it is like to see them. No description will help the unacquainted (Nordby, 1990, p. 305; Raffman, 1995; Tye, 1995, pp. 169, 172-174; 2000, Chs. 1-2; Jakab, 2000). So if it is indeed the essence of color that color vision reveals, then it appears that nothing other than color experience – not at least language – can reveal to us the essence of color.[1]

So we have one notion of revelation: we might call it *primitive revelation*.[2] I think primitive revelation is a profoundly mistaken idea. However, the goal of this paper is not to argue against primitive revelation. As for my concerns about this notion, what makes it sound implausible is that it implies a sort of "agnosticism" about color, namely that no matter what concepts we might come to have, none of them will ever give us any understanding of color. Experience is supposed to remain the only teacher in this respect (McLaughlin, 2002, pp. 97-98). In addition, primitive revelation seems to imply that colors are not the canonical causes of our experiences of color[3]. For color science surely can describe the physical properties that are the canonical causes of color experience. But color science also makes it clear that color perception does not reveal the essence of those properties in any obvious way.[4]

The reason I began this paper by introducing this notion of revelation is because I am going to offer an alternative notion of perceptual revelation, one that is arguably true of some perceptual modalities – though not of color perception. My alternative notion, I claim, is consistent with our knowledge about the psychology of perception, and I hope it will also help to clarify some issues about perceptual representation.

The outline of the present paper is as follows. Section 2 draws the reader's attention to certain differences between the perception of shape and that of color, and introduces an alternative notion of revelation. Section 3 supports the claims of Section 2 by some psychological data and reasoning. Section 4 outlines the two main brands of physicalism about color, and the difference between them that is most important in this context. Section 5 argues against one of them, the so-called absolutist verison of color physicalism, and for the other, relativist approach. Section 6 considers some defending lines for color absolutism. Section 7 formulates the key claims of this paper, namely how the notion of revelation proposed in Section 2 makes color relativism plausible while preserving the idea that relativism about shape would be absurd. Finally, Section 8 critically discusses Martin Davies's paper called 'Externalism and Experience' (Davies, 1997) and argues that representational externalist accounts of visual experience are not very well supported – indeed, I think they are mistaken.

Here are some termionological remarks. I shall use the terms 'color', 'color category', 'narrow shade', 'determinate', 'determinable' to designate object-color (i.e., stimulus) properties (for the latter two terms see Byrne and Hilbert, 1997, pp. 266-267, 276-278, 280-281). To refer to the relevant internal perceptual states I shall use the terms 'perceptual color category', 'perceived narrow shade', 'color percept', 'unique hue experience', and 'binary hue experience'. Of these, 'perceived narrow shade' and 'color percept' mean the same. Color percepts or perceived narrow shades are perceptually determinate color experiences, perceptions that we get when looking at particular colored surfaces on particular occasions. For instance, we never perceive a surface that is *red, full stop*: we always perceive some determinate shade of red. Unique hue experiences and binary hue experiences are different types of color percepts. Narrow shades, on the other hand, are specific color stimuli corresponding to color percepts. Narrow shades do not have discriminable sub-shades – they are probably best characterized in terms of metamer sets (see Finlayson and Morovic, 2000a, 2000b). A metamer set that is a narrow shade is such that any two subsets of it would look indistinguishable in color.[5] In Byrne and Hilbert's terms, narrow shades are determinates, whereas broader or narrower color categories are determinables, under which many different narrow shades or determinates belong.[6]

## 2. Perception of color and perception of shape: a difference

We visually perceive shapes and colors. It is reasonable to distinguish shapes from what it is like to see them, because shapes are perceiver-independent: they are physically realized in the absence of perceivers. What it is like to see the shapes has to do with what information the visual system picks up about three-dimensional scenes, and how it processes that information. According to physicalism about color, colors are also perceiver-independent – just like shapes, colors are physical properties that are realized in the absence of perceivers. On the other hand, what it is like to see the colors has to do with how our vision accesses and processes information about surfaces and illuminants. Thus it is also reasonable to distinguish between colors and what it is like to see them.

In addition to this, there is no doubt that shapes are the canonical causes of our shape percepts (visual or tactile). Analogously, color physicalists maintain (quite rightly, in my opinion), that colors are the canonical causes of color percepts. That is, in anything like "broadly normal" or "optimal" conditions of perception – in circumstances where humans and animals with color and shape perception evolved and have been living – it is shapes that evoke our shape percepts, and it is colors that evoke our color percepts.

However, here is a difference between the perception of shape and color. Philosophers have been obsessed with the following questions: what it color in objects? Are external objects colored at all? If yes, are colors the canonical causes of our perceptions of color? A number of different answers have been proposed to these questions. Some philosophers think objects are colored, whereas others hold that color perception is a pervasive visual illusion (Hilbert, 1987; Jackson and Pargetter, 1997; Byrne and Hilbert, 1997, 2003; Tye, 2000; Ch. 7; Boghossian and Velleman, 1997a, 1997b; Hardin, 1988). Some think colors are physical properties, while others disagree (e.g., Broad, 1923; Johnston, 1997). There is no general agreement about whether colors are the canonical causes of color experience (Tye, 2000, pp. 148-149; Shoemaker, 1994; Campbell, 1997; Broad, 1923; Hardin, 1988).

There is no parallel problem for the notion of shape and shape perception. When one asks: 'What are the shapes of objects?' we can reply: shapes are types of spatial distribution of matter. We also have abstract shape concepts designating these types. We can readily describe shapes: regular ones by the well-known shape concepts of Euclidean (or some other) geometry, irregular ones by the notion of coordinate systems and lists of pairs (n-tuples) of numbers characterizing points in coordinate systems.[7]

The disanalogy continues. Our abstract concept of shapes in general, and an exact characterization of many particular shapes, arose from visually perceiving (and manipulating, e.g., drawing) shapes, plus intellectual reflection. Ancient Greeks did not have empirical science, nor did they need it for coming up with Euclidean geometry.

After that it took a long time until others formulated non-Euclidean geometries, which made assumptions that contrast with perception-driven intuition (e.g., changing Euclid's fifth axiom). Apparently, visual perception happens to be the most powerful source of our concepts of shape. We can of course acquire abstract concepts of shape without having vision, as do congenitally blind people. Even though vision is a very powerful, and the most typical, means of learning about shape, it does not seem necessary that we learn shape concepts through vision.

The corresponding story about color is entirely different. Color perception plus intellectual reflection alone have never given us a non-controversial notion of object color. Indeed, philosophical reflections on our perception of color sparked the very debate about color that I just mentioned. That is still an ongoing debate. On the other hand, empirical science (color science and psychophysics) taught us a whole lot *about the canonical causes of color experience* – surface reflectances, color signals, spectral sensitivities of the photoreceptors, and the like. Science taught us things that color perception and intellectual reflection, not conjoined with empirical methods, were unable to teach us. In particular, color perception and reflection alone never gave us ideas like color is the same as, or is intimately related to, surface reflectance, relative energy distribution of emitted light, and so on.

Let me elaborate on this claim a bit. Even ancient Greeks were aware of the fact that surfaces modify the incoming light and that that process is largely responsible for color. Medieval thinkers knew this fact too (????**sources in ZemplenValasz.txt, and Maloney, 1999's motto**). The idea that white light is complex, colored lights are components of white light, and that the color of surfaces depends on what component of the incoming light they reflect was apparently not known to Greeks. It is often attributed to Newton; however it can also be found in Albertus Magnus' writings in the 13th century (Gage, 1990, p. 140), and Newton was likely aware of that. In sum, it is reasonable to assume that color perception together with intellectual reflection can give us a rough and ready idea of what *colors in general are*. But this idea is not very precise and, most importantly, it does not extend to any (let alone exact) characterization of particular colors. In contrast, shape perception plus intellectual reflection can lead us to exact and largely uncontroversial characterization of particular shapes.

Now if one agrees that colors are the canonical causes of color experience, then the conclusion is that empirical science taught us a lot of new facts *about the nature of color*. If one denies that colors are the canonical causes of color experience, then one can insist that color science did not teach us the least bit about color as such. In this case the question arises what can possibly teach us about the nature of color?[8] Different answers are possible to this question. One is that philosophical reflection on color perception can give us an understanding of the essence of color (including that of particular colors), independently of whatever empirical science might find. Another answer might be that the essential nature of color is accessible only to color perception, and no abstract concept that we might form about color could ever capture that essence (this is primitive revelation). Third, it is also possible to hold that nothing – neither abstract concepts, nor color perception – could ever reveal the nature of color. (Though I'm not sure if anyone holds such a view about color.) A fourth answer is to hold that colors in objects do not exist, and so color perception is a pervasive illusion. In this case psychology might be able to tell us a lot about the nature of color by offering an understanding of how such a pervasive illusion can arise.

In what follows I shall assume that colors *are* the canonical causes of color experience. I shall also assume that colors exist and that they are physical properties of environmental surfaces. For support of these assumptions, I rely on the literature that defends physicalism about color (especially McLaughlin, 2003a, 2003b; Jackson and Pargetter, 1997, 2000, but see also Hilbert, 1987; Byrne and Hilbert, 1997, 2003; Tye, 2000, Ch. 7). With these assumptions at hand I can reach the conclusion that empirical science taught us facts about the nature of color (and especially that of particular colors) that color perception plus intellectual reflection were unable to teach us.

On the contrary, as I argued, shape perception plus intellectual reflection were able to teach us the essence of shapes in general, and also the essence of particular shapes. No empirical inquiry was necessary for this achievement. Now consider the following principle:

**[Conceptual revelation]** Stimulus property P is revealed in perception iff perceiving P plus intellectual reflection together can lead us to a conception of P's essence in perception-independent terms, that is, a conception that does not make reference to our perception, or experience, of P.[9]

As I argued above, conceptual revelation is true of shape perception. On the other hand, it is not obvious, to put it mildly, that it is true of color perception. If what color science can teach us includes the essential nature of color, then the essential nature of color is not accessible to just color perception and intellectual reflection – because what color science teaches us (about wavelength distributions, surface reflectances, and so on) is not so accessible. If, on the other hand, it is held that color science cannot teach us the essence of color, then it is possible to argue that conceptual revelation is true of color perception. But the best argument in favor of such a claim would be to come forward with a relatively uncontroversial philosophical theory of color that includes an exact characterization of particular colors in perceiver-independent terms[10] (like Euclidean geometry characterizes shapes) and that does not rely on the findings of color science. Anyone who knows the philosophical literature on color knows that there is no theory up to date that comes anywhere close to this requirement.[11]

At this point, the following question arises: what psychological difference between shape perception and color perception is responsible for the fact that shape perception conceptually reveals its own canonical causes (which are no doubt the shapes) whereas color perception does not conceptually reveal its own canonical causes (which, I assume, are the colors)?

**3. Reason for the difference**

As far as we can tell, visual representations of particular shapes are compositional. At higher levels of visual processing, representations of complex shapes are built up out of shape primitives. Marr (1982) suggested that the shape primitives might be something like cylinders of different height and diameter.[12] More recently, Irving Biederman (1990) proposed a more sophisticated set of shape primitives that he calls *geons*, out of which representations of complex shapes are constructed. The number of shape primitives is small (approximate number????), but if they are properly chosen, then combinations of them can approximate a large number of different shapes that might arise in our environment. Biederman proposes that the geons are representations of simple regular shapes (e.g., cuboid, cylinder, cone, prism, etc.) – shapes that are easy to distinguish from one another from almost any viewing perspective. Another criterion is that the geons remain extractable from the lower-level visual information after partial deletion of that information (and so the corresponding real shapes are recognizable when partially occluded in different ways). Marr's and Biederman's models of shape perception focus on shapes that are of special interest to humans (and other vertebrates): animals, plants, familiar artificial objects, and other shapes that exhibit some regularity, or symmetry.[13] The shape representations that we store in our memory are thought to be aggregations of the geons. According to Biederman, shape recognition consists of an approximation of the lower-level visual information about a seen object by some pattern of geons, then matching this pattern against memory representations of previously seen objects (Sekuler and Blake, 2002, p. 224). In effect, shape primitives are thought to constitute a generative system of representation in which the rules of combination are fairly simple and flexible, somewhat similarly to the combination of Lego blocks.

At lower levels of visual processing the principles of organizing visual information are different, but other forms of compositionality are very likely present at these lower level visual representations: for instance, representations of surfaces are constructed out of information about edges and that about texture (Marr, 1982; Stillings et al., 1995, pp. 464-490). There is reason to believe that the output of processing at these lower levels sometimes reaches consciousness. One such example is Julesz's random dot stereograms that demonstrate that perceptions of shape and depth can arise from binocular disparity data alone (Julesz, 1971). However, it is mostly the higher level visual processing the output of which reaches consciousness, or becomes accessible to the rest of the cognitive system (like reasoning or action planning).

It is worth noting that we can also discern the details of two-dimensional shapes and drawings – proportions of lengths of lines and edges, numbers and types of angles, and so on – suggesting that such shapes are also compositionally represented. This might be a special case to which Biederman's account can be applied, though that account is primarily concerned with the recognition of three-dimensional objects.

For my purposes it is not highly relevant whether Biederman's account of shape perception is exactly correct or not. At any rate, it is currently one of the top contenders, and so it can illustrate my general point that the visual representation of particular shapes is compositional. Due to this compositionality, shape perception delivers abundant information about particular shapes.

Now compare color perception. The information accessed by trichromatic color vision about color consists of the outputs of three wavelength-selective cone types that are sensitive to three different and overlapping broad ranges of wavelengths in the visible spectrum. The cone types themselves exhibit a characteristic sensitivity profile within their sensitivity range: sensitivity is highest in the middle of the sensitivity range and gradually decreases toward the two ends. This differential sensitivity introduces a filtering of the incoming light at the entry level of visual processing, namely absorbing light at the retina. This filtering is followed by a number of subsequent steps of processing, the most well-known of which is opponent recombination (Hardin, 1988, pp.

~36????; Wandell, 1995, Chs. 4, 9; Maloney, 1999, pp. 409-413; DeValois and DeValois, 1997). The output of all these stages of visual processing is our phenomenal experience of color. Even though particular colors are represented, at low levels of processing, by differential wavelength composition of the incoming light (e.g., reds correspond to a dominance of light at the long end of the spectrum accompanied by little light in the 400-580 nm range; greens correspond to light predominantly in the middle of the visible spectrum and little at the two ends, etc.), no such information is discernible from our conscious experience of color. The output representation of color processing, that is available to the rest of the cognitive system, is some position in a three-dimensional perceptual similarity space called color space. Each perceived color corresponds to a point in color space. There are different representations in color science of color space, but from each of these three dimensions of perceived color can be derived (directly or indirectly): hue, lightness, and saturation. There is another aspect of perceived color that is not well expressed by the dimensions of hue, lightness and saturation, and so alternative representations of color space are used to contain it explicitly. This is the unique-binary division and opponent organization of colors as we perceive them.[14] There are four chromatic colors that do not look mixed of two other colors: unique red, green, yellow, and blue. The rest of colors look mixed, in different proportions, of two other chromatic constituents. Oranges look reddish and yellowish; purples look bluish and reddish; lime looks yellowish and greenish, and chartreuse, bluish and greenish. Opponent organization means that red does not perceptually mix with green, and yellow does not perceptually mix with blue – no surface or volume ever looks reddish green or bluish yellow.[15]

By this brief exposition of color processing I meant to illustrate the point that the representation of particular colors in color vision is not compositional, or only very minimally so. In particular, unique hue experiences are chromatically non-compositional, whereas binary hue experiences are minimally compositional: they, as representational states, have two chromatic constituents (that can be activated at different levels, and in different proportions). Unique hue experiences, as perceptual representations, consist of a single chromatic constituent (or process). Binary hue experiences consist of two chromatic constituents.[16] At another place (Jakab, 2000) I supported this point with more detailed reasoning. The argument there was, roughly, that something is a constituent in a complex perceptual experience if and only if (1) it is *discernible* in the complex experience (i.e., discernible to the rest of the cognitive system; discernible for the cognitive processes that access, and operate on, perceptual representations), and (2) it can be undergone on its own, separately from "the rest"; it can occur as an experience itself. Orange surfaces look reddish and yellowish, because perceived reddishness and perceived yellowishness are discernible constituents in the experience of orange, and they can be undergone as experiences of their own: those of red and yellow. The observation that orange does not appear both red and yellow (i.e., many oranges appear neither red nor yellow), does not threaten the idea that the experience of red and that of yellow are constituents of experiences of orange (see Byrne and Hilbert, 1997, p. 280; see also Thompson, 2000, pp. 169-173).[17] Vodka is neither water, nor alcohol, yet it has two key constituents: water and alcohol. Vodka is pretty much waterized alcohol, or alcoholized water; similarly, at the level of perceptual experience, orange is *red "diluted" by yellow*, or *yellow "contaminated" by red* (Jakab, 2000, p. 339). This is how orange – the object color – appears both reddish and yellowish (see Byrne and Hilbert, 1997, pp. 280-281). On the other hand, looking only yellowish (chromatically), or predominantly yellowish, is the same as looking yellow. The compositionality in the experience of orange should be understood as a combination of representational states. The experience of orange is not the experience of red *and* the experience of yellow (it is neither); it is the concatenation of these two experiences. Similarly, the complex symbol A&B is neither the symbol A nor the symbol B – it is the concatenation of these two symbols.

It is also arguable that whatever simple syntactic, or constituent, structure color experiences might have is representationally irrelevant in the sense that it does not successfully, or veridically, represent a corresponding physical structure present in the canonical causes of color experience. The unique-binary division is a perfect example of this (Hardin, 1988, pp. 66-67; 1997, p. 291; Thompson, 1995, pp. 123-124, 2000, p. 169). There are theorists who contend that the unique-binary division in color experience does represent a true physical division in the realm of object color (Byrne and Hilbert, 1997, pp. 280-281, 2003; Tye, 2000, pp. 162-165; Bradley and Tye, 2001). In my opinion, this is less than obvious (see Hardin, 1988, pp. 66-67; Clark, 1996, p. S145; Matthen, 1999; Thompson, 2000, pp. 172-173; Jakab, 2003a, 2003b; Jakab and McLaughlin, 2003). For this reason I think that singular color experiences are *representationally atomic* (Fodor, 1987, 1998, pp. ????; Jakab, 2000, pp. 337-338). Their syntactic structure is minimal and "residual": it does not reliably indicate any corresponding structure in the environment. In addition, some color experiences (like those corresponding to the Hering primaries) may not have any proper constituents (constituents other than themselves) at all, and so they are *elementary perceptual states* (Jakab, 2000, p. 337).[18]

Note that a better-known analysis of binary hue experiences rejects the idea that such experiences have constituent structure (Hardin, 1988, p. 43; Thompson, 2000, p. 171). Hardin says that particular colors are represented by vectors in a three-dimensional vector space (i.e., color space), and that such vectorial representations of colors have *vector components*, and vector components are not parts of vectors. Hence, on this analysis, no color experience has any constituent structure in the sense I just presented. In response to this view I remarked that in some cases, dimensions of vectors might correspond to some sort of constituent structure[19] (e.g., the proportion of activation of different states or processes, as happens with the two chromatic opponent channels in color experience) whereas in other cases vector dimensions do not correspond to any constituent structure (Jakab, 2000, pp. 338-339, 341). Let me emphasize that if one endorses the conclusion that no color experiences have any constituent structure at all, that does not jeopardize the point I want to make here. Quite the contrary, such a conclusion makes my main point even more straightforward.

To summarize, my point is that visual representations of particular shapes are compositional and often quite complex states whereas visual representations of particular colors are much simpler states that are either not themselves compositional, or very minimally so. For this reason, shape perception appears to give us much more information about particular shapes than does color perception about particular colors. I offer this as an explanation of why conceptual revelation is true of shape perception but false of color perception.

**4. Two versions of physicalism about color**

Physicalism about color is the thesis that object colors are physical properties of surfaces, and they are the canonical causes of our experiences of color. On physicalism, object colors are observer-independent. This means that (1) they can exist (i.e., be physically realized) in the absence of perceiving organisms, and (2) they are specifiable without making reference to observers' perceptual responses. Current versions of color physicalism propose to identify colors with types of surface reflectance (Hilbert, 1987; Byrne and Hilbert, 1997, 2003; Matthen, 1988, pp. 24-25; Tye, 1995, pp. 147-148; 2000, Ch. 7; Dretske, 1995, pp. 88-93????). This is so because it seems very unlikely that colors can be identified with some fundamental physical, or chemical properties (Hardin, 1988; Nassau, 1997), and surface reflectance is at any rate the closest empirical correlate of our color sensations. It is also true that surface reflectances are key factors in causing our color sensations.[20] Physicalists about color also attempt to identify colors with local, in some sense intrinsic, properties of surfaces, as opposed to, say, relations between a target surface (to which a color is attributed) and its surround.[21]

There exist, however, two rather different brands of physicalism about color. According to Hilbert, Byrne, Tye, and Dretske, the property of being a color (say, red), does not in any way include a relation to perceivers. Indeed, being red is one and the same property in every possible world: color names are rigid designators (Tye, 2000, note 4 on p. 167). On the alternative account (Jackson and Pargetter, 1997; McLaughlin, 2002), colors are identified with the bases of the dispositions to elicit experiences of color. Something is red only if it fills *the redness-role* (McLaughlin, 2002), that is, it disposes its bearers to look red (in specified circumstances C, to specified types of perceivers P). To this view McLauglin adds that some surface property R is redness only if it is common to all surfaces that are disposed to look red (to perceivers of type P, in circumstances C).[22] McLaughlin explicitly rejects the idea that color names are rigid designators. For instance, the surface property R that disposes its bearers to look red to trichromat humans in average daylight, might not have done so – it might not have played the redness role.[23] Had it not done so, it would not have been redness.

Note some subtleties about this second brand of color physicalism. It is still true on this view that colors can be physically realized in the absence of perceivers, since even if there were, say, no trichromat humans, the bases of the dispositions to look red (to trichromat humans in average daylight) could still be present. Such properties can also be characterized without making reference to reactions of perceivers. Redness can be characterized in terms of surface reflectance, composition of radiant light, and so on. However, notice that redness can be characterized in perceiver-independent terms only if we refer to some of its contingent features that happen to obtain in the actual world, but remain silent about its essential attributes. For playing the redness-role is essential for a property to be redness, but the redness-role consists in eliciting experiences of red (in suitable perceivers, in suitable circumstances). Therefore, characterizing the colors by this essential attribute of them inevitably makes reference to reactions of perceivers.

Perhaps the most dramatic difference between the two approaches to color physicalism just presented is that on the rigid designator account colors are absolute whereas on the non-rigid version they are relative to perceivers and circumstances. In fact, both Jackson and Pargetter (1997) and McLaughlin (2003) are color-relativists. This means that color-roles like the redness-role necessarily include the specification of circumstances and that of perceivers. According to the notion of relativized color, there is no such property as *red, full stop*. What there is is *red, for subject S in circumstance C*. Relativization seems to be a very useful move since dependence of color perception on the circumstances and on perceivers is obvious. What looks bluish green to me in average daylight looks black to me in red light. It might easily happen that, on purchasing a shirt, it looks to me one shade in the store, and a noticeably different one in the street. Similarly, if the shirt looks predominantly bluish green to me in daylight, it might still look predominantly greenish blue to some other person in the very same circumstance.[24] It is widely believed that there are even greater differences between the color perception of different species (Thompson et al., 1992; Matthen, 1999).

To the contrary, the rigid-designator version of color physicalism is absolutist about color. According to Dretske, Tye, Byrne, and Hilbert, a given type of surface reflectance R is one and the same color Q for everyone, therefore it looks Q to trichromat humans in "normal", or "optimal" circumstances of perception. It may not look Q to dichromats, but that's because dichromats do not have "normal" color vision. Similarly, if a surface exhibiting R looks some color other than Q to a trichromat human, then either the circumstances fail to be normal or some kind of *normal misperception* obtains (Matthen, 1988; Dretske, 1995, pp. 88-93; Tye, 2000, pp. 151-162; Haugeland, 1981, p. 18; Matthen and Levy, 1984).

## 5. Problems with color absolutism.

The problem with color absolutism is the following. First, “normal” or “optimal” conditions of perception are pretty variable. Second, independently of the relevant circumstances, normal trichromat humans vary substantially in their color perception. Regarding the circumstances, there are a number of crucial factors influencing color perception: illumination, surround effects (simultaneous contrast), state of adaptation of the eye, and so on. Individual differences in color vision are, to a large extent, due to differences in the spectral sensitivity profiles of the three wavelength-selective retinal photoreceptor classes (Lutze et al., 1990; Neitz and Neitz, 1998; Hardin, 1988, 76-82). The changes in perceived color that these factors produce are, most often, only slight: they are changes in the perceived narrow shade of surfaces, but not in their broader perceptual color category. Ripe tomatoes look red to every trichromat, but the exact shade of red that a particular tomato looks to different trichromats in the same illumination, or to the same observer under two different illuminants seems to vary to a measurable extent.[25]

For instance, take sunlight as the best norm of illumination. Natural lighting by the sun differs widely in different parts of the day, and in different weather conditions (see Shepard, 1997). Despite approximate color constancy (Fairchild, 1998, pp. 156-157; Wandell, 1995, pp. 314-315), perceived narrow shades change with illumination to some extent. In everyday situations we tend not to notice such changes, but they can be demonstrated in the laboratory (Brainard and Wandell, 1992).[26] Simultaneous contrast effects on color perception are also ubiquitous. The perceived narrow shade of a particular surface changes slightly with changes in the color of its surround (Wandell, 1995, Chs. 4, 9; Fairchild, 1998, pp. 135-139; Chichilnisky and Wandell, 1995;

Shepherd, 1999). Finally, due to individual differences in color perception, one and the same surface in the same circumstance often looks different in color to different color-normal subjects.

For these reasons, color science introduced standards. Standard illuminants are lights with  precisely specified relative energy distribution. Standard color contexts for viewing color samples often consist of achromatic grays, blacks, ot otherwise attempt to minimize chromatic induction (Fairchild, 1998, p. 135). To eliminate variation in color perception due to differences in adaptation of the eye, experimental subjects can all be adapted to the same illumination (e.g., dark-adapted before viewing color stimuli on a computer monitor).[27] Any particular standard circumstance for perception can minimize or eliminate within-subject changes in color perception. However, there exist many different standards in color science, each having a different effect on subjects' color perception. Switching between standards does cause changes in the perception of particular color samples by particular subjects.

This problem of standard variation (McLaughlin, 2002, sections 10-11) raises a question for color absolutists. A given sample looks different in color (i.e., narrow shade) in different standard circumstances to the same observer; but which of these circumstances reveals *the true color* of the sample? Remember, for color absolutists a given sample X has one and the same color for everyone, regardless of changes in circumstances of perception.[28] But the color X looks to any perceiver differs in different standard circumstances. So, even if we admit, with Matthen (1988) and Tye (2002) the possibility of normal misperception, we first have to pick and choose which of color science's standard conditions to take as revelatory with respect of the true shades of objects. There seems to be no non-arbitrary ground for such a choice. Furthermore, if we do make this choice, the unavoidable consequence will be that in our everyday life we are virtually never in the circumstance that reveals the true shades of objects, so with very rare exceptions we always misperceive the colors. Individual differences cause a similar problem: Whose color perception reveals the true colors (narrow shades) of objects, given that color-normals often explicitly disagree in their color perception and judgment of one and the same sample in the same circumstance? Pick any one subject and it follows that anyone who disagrees with her about the exact shades of objects (at least 90 per cent of color-normals, on closer scrutiny) will be in error. This is an utterly absurd consequence, and what generated it is the assumption of color absolutism. That's how relativism about color acquires motivation (see McLaughlin, 2002).

## 6. Some replies to the challenge

Naturally, color absolutists have tried to counter the challenge just presented. Tye (2000, pp. 89-93) argues that individual differences in color perception are properly understood as differences in one's capacity to discriminate colors. When two subjects A and B look at the same color sample S in the same circumstances, and S looks pure green to A whereas it looks slightly bluish green to B, what is going on, according to Tye, is that B has better color-discrimination than A. B is capable of discerning the tinge of blueness in the sample that A does not notice. The difference is like that between two gauges, one more finely calibrated than the other. A blunt ruler may say of a steel rod that it is roughly 19 inches long; a better ruler may say of the same rod that it is 19.35 inches long – neither ruler is mistaken. The same moral applies to individual differences in color vision, Tye contends. Thus the existing individual differences between trichromat subjects do not support the conclusion arising from color absolutism that some, perhaps many, color-normals misperceive the narrow shades of objects. Tye also thinks that the phenomenon of color constancy is robust enough so that changes in normal illumination cause at most ignorable changes in perceived shades (Tye, 2000, pp. 147, 150), and that simultaneous contrast effects occur only occasionally, therefore they can be safely regarded as cases of normal misperception, on a par with shape illusions.

Regarding the latter two claims, they are simply false (see Fairchild, 1998, pp. 156-157; Wandell, 1995, pp. 314-315). Tye's suggestion about color constancy (2000, pp. 147, 150) is made to sound plausible by effectively confusing broad color categories and narrow shades under the term 'color'.[29] Moreover, simultaneous contrast effects are ubiquitous, not occasional. Any particular reflecting surface under constant illumination can look a whole variety of different shades (even very different ones) to the same subject, depending on its surround. So the question which of the surrounds reveals *the true color* of the sample continues to bother the absolutist. Perhaps some neutral mid-gray background reveals the true shades of objects? Again, any such choice seems entirely arbitrary (McLaughlin, 2002), and it implies that in everyday life we very rarely see the true shades of objects – perhaps we do so only in the color scientist's laboratory.

Tye's proposal about individual differences, despite being ingenious, does not work either. Variation in trichromat color perception can take forms that cannot be accounted for by Tye's calibration approach. If a green sample looks slightly yellowish green to subject A, and slightly bluish green to B, this cannot be a mere difference in their ability to discriminate colors. A discerns a tinge of yellowishness, whereas B discerns a tinge of bluishness; now the question, which one of them is right, arises in a nastier way for the color absolutist. The sample is either yellowish or bluish; there is strong reason to assume, within the absolutist approach, that it cannot simultaneously be both. Moreover, if, by assumption, the sample is objectively bluish green, whereas it looks to subject A yellowish green, then it seems very hard for Tye to avoid the conclusion that A, a perfectly normal color perceiver, misperceives the sample. There exist some data suggesting that this sort of individual variation might actually exist (Jakab, 2001, Ch. 6, 2003).

Byrne and Hilbert (1997, 2003) propose two other ways to accommodate variation in trichromat color perception in the absolutist approach. Their first proposal (Byrne and Hilbert, 1997, pp. 272-274) is that narrow shades or certain specific color categories like *unique green* or *slightly bluish green* are not contraries: if a surface is unique green, it does not follow that it is thereby not slightly bluish green. Some bluish greens are certainly not unique greens, but others may be. Byrne and Hilbert offer an analogy: there are shapes that are both squares and diamonds, namely, any square standing on one of its corners. Yet it might mistakenly appear to someone that no diamonds are ever squares – perhaps the idea that no bluish greens are ever unique greens is the same way mistaken, suggest Byrne and Hilbert. Of course, there are many diamonds that aren't squares, but there are also bluish greens that are not unique greens, so this analogy between shape and color seems to obtain.

If it is accepted that unique green and bluish green *are* contraries, then it still remains true that seeing a unique green surface as green with a tinge of blue is only a marginal error and such a color perception is for the most part veridical. (I.e., that the surface is predominantly green is correctly perceived).

In response, it is worth noting that even if being bluish green is compatible with being unique green in the way Byrne and Hilbert assume (something I do not believe), being slightly yellowish green and being slightly bluish green are certainly incompatible properties simply because being bluish and being yellowish are incompatible.[30] So if a sample looks slightly yellowish green to A and slightly bluish green to B, then at least one of them must be in error. In a *marginal* error only, Byrne and Hilbert would add; in a marginal error *with respect to broad color categories*, I would add. But what is a marginal error with respect to broad color categories is a total error with respect to narrow shades. If, by assumption, a particular surface is slightly bluish green, but it looks to me slightly yellowish green, then I completely misidentify the narrow shade (the *determinate*, in Byrne and Hilbert's terms: 1997, pp. 266-267, 276-278, 280-281) in question.[31] I might still be correct about which broad color category (determinable) the perceived surface belongs in, but regarding the narrow shade, I am simply off the mark. Thus variation in trichromat color perception either prevents us from specifying the narrow shades in terms of surface reflectance (i.e., as metameric sets), or it forces us to admit that the overwhelming majority of color-normal people, in most circumstances, misperceive the narrow shades of objects in normal circumstances of perception.

In their most recent article, Byrne and Hilbert (2003) choose the latter way. They argue that a slight but ubiquitous error with respect to the visual perception of shapes occurs fairly frequently in humans.[32] So it should come as no surprise if such errors obtain in "normal" color vision as well. Byrne and Hilbert would ask that since the existence of this kind of error does not make relativism about shape the least bit plausible, why should it make color relativism any more plausible? Byrne and Hilbert note that there is a relevant difference between color perception and shape perception, namely that in the latter case there are "independent tests" for the exact veridicality of shape perception, whereas there is no comparable test for color perception (Byrne and Hilbert, 2003, section 3.4.). But, they suggest, this difference does not support color-relativism – it only leads to a strange form of agnosticism about narrow shades. Since to different trichromat subjects different metameric sets will look unique green, and since there is no independent test for the correctness of the perception of narrow shades, it is in principle impossible to decide who is right about perceiving unique green, and so it is in principle impossible to establish which metameric set *is* unique green – objectively, or in absolute terms. Mutatis mutandis for all other narrow shades, we should add. We can probably learn the reflectance basis of the eight broad chromatic categories (red, green, yellow, blue, orange, purple, yellowish green and bluish green) plus that of black, mid-gray and white. Perhaps we can do a little better, characterizing narrower color categories like scarlet or navy blue in terms of reflectance. But just go to a paint shop, look at the thousands of available narrow shades each one of which is affected by the problem of agnosticism. The sample called Patio Pink very likely appears slightly different in color to different normal color perceivers, even in the same circumstances. In the other direction, the color that *is* 70 per cent reddish and 30 per cent bluish, cannot be characterized in terms of reflectance. This is because to different normal color perceivers different and often non-overlapping narrow ranges of reflectance (metameric sets) will appear 70 per cent reddish and 30 per cent bluish. Thus on Byrne and Hilbert's view, the reflectance basis of all narrow shades is in principle unknowable. However, there still exist the narrow shades that are objectively unique green, or objectively 70 per cent bluish and 30 per cent reddish, and so on. This is Byrne and Hilbert's absolutist reply to the problem of individual differences.

**7. Why relativizing colors is plausible whereas relativizing shapes would be absurd**

Here is a difference between color perception and shape perception that follows from Byrne and Hilbert's view. Individual differences in shape perception together with the (very plausible) view of type physicalism and absolutism about shape do not lead us to agnosticism about shape. However, individual differences in normal color perception together with type physicalism and absolutism about color do lead us to agnosticism about color, as Byrne and Hilbert admit (2003, note 50).

Question: what is the reason for this difference between color and shape? Byrne and Hilbert's answer: there is an independent test for *correctness* of shape perception, whereas no such test exists for color perception. This sounds right, but needs a little explication. So the next question is, what is the reason for this latter difference? Why isn't there an independent test for the correctness of color vision? Byrne and Hilbert's answer: the lack of such an independent test for color vision is partly due to the fact that colors are perceived by only one modality (unlike shapes); in addition, colors, contrary to shapes, do not play any interesting causal role, hence they do not figure significantly in the data and theories of other sciences.[33]

I think this latter reply is wrong. In my opinion, we have an independent test for the correctness of shape perception because shape perception is *conceptually revelatory* with respect to shapes, therefore there exists a normative connection between shapes (the relevant stimulus properties) and the corresponding types of visual shape experience. For instance, the visual experience of something being circular is supposed to be such that it makes available to intellectual reflection the essence of being circular. Here is the story in more detail.

As we saw above, shape percepts are compositional. Perceptual states representing particular shapes on particular occasions have constituent structure, and so they encode structural information about the physical conditions that are their canonical causes, namely particular shapes. Color percepts, on the other hand, are representationally atomic – they do not encode structural information about their causes (surface reflectances, wavelength compositions, etc.). Independently of this, in both cases (that of color perception and that of shape perception), the canonical causes of the percepts can be described in perception-independent terms.

Due to the structural information encoding just mentioned, there are two ways in which shape percepts carry information about shapes:

(1) Lawlike covariation or tracking. Optimally, percepts as of squares track squares; percepts as of spheres track spheres, and so on.

(2) Structural information encoding (as explained above).

Now, shape perception is mistaken when there is a discrepancy between (1) and (2): for instance, when a shape percept $P_s$ covaries with certain ovals (or egg-shapes) in the environment, yet it delivers to the processes that operate on it the information that each point on the indicated shape’s perimeter (or 3D surface) is at equal distance from its center. (I.e., the structural-information-encoding part conveys information about something round or spherical, when in fact an oval or egg-shape appears in the scene).

Due to the lack of structural information encoding, there is only one way color percepts carry information about the colors, and that is tracking. Since there is no structural encoding in color percepts, there is no possible discrepancy between the two sources of information.

Here is the most important implication of this story. The required synchrony or harmony between the two sources of information (1) and (2) is the source of the normative link between shapes and shape percepts. This is why, on finding individual differences in shape perception, we can establish who is right and how shape perception *should* work. And the lack of (2) in the case of color perception is the reason why the normative link is missing between colors and color percepts.

This difference in normativity between shape perception and color perception explains why it is reasonable to relativize colors, but not shapes. No matter whether a subject locates unique green on one Munsell chip or rather another (see Kuehni, 2001), as long as she performs well in color discrimination, her color vison is normal *and veridical* – there is no non-question-begging reason to suppose otherwise. There is no theoretical reason whatsoever to anchor unique green to this, rather that that, narrow reflectance range. That's why 'unique green full stop' is indeterminate in its reference. However, 'unique green for subject S, in circumstances C, at time t', is not the same way indeterminate.

Moreover, recognizing the difference in normativity between shape perception and color perception helps to come to terms with the independent test problem raised by

Byrne and Hilbert. What Byrne and Hilbert call an independent test for shape perception is checking for the harmony between the two sources of information in shape perception: tracking and structural encoding. For instance, does a subject find circular objects circular? Do her visual percepts that track circular shapes induce the behavioral output in her that is a report of a circular shape? Note that this is done simply by showing circles to the subject (the tacit assumption is that doing so will elicit the type of visual percept in her that tracks circles), and asking her what shape they look. If she reports an ellipse, she might have astigmatism, aniseikonia, or some other opthalmological (or neurological) condition.

The normative element in shape perception is quite important. Shape perception is supposed to reveal the properties that are the *particular* shapes in rich (and correct) detail. In Dretske's evolutionary terminology (Dretske, 1988, 1995, chapter 1) shape perception was selected for this achievement, and so if function is understood in terms of evolutionary history, then shape perception has the function of providing rich detail about particular shapes to humans and various animals. Of course, providing such detail served skills like spatial navigation and recognition – that's how it added to the organism's fitness. (Conceptual revelation was hardly the performance for which perception of 3D space and shape were selected.) Note also that the mere need for effective discrimination and recognition of shapes does not necessitate structural encoding, but complex behavioral interaction with shapes, like navigation or manipulation, do. If all an organism needed would be to discriminate and recognize shapes, then its visual system could in principle succeed by encoding them in an low-dimensional similarity space where particular shapes would be represented as single points, quite similarly to the representation of colors in color space.

On the contrary, color vision does not structurally encode particular object colors. However, this simpler form of representation already helps the organism to discriminate surfaces based on their narrow shade, and recognize them based on broad color categories. And that's all animals and most humans need color for. Animals and humans discriminate and recognize objects and surfaces[34] based on their color, and they interact in complicated ways with the surfaces and objects themselves, not with their colors. They pick the berry, not its redness. Probably for these reasons, color vision does not provide any detailed information about the rather complex states of affairs that are the particular colors. These states of affairs include surface reflectances of target objects, their interaction with reflectances of surrounding surfaces and the illuminant, and a number of other factors. Color perception does handle such information (see Maloney, 1999, 2002), but it does not make it available to the rest of the cognitive system.

## 8. Martin Davies on externalism and experience

### 8.1. An externalist defense

In one of his papers, Martin Davies sets out to establish representational externalism about the contents of perceptual experience (Davies, 1997).[35] Perceptual content, the representational content of perceptual states is supposed to differ from belief content (or propositional content), in some characteristic ways, and these differences make externalism harder to esablish for perceptual content than for belief content.

First, perceptual content is thought to be *non-conceptual*, in the sense that a subject can have an experience E without possessing those concepts that would be used in specifying the content of E (Davies, 1997, pp. 309-310; see also Tye, 1995, p. 139; 2000, p. 62). For instance, it is possible to undergo the experience of red without having the concept RED. Second, perceptual content is *not object-involving* (Davies, 1997, p. 310), or *abstract* (Tye, 1995, p. 138; 2000, p. 62). This means that the numerical identity of perceived objects has no role in determining the sameness or difference of contents, only their qualitative identity does. Two objects that are exactly alike regarding their perceivable properties can be substituted for each other without altering the perceptual content they give rise to. Third, perceptual content is *fully representational*, meaning that it has correctness conditions (Davies, 1997, p. 310).[36]

The externalist claim for perceptual experience consists of a refutation of modal individualism, or the local supervenience thesis. According to this thesis, the representational content of perceptual states supervenes on the internal state of the subject, and so this sort of content cannot differ in exact duplicates no matter what counterfactual situations we devise for them. This amounts to constitutive externalism, the idea that the representational content of perceptual states (and hence the type of contentful mental states) is essentially dependent on relations between those states and the environment (Davies, 1997, p. 313).

In defending this position for perceptual experience, the externalist has a much more limited degree of freedom than in the case of belief content. For instance, some externalist arguments for belief content appeal to the social environment, and try to show that belief content is dependent on the individual's social context (Burge, 1979; Davies, 1997, p. 313). But since perceptual content is non-conceptual, it sounds implausible to make a similar claim for perception. On the other hand, since perceptual content is not object-involving (Davies, 1997, pp. 310, 314), modal externalist examples that vary the numerical identity of the object represented while leaving all its perceivable (or other) attributes unchanged in order to change content do not work for perceptual content either. Two percepts that represent their objects by the same perceivable properties have the same content, differences in the numerical identity of their objects notwithstanding.[37] The idea that perceptual content is not object-involving means that perceptual experiences have only what is called *narrow content*, but not wide content (Fodor, 1987, Ch. 2, 1991; Davies, 1997, pp. 314-315). Narrow content includes perceivable properties, but not the numerical identity of the object represented. Thus perceptual states are fully representational: their correctness conditions simply do not include numerical identity (Davies, 1997, p. 315).

To this I would add that the same applies to externalist arguments varying "hidden essences". Even though my concept WATER and my duplicate's concept WATER on Twin Earth refer to different natural kinds ($H_2O$ and XYZ respectively), and so they have different content, *water percepts* have the same content in both the denizens of Earth and those of Twin Earth, since the content of percepts is limited to perceivable properties, and by assumption, the difference between $H_2O$ and XYZ is not perceivable.

The challenge for the externalist is to provide an example, despite the above limitations, in which two duplicate subjects $D_1$ and $D_2$, in different environments $E_1$ and $E_2$ respectively, undergo some physiologically and functionally identical perceptual state type $S_0$, yet due to environmental differences, the $S_0$ tokens in $D_1$ and $D_2$ acquire different perceptual content. There are additional challenges that the externalist has to face. The content of perceptual experiences is arguably dependent on the resulting behavior as well, not just on the causal antecedent that those experiences track or covary with. Assume that subject $D_1$ in some actual-world situation $E_1$ undergoes perceptual state type $S_0$; his duplicate $D_2$ in counterfactual situation $E_2$ also undergoes $S_0$, and the only difference between the two situations is that $S_0$ has different causal antecedents in them. If subject $D_1$ exhibits some behavior in $E_1$ that is specific and adequate to $S_0$'s object, and this behavior perseveres into the counterfactual situation, then the individualist can claim that $S_0$'s content is preserved through the transiton: it is the same in $E_2$ as in $E_1$. This is called the *conservative stance* (Davies, 1997, pp. 315-318). If, on the other hand, subject $D_1$ in situation $E_1$ does not exhibit any behavior that is specifically adequate to $S_0$'s object, then the individualist can adopt a *revisionary stance* claiming that the content of $S_0$ in both situations is indeterminate: it is $X_1$-or-$X_2$, where $X_1$ is $S_0$'s stimulus in $E_1$, and $X_2$ is $S_0$'s stimulus in $E_2$.

Davies puts this latter point in a less abstract way using Tyler Burge's example of shadows and cracks (Burge, 1986, 1988). There $X_1$ in $E_1$ is some strip of dark shadow whereas $X_2$ in $E_2$ is some crack (e.g., in the rocky soil), the two being perceptually indistinguishable. $D_1$ and $D_2$ are individuals of some animal species. First assume that the shadows and the cracks are sufficiently wide so that the shadows ($X_1$) in $E_1$ provide protection form direct sunlight whereas the cracks ($X_2$) in $E_2$ are hazardous as the animal can fall into them. If $D_1$ in $E_1$ sees the shadows as shadows, then it will approach them. If this approaching behavior carries over into $E_2$, then the duplicates will fall into the cracks, and this motivates the the ascription of the same content (namely shadow) to $S_0$ in

$E_2$ as in $E_1$ – the conservative stance is in order. If the shadows and cracks are narrow, and they are neither protection nor threat to the animal, consequently evoking no specific behavior in any of the two situations, then the indeterminacy in content ascription sounds more reasonable: in both situations, $S_0$'s content is *shadow-or-crack*. Either way, the switch between $E_1$ and $E_2$ makes no difference to $S_0$'s content, and the individualist apparently wins the game.

The way out for the externalist is to create a difference between the subject's and its duplicate's behavior in a way that is consistent with their being duplicates (Davies, 1997, p. 318). In addition, the externalist has to assure that in both the actual and the counterfactual situation all factors that play a role in determining content are in harmony, that is, support the same content ascription. Three such factors are taken into account by Davies (p. 319): input (causal antecedents of percepts), output (behavior), and evolutionary history (or function, understood in terms of evolutionary history). In Davies's most developed example (pp. 321-322), the subject Percy in the actual situation ($E_1$) sees ellipses as ellipses and circles as circles. In $E_1$ ellipses evoke perceptual state $S_1$ in Percy, and circles, state $S_2$. Moreover, in $E_1$ the ensuing behavior on $S_1$-tokenings is specifically adequate to ellipses, and $S_2$-tokenings result in "circlish" behavior. In the counterfactual situation $E_2$, $S_1$ is caused by circles, due to some systematic optical distortion. Moreover, the assumption is that in $E_2$, $S_1$ initiates behaviors that are internally indistinguishable from what it initiates in $E_1$ (i.e., the evoked nerve firings, muscle contractions, etc. are all the same), but due to some gravitational or so distortion, the actual trajectories of the duplicate's movements will be specifically adequate to circular shapes. Finally, in the counterfactual situation, this has been so all along in the course of evolution. This means, Davies contends, that $S_1$ acquires a content in $E_2$ that is different form its content in $E_1$. All three factors enumerated support this view, whereas back in $E_1$, all three factors support the alternative content-ascription to $S_1$ (i.e., ellipse).

Given this result, Davies claims that it would be question-begging against externalism to insist that the duplicate must see circles as having just two axes of symmetry. Presumably by this remark he wishes to alleviate worries that the intrinsic character of the evoked perceptual states and their processing might determine how many axes of symmetry we see shapes as having. As I shall argue in a moment, there is no question-begging here, but rather a genuine problem for the externalist. Another clarificatory remark Davies adds is that his externalist example is somewhat unrealistic; in particular, the range of relevant behavioral interactions with the environment is extremely simplified. This, Davies claims, does not undermine the purpose of his argument, but it does indicate that it will be difficult to devise externalist examples involving shape perception.

With these remarks and the key example at hand, Davies rejects both the conservative and the revisionary interpretations of his example. The conservative stance is rejected since none of the three factors supports the original content ascription in the counterfactual situation. The revisionary stance is rejected as in both situations, the subject's behavior is supposed to be distinctively appropriate to the object of $S_1$.

### 8.2. A problem

In a nutshell, the problem with Davies's proposal is that there is a fourth factor of content determination that he does not consider in his argument: inferential role. Due to structural encoding, shape percepts acquire quite specific inferential roles not characteristic of color percepts (see Section 7 above for explanation, and below for examples). These roles are internally supervenient, and for this reason they break the harmony of the content-endowing factors in the counterfactual situation. Inferential role is dependent solely on the causal or computational processes that operate on representations – that is, the vehicles of representation, the physical substrates (datastructures, internal symbols, etc.) that, one way or another, acquire content.[38] On the other hand, Davies's three factors of content determination (tracking stimuli, evolutionary history, and behavior) are externalist, that is, they essentially involve relations between the organism and its environment. It seems to me that Davies tacitly assumes, throughout his paper, that visual shape percepts are representationally atomic states, therefore they do not assume specific inferential roles, and so it is enough to consider the other three factors of content determination. This, however, is not the case. Here is the point in more detail.

Assume that Percy's duplicate in the counterfactual situation undergues a perceptual state $S_1$ that is derived from retinal images that are themselves elliptical in shape, and so in Percy in the actual situation such retinal images, coupled with $S_1$, would track ellipses (Davies, 1997, p. 321). In Percy's duplicate, however, $S_1$ and the corresponding retinal image covary with circles, and give rise to "circlish" behavior, and this has always been so in evolution. Davies argues that in such a case $S_1$ is a full-blown representation of circles in the counterfactual situation – the harmony of causal antecedents, behavioral consequences, and evolutionary history support this conclusion.

The problem, however, is that an inner state that is to count as the (visual) perceptual representation of circles, *must be capable of giving rise* to the inference (via some computational derivation) that circles have infinitely many axes of symmetry. This is not satisfied for Percy's duplicate in the counterfactual situation, because the duplicate, on looking at circles in $E_2$, undergoes the same representational state $S_1$ as Percy does in $E_1$ on looking at ellipses. In addition, all processes of whatever computational derivation are qualitatively identical in Percy and his duplicate. For this reason, $S_1$ in the duplicate's mind in the counterfactual situation is not a full-blown perceptual representation of circles. Note that Davies's reply to this problem (1997, pp. 321-322) is unsatisfactory. It is true that subjects sometimes see shapes as having fewer axes of symmetry than they really have. But the real question is, *can* the subject recognize, on more careful looking, or reflection, the true number of axes? Normally people can recognize, on reflection, that squares have four axes of symmetry, even if they originally saw these shapes as having only two such axes. However, Percy, on undergoing $S_1$ in the counterfactual situation, will not be able to infer, or recognize, that circles have infinitely many axes of symmetry. This is because in him it is $S_1$ that tracks circles, but $S_1$ is a perceptual representation that structurally encodes ellipses, and so the internal processing routines that operate on it cannot, without altering or distortion, extract from it the information that the shape it representes has infinitely many axes of symmetry. These routines, on processing $S_1$, can only output two as the number of axes of symmetry. As we now see, it is the *capacity* of perceptual representations to give rise to certain inferences (or *protopropositional content*: Peacocke, 1992, p. 77; Davies, 1997, pp. 310, 321) that is relevant to the attribution of perceptual content, not the inferences actually drawn, or the protopropositional content actually derived, at any point of time.

We can reformulate this argument at the level of personal, instead of sub-personal level of description. I think of, or imagine, a certain type of ellipse E, with an approximately 2:1 ratio of its major and minor axes. Based on the arising representation, I will not be able to reasonably infer that the shape I'm now thinking of, or imagining, has infinitely many axes of symmetry. Nor would I be able to make such an inference on looking at ellipses like E. Now assume that to my duplicate in a counterfactual world $W_{6623}$ it is circles that visually appear like E. That is, on looking at circles, my duplicate will not be able to reasonably infer that the shape he is looking at has infinitely many axes of symmetry. (Obviously, my visual system in the actual world does make such an inference available to me.) If circles produce ellipsis-encoding visual representations in my duplicate, then, on looking at circles, he will be inclined to infer that the shapes he is looking at have only two axes of symmetry. But the shapes he is looking at do, in fact, have infinitely many axes of symmetry. This means that my duplicate will be significantly *misrepresenting* circles, at least at some (but arguably all) levels of consciously accessible representation. This is so despite the fact that his movements, being also "squashed along the vertical axis" (Davies, 1997, p. 321), are in harmony with his distorted perception. Circles evoke ellipsis-encoding visual representations in him, the corresponding outgoing motor commands are for ellipsis-related movements, but some distorting environmental effect makes the actually occurring movements adequate to circles. Moreover, in world $W_{6623}$, this has been so all along in the course of evolution. So why stimulus correlates, ensuing behavior, and evolutionary history, all in harmony, still cannot uncontroversially fix perceptual content? The problem is that my duplicate will think “ellipsis thoughts”, and draw “ellipsis inferences”, on the basis of perceptual representations structurally encoding ellipses, just like I do. The distorted perceptual and behavioral relations with the environment do not affect this internal relation: this internal relation cannot change, or else he isn’t my duplicate. Thus despite the envisaged harmony between stimuli, behavior, and evolutionary history (function), the externalist position is not secured for visual experiences of shape.

Davies’examples for establishing the externalist case are about shape perception. He admits that there are important simplifying assumptions in his examples (the most important of which is that the range of behavioral interactions with the relevant shapes is extremely limited), and that a sufficiently complex externalist example about shape experience will be difficult to provide (p. 322). I agree with these cautions entirely. In addition, I have argued in this section that even Davies’s simplified externalist example about visual experiences of shape faces serious problems. One can try to solve the above controversy by endorsing the conservative stance (pp. 315, 322). This would mean claiming that the duplicate’s perceptual representation in the counterfactual situation has the same content as in the actual situation. (In the particular example I provided here, the content is in both cases that an elliptical object is present.) This may strike the reader as implausible, since in the counterfactual situation the stimulus correlates (circles), behavior (circle-related behavior), and evolutionary history oppose this interpretation in concert.

If we refuse to accept the conservative stance, then we can try out the revisionary stance (pp. 315, 322), claiming that perceptual state $S_1$ has the content circle-or-ellipse in both Percy and his duplicate. This does not sound plausible for more than one reason. First, based on the occurrence of percept $S_1$, both Percy and his duplicate will be disposed to infer that the shape he is looking at has two axes of symmetry, period. But this is not an adequate inference to ellipses-or-circles. If something is an ellipse-or-circle, then that thing has either two or infinitely many axes of symmetry. But neither Percy's nor his duplicate's percept $S_1$ will support such an inference. So the structural-encoding-based inferential role aspect of $S_1$'s content is definitely not about circles-or-ellipses. This aspect of $S_1$'s content is determinately about ellipses. Second, in the actual and the counterfactual situation the arising behavioral responses are different and in both cases specifically adequate to the stimulus.

Controversy attends the externalist content ascription as well. In the actual situation everything is in harmony (structural encoding, stimulus correlate, behavior, and evolutionary history). In the counterfactual situation, only three of these factors support the new content attribution, the fourth (structural encoding) opposes it.

It seems that two alternative replies can be given to this apparent dead end. First, one might remark that even the oversimplified externalist example for the content of shape percepts isn't working, so we might just as well abandon extenalism about perceptual content altogether. On careful scrutiny, Davies's example, instead of establishing externalism, has lead us in an impasse. So perhaps we just don't know what to say about his counterfactual situations. Second, it is possible to argue that the conservative stance is not so bad after all. Evolution sometimes succeeds by systematically misrepresenting, or distorting, types, or ranges, of stimuli in perceptual representation. Perceptual representation need not be veridical in every aspect, in order to be adaptive (Akins, 1996, p. 364**????check what she says there**). The perceived similarity of color is one often-cited example of this phenomenon (Matthen, 1999; Thompson, 1995, pp. 122-133; 2000). For instance, in terms of wavelength distribution of emitted or reflected light, or surface reflectance, violet is more similar to green than to red. However, perceptually, violet looks more similar to red than to green. There are numerous similar examples from perceived similarity relations of the colors. Strictly speaking, this means that color perception systematically misrepresents object colors in terms of their similarity relations. The content of our color perceptions has it that violet is more similar to red than to green, notwithstanding the opposite relation in terms of the relevant stimulus properties.[39] Note also that behavior in this case is in harmony with (mis)perceived color similarity, not with actual stimulus similarity. For instance we do claim in everyday situations that red and violet objects are more similar in respect of color than violets and green ones.

In the spirit of the conservative stance, if circles are misperceived *as ellipses* in the counterfactual situation, then the content of the percept in question (i.e., P) is *about ellipses* in that situation, just like in the actual one. In the spirit of the externalist stance, all one can say is that even though $S_1$ veridically represents circles as circles in the counterfactual situation (due to the harmony between stimuli, behavior, and evolutionary history), still Percy's duplicate in that situation simply cannot draw the inferences he should be able to draw about circles. But he can do so in the actual situation; so the externalist has to come to terms with the difference in inferential ability between Percy in the actual situation and his duplicate in the counterfactual one. We can put this point by saying that if we take the conservative stance, then we have a case of *effective* *misrepresentation*: in the counterfactual situation circles are misrepresented as ellipses, whereas within the subject's cognitive system everything works as one would expect on the basis of the content ascription. If, however, we take the externalist stance, then what we have is a case of *ineffective veridical representation*, where in the counterfactual situation circles are by assumption veridically represented as circles, yet the rest of the cognitive system cannot make proper use of this veridical representation. I personally prefer the conservative stance, though I am not going to provide more arguments in favor of it.

Once again, note that if we allow a fine-tuning between Percy's visual representations and his propositional ones in order to bring the inferential role aspect of perceptual content in harmony with the other three factors, then we evidently violate the externalist assumption – we no longer have duplicates. Thus at a minimum we have shown that even Davies's oversimplified examples designed to support representational externalism about perceptual content cannot do their job. I think this result corroborates the intuition that the content of visual shape percepts is not a matter of relations between the organism's perceptual states and its environment.

### 8.3. Another externalist-internalist exchange about perceptual content and phenomenal character

After supporting his externalist stance about experience, Davies examines two antecedently plausible ideas from which an anti-externalist conclusion seems to follow. The first idea is that perceptual content supervenes on phenomenal character (perceptual content is a matter of how things seem to the conscious subject). The second is that phenomenal character supervenes on internal constitution. These two propositions, together with the transitivity of supervenience, entail local supervenience for perceptual content – the very idea that Davies is set out to defy. The problem is, how can perceptual content be externalist if it supervenes on phenomenology, and phenomenology supervenes on internal constitution (and supervenience is transitive)?

Davies's reply to this counterargument is to distinguish two different kinds of supervenience that are involved in the two propositions. For the sake of argument he accepts the idea that phenomenal character is modally strongly internally supervenient (a so-called XYWW' supervenience is in action: Davies, 1997, pp. 312-313; McFetridge, 1985) and is also non-representational. However, perceptual content supervenes on phenomenonogy in a weaker sense only: within a subject and within a world, there is no difference in content without some difference in phenomenology (XXWW supervenience).

This way, the clash between Davies's externalist conclusion (Davies, 1997, section 5) and the two antecedently plausible intuitions is avoided (pp. 322-323). This result also implies that exact physical *and phenomenal* duplicates in two different possible worlds can still differ in perceptual content in those two worlds (p.324).

Following Peacocke (1983, chapter 1; see also Peacocke, 1997), Davies considers the position of qualia being *sensational properties*. Phenomenal characters on this approach are understood as internally supervenient and non-representational. Davies raises the question: are such sensational properties rarities, like Peacocke's examples suggest (like the difference between looking at the same scene with one eye or two: Davies, p. 325), or, to the contrary, there exist a host of such non-representational phenomenal properties “upon which the representational superstructure of perceptual content is erected” (p. 324), or which subvene under the myriad representational properties of every perceptual experience (p. 325).

If we take the second route, the following dilemma arises (p. 325). How strong is the correlation between the sensational substrate and the representational superstructure?

(H1) Assume it is relatively strong. Then there must be some theoretical principles that govern, or secure, this correlation. Where might these principles come from?

(H2) Assume that the correlation is virtually nonexistent – there is essentially a free variation between sensational properties and representational ones. Then it follows that a duplicate of mine can enjoy experiences that are phenomenally identical with mine, yet representationally entirely different.

Davies finds this a daunting dilemma, so he prefers to avoid it. His suggestion is that we should reconsider the commitment to a sensational substrate. In particular, we should reconsider the idea *that there is a phenomenal level of perceptual experience that is modally strongly supervenient on internal constitution* (p. 325). If we take this route, we can still retain sensational properties as relatively rare peculiarities, but not as a comprehensive subvening basis for all perceptual representational properties. In addition, we can still consistently hold that perceptual content is a phenomenological notion. That is, we can maintain that perceptual content is essentially an aspect of what it is like to have an experience. But the second idea, namely that phenomenal character is intrinsic and non-representational, has to be given up.

On this proposal, phenomenal character becomes (A1) a representational notion, that is (A2) not independent of the externalist factors that determine perceptual content. This is the tentative suggestion with which Davies ends his paper.

I wish to comment on this line of argument. In particular, I want to argue that it is less convincing than it might appear at first. For Davies does not consider a view of phenomenal character that is highly relevant in this context. It is possible to claim that phenomenal character is fully representational, yet it is modally strongly internally supervenient, that is, accepting A1, but not A2. This view is called *narrow intentionalism* (McLaughlin, 2002, Sections 14-16). The idea that we *should* reconsider, with Davies, is that there is a phenomenal level of perceptual experience that is modally strongly supervenient on internal constitution AND is also non-representational. I think we should reject this idea, just because we should reject the second conjunct of it. But this does not amount to rejecting the assumption of a modally strongly internally supervenient phenomenal level. Among other things, the phenomenon of *transparency* (Tye, 2000, Ch. 3; McLaughlin, 2003b, Sec. 15; Harman, 1990) supports the idea that phenomenal characters are inherently representational. We do not perceive the phenomenal characters of our experiences. For instance, we do not perceive *what it is like to see red*. What we perceive is the color red. What it is like to see the colors is the phenomenal character of color experience. Phenomenal character is not the *object* of perception – it is the *mode* of perception. So due to the phenomenal character of our experience of red, ripe tomatoes are presented to us in vision as being in a certain way. However, there is reason to assume that phenomenal character is not an externalist notion – not at least for color. What makes this assertion plausible is that so far the most sophisticated representational externalist accounts of phenomenal character have encountered apparently insurmountable obstacles in accounting for color expeirence. For the views I have in mind see Matthen, 1988, pp. 24-25; Dretske, 1995, pp. 88-93; Tye, 1995, Ch. 5; 2000, Ch. 7; Byrne and Hilbert, 1997, 2003. For critique see Thompson et al., 1992; Matthen, 1999; McLaughlin, 2003; Jakab and McLaughlin, 2003; Jakab, 2003a, 2003b, 2001. For a related critique that uses other perceptual modalities as examples see Akins (1996).

If we take the narrow intentionalist route (as I think we should), then we can say the following.

(a) Phenomenal character is internally supervenient.

(b) Every phenomenal difference is *ipso facto* a representational difference.[40]

(c) There might still be an externalist aspect of perceptual content arising from stimulus correlates, but this aspect of content does not show up in phenomenal character. There is no modally strong link of supervenience between externalist content and phenomenal character.

This leads us close to position H2 above (for Davies's statement, see p. 325): there is no across-world, across-subject supervenience of phenomenal character upon *externalist* content. There might still be a strong within-world, within-subject correlation between the two. On narrow intentionalism, phenomenal character inversion (McLaughlin, 1997, sec. 14; Byrne and Hilbert, 1997, pp. 267-272) in general is impossible, simply because of (b). However, phenomenal character inversion, with respect to the externalist aspect of content only, is straightforwardly possible. For the case of color, Ned Block's Inverted Earth scenario (Block, 1997) demonstrates such a case.[41] Content in general does not supervene on phenomenal character, but the internally supervenient part of it – whatever content that phenomenal character gives rise to – does. Note, however, some differences between the view rejected by Davies in favor of externalism, and the view I am proposing here. On the former, phenomenal character is thought to have a non-representational aspect, and that is why it can vary without perceptual content varying. On the latter, externalist content does not affect phenomenal character, and so externalist content can vary independently of phenomenal character, even though phenomenal character has no non-representational aspect. So the reasons for the de-coupling of perceptual content and phenomenal character are different in the two cases.

**9. Summary**

I argued that there is an important difference with respect to compositionality between color perception and the visual perception of shape. In addition, this difference helps to explain certain phenomena in the realm of visual perception. These phenomena include perceptual revelation and the normativity inherent in visual perception. The same difference in compositionality also helps us to properly formulate certain objections to externalist accounts of perceptual content.

**References:**

Block, N. (1999). Sexism, Racism, Ageism, and the Nature of Consciousness. *Philosophical Topics*, **26**, (1-2), 39-70.

Bradley, P., & Tye, M. (2001). Of colors, kestrels, caterpillars and leaves. *Journal of Philosophy*, **98**, 469-487.

Brainard, D., H., Wandell, B. A. (1992). Asymmetric colr matching: how color appearance depends on the illuminant, *Journal of the Optical Society of America A*, **9** (9), 1433-1448.

Byrne, A., Hilbert, D., R. (1997b): Colors and Reflectances. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 1. (The Philosophy of Color), Cambridge Mass. The MIT Press, 263-288.

Campbell, J. (1993). A Simple View of Color. In John Haldane and Crispin Wright (Eds.): *Reality, Representation and Projection*. Oxford University Press

Chichilnisky, E., J., Wandell, B., A. (1995). Photoreceptor Sensitivity Changes Explain Color Appearance Shifts Induced by Large Uniform Backgrounds in Dichoptic Matching, *Vision Research*, **35** (2), 239-254.

Clark, A. (1996). True theories, false colors. *Philosophy of Science*, **63**, (Proceedings), S143-S150.

Finlayson, G., D., Morovic P., M. (2000a). Crossover Wavelengths of Natural Metamers. In: *Color Image Science 2000*, April 10-12, 2000. Color Imaging Institute, University of Derby

Finlayson, G., D., Morovic P., M. (2000b). Metamer Crossovers of Infinite Metamer Sets, Proceedings Volume of the 8th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications, Scottsdale, Arizona, Nov. 7-10, 2000.

Fodor, J., A. (1987). Why there still has to be a language of thought. In Jerry A. Fodor: *Psychosemantics*. Cambridge Mass. MIT Press, 135-154.

Gilchrist, A., Annan, V. (2002). Articulation effects in lightness: Historical background and theoretical implications. *Perception*, **31** (2), 141-150.

Hardin, C., L. (1988). *Color for Philosophers: Unweaving the Rainbow*. Indianapolis, MA: Hackett.

Harman, G. (1990). The intrinsic quality of experience. Philosophical Perspectives, **4**, 31-52.

Haugeland, J. (1981). Semantic Engines: An Introduction to Mind Design, In: John Haugeland (Ed.): *Mind Design*, Cambridge Mass. The MIT Press, 1-34.

Jakab, Z. (2003a). For a truly colored world. (Review of M. Tye: Consciousness, Color and Content). *Color Research and Application*, forthcoming.

Jakab, Z. (2003b). Intrinsic colors – and what it is like to see them. In Mausfeld R., and Heyer, D. (Eds.): *Colour Vision: From Light to Object*. Oxford: Oxford University Press, forthcoming.

Jakab, Z., McLaughlin, B., P. (2003). Why Not Color Physicalism Without Color Absolutism? (Commentary on A. Byrne and D. Hilbert: Color Realism and Color Science). *Behavioral and Brain Sciences*, forthcoming.

Johnston, M. (1998). Are manifest qualities response-dependent? *The Monist*, **81**, 3-43.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press.

Kardos, L. (1934). “Ding und Schatten” [Thing and Shadow]. *Zeitschrift für Psychologie* Ergänzungband 23

Katona, G. (1929). Zur Analyse der Helligkeitskonstanz [On the analysis of lightness constancy]. *Psychologische Forschung*, **12**, 94-126.

Marr, D. (1982) *Vision. A computational investigation into the human representation and processing of visual information*. New York: Freeman

Matthen, M., Levy, E. (1984). Teleology, Error, and the Human Immune System. *Journal of Philosophy*, **81**, (7), 351-372.

McGinn, C. (1996). Another look at color. *Journal of Philosophy*, **93**, 537-553.

McLaughin, B., P. (2003a). The place of color in nature. In Mausfeld R., and Heyer, D. (Eds.): *Colour Vision: From Light to Object*. Oxford: Oxford University Press, forthcoming.

McLaughlin, B., P. (2003b). Color, consciousness, and color consciousness. In Quentin Smith and Aleksandar Jokic (Eds.): *New philosophical percpcetives.* Oxford: Oxford University Press, 97-152.

Peacocke, C. (1997). Colour Concepts and Colour Experience. In Alex Byrne and David R. Hilbert (Eds.), *Readings on Color*, Vol. 1. (The Philosophy of Color), Cambridge Mass. The MIT Press, 51-65.

Sekuler, R., Blake, R. (2002). *Perception*. McGraw-Hill.

Shepherd, A., J. (1999). Remodelling color contrast: implications for visual processing and color representation. Vision Research, 39, 1329-1345

Shoemaker, S. (1994). Phenomenal Character. *Noûs*, **28**, (1), 21-38

Smith, M. (1993). Color, Transparency, Mind-Independence. In John Haldane and Crispin Wright (Eds.): *Reality, Representation and Projection*. Oxford University Press

Stillings, N., A., Weisler, S., E., Chase, C., H., Feinstein, M., H., Garfield, J., L., & Rissland E., L. (1995). *Cognitive Science: An Introduction*. Cambridge Mass.: MIT Press

Stroud, B. (2000). *The Quest for Reality; Subjectivism and the Metaphysics of Color*. Oxford University Press, New York, Oxford.

Thompson, E., Palacios A., G., Varela F., J. (1992). Ways of Coloring: Comparative Color Vision as a Case Study for Cognitive Science, *Behavioral and Brain Sciences*, **15**, 1-74.

Yablo, S. (1995). Singling out properties. In: *Philosophical Perspectives*, Ed. J. Tomberlin. Vol. 9. Atascadero, CA, Ridgeview.

---

[1] Two remarks. First, Campbell (1997, pp. 178-179) uses the term ‘transparency’ to mean what I mean by revelation here. I follow McLaughlin (2003) in using ‘transparency’ and ‘revelation’.

Second, Tye (2000, chapter 2, esp. pp. 26-32; pp. 55-56, 149-150, note 4 on p. 167; Tye, 1995, pp. 169, 174) has a philosophically more attractive approach to this problem. He says that we should assume a relation of metaphysical identity between redness and a certain type of surface reflectance $S_R$. Then our concept SURFACE REFLECTANCE OF TYPE $S_R$ will reveal the essence of redness under a conceptual mode of presentation, and our visual experience of red will reveal the essence of redness under a perceptual mode of presentation. There is, however, not much of a transition between the two modes of presentation, since phenomenal concepts are conceptually irreducible – they are *simple recognitional concepts* (2000, pp. 26-29). This explains why congenitally blind or colorblind people can never learn what it is like to see the colors. The same assumptions presumably also explain why Tye thinks that red objects, in looking red, look disposed to reflect such-and-such percentages of the incoming light (p. 55), even though color experience alone would never teach us what reflectance type redness is (i.e., it would never endow us with the concept SURFACE REFLECTANCE OF TYPE $S_R$). As it will become clear below, I deny that red objects, in looking red, look *disposed to reflect such-and-such percentages of the incoming light* – because the experience of red is a simple representational state, and so it can no way represent such a complex state of affairs.

[2] The term ‘primitive revelation’ is my invention. In it, ‘primitive’ refers to primitivism, a philosophical theory of object color. The reason for this is that the idea of revelation that I just characterized is at the heart of this theory. For exposition and discussion of primitivism see Campbell, 1993; Smith, 1993; Yablo, 1995; McGinn, 1996; Johnston, 1998; Stroud, 2000.

[3] Though some would deny this: see in particular Campbell, 1997. See, however, Tye (2000, p. 149) for an objection against Campbell’s view.

[4] Not everybody would agree with this claim. Byrne and Hilbert (1997, 2003), and Tye (2000, Ch. 7) formulate a view of color that (1) attempts to be consistent with the findings of color science, (2) assumes that colors are the canonical causes of color experience, (3) claims that colors can be adequately captured by scientific concepts (4) argues that object colors crucially determine what it is like to see them, and so in this most important sense comes close to the idea of primitive revelation. However, McLaughlin (2003b),

and Jakab (2001, 2003a, 2003b, Jakab and McLaughlin, 2003) argue that both Tye's and Byrne and Hilbert's views fail right on the first count – they cannot be made consistent with findings about color perception. For more on (primitive) revelation, see McLaughlin (2003b, esp. sections 1,2,7,18).

[5] We have to take into account that metamerism is both illumination-dependent and observer-dependent. I'll address this problem abundantly below.

[6] For instance, the broad color category green includes many shades of green with differing lightness, saturation, and chromatic composition (i.e., only slightly yellowish or bluish greens we might still classify as green). Physically, the color category green fairly well corresponds to surfaces that reflect or emit or transmit light dominantly between 500 and 600 nm, and little light between 400-500 nm and 600-700 nm. Olive greens (i.e., dark, less satureted greens) mostly correspond to surfaces whose reflectance or emittance or transmittance is low overall, but still it is relatively higher in the middle of the visible spectrum. Narrow shades of green correspond to metameric sets of surface reflectances, transmittances, or relative energy distributions of light emission (Byrne and Hilbert, 1997, p. 266; Finlayson and Morovic, 2000a, 2000b), but each member of such a set has the more general feature characteristic of the broad category green.

[7] I do not mean to assert that the nature of three-dimensional space and shapes does not raise any philosophical problems. It probably does so at some more basic levels of metaphysical theorizing. But my point is that on reflection the nature of color raises immediate problems and fierce disagreement between different philosphers, a kind not characteristic of space and shape. I just offered a crude starting characterization of the nature of shapes that is, I think, fairly uncontroversial. However, analogous characterizations of color have faced serious disagreement.

[8] And the corresponding metaphysical question is what on earth the colors are if they are not the canonical causes of color experience.

[9] Again, the emphasis is on specific stimulus properties, not highly generalized ones. Substitute for P something like being rectangular (as opposed to having shape), or being unique red (as opposed to being colored). The principle should work at both levels of generality in order for conceptual revelation to obtain in a perceptual modality. However, as I argued in the main text, what really makes the difference between shape perception and color perception is the level of specific properties.

[10] That is, in terms that are *non-question-beggingly perceiver-independent*. Primitivists like John Campbell or Barry Stroud would claim that 'red' refers to a property whose essence does not include any reaction of perceivers, and for this reason it is 'looks red' that is to be explained by reference to redness, and not the other way round. Since this claim is not generally agreed upon (to put it mildly: see among others Hardin, 1988; Shoemaker, 1994; Peacocke, 1997; Block, 1999; McLaughlin, 2003b), 'red' does not appear to be a non-question-beggingly perceiver-independent term to characterize particular colors.

[11] Note that physicalist views of color that I accept (Jackson and Pargetter, 1997, 2000; McLaughlin, 2002, 2003) do not satisfy this requirement, nor is this their goal. For instance, on McLaughlin's approach, the perceiver-independent characterization of colors does rely on the findings of color science. Moreover, on that account, the essence of being a certain color (e.g., red) is not perceiver-independent. Being red is being the property that fills the redness role (i.e., evokes experiences of red in certain perceivers in certain circumstances), plus some other requirements, according to McLaughlin. Finally, McLaughlin explicitly denies that color perception is revelatory with respect to colors.

[12] The primitives need not literally be cylinders. Marr's point may be taken to be that shape primitives are simple representational structures whose content can be characterized by some generic shape and a few additional parameters. This implies, of course, that the shape primitives are not literally cylindrical (that would be absurd); they are properly thought of as *perceptual representations of* cylinders. Marr's idea was apparently that just as we can fairly well approximate many complex shapes by building them up out of cylindrical building blocks of different height and diameter, an analogous approximation can happen at the representational level – complex representations of complex shapes are composed of simple representations of some simple shapes, that are treated, for representational purposes, as basic.

[13] Sometimes we are inclined to squeeze utterly irregular shapes into simple schemes that approximate them quite badly. For instance, sometimes it is remarked that the borders of the United States (excluding Alaska and Hawaii) very poorly approximate a rectangle (or that the borders of France very poorly approximate a hexagon). For another example, think of humans' strong tendency to "see meaningful shapes into" irregular ones, as it happens when the Rorschach clinical test is administered to subjects.

[14] For instance, the CIELAB color space uses three rectangular dimensions: lightness (L*), redness-greenness (a*) and yelowness-blueness (b*), thereby making explicit the opponent organization and

unique-binary division of perceived color. Saturation is not an explicit dimension in this color space, but it is implicitly represented. Colors with low values of both a* and b* are unsaturated whereas high values of a* or b* (or both) correspond to higly saturated colors. A very simple mathematical transformation takes the L*a*b* color space into the L*C*h color space whose three dimensions correspond to lightness (L*), saturation (or chroma: C*), and hue (h). Hue is a polar (cylindrical) dimension in this color space. The L*C*h space does not explicitly represent the opponent organization and unique-binary division of perceived color. (see Wyszecki and Stiles, 2000, pp. 164-169 for these color spaces).

[15] When physically mixed, red and green yield either yellow (in additive mixing, like mixing the light of light sources) or achromatic gray (in subtractive mixing like superimposing color filters or mixing paints). Blue and yellow yield white or gray in additive mixing and green (often perceptually unique green) in subtractive mixing. See Hardin, 1988, and Fairchild, 1998, for more on the phenomenology of color perception.

[16] Color experiences can also be thought of as having achromatic grayness as constituents (and so the experience of an unsaturated unique red, or reddish gray, has two constituents: a red process and a gray (achromatic lightness) process. Grays can also be thought of as perceptual mixtures of black and white, the achromatic Hering primaries, but with this move we reach the maximum limit of compositionality in color experience. Alternatively, hue, lightness, and saturation might be thought of not as constituents, but as *dimensional positions* of color experiences (Jakab, 2000, pp. 336-341).

[17] I think such a compositionality pretty much follows from the opponent processing theory of color vision, that posits two “orthogonal” chromatic channels (red-green and blue-yellow).

[18] Constituent structure here is always understood relative to the processing system that accesses the representational states in question, that is, what interpretable parts this processing system would decipher in a given representational vehicle (Jakab, 2000, p. 336).

[19] Because simple constituent structures (in which the ordering or configuration of the constituents is not relevant) can pretty well be represented by vector spaces. Think of the water-alcohol mixture: the percentage of those two constituents can be represented in a two-dimensional vector space.

[20] Reflectance is a disposition, so, strictly speaking, it is the event of manifestation of this disposition (surfaces’ actually reflecting light) that causes our experience of color.

[21] Tye (2000, note 7 on pp. 167-168 considers a view according to which the color C of a surface S is some sort of a ratio, or difference, of S’s surface reflectance and the reflectance(s) in S’s surround. The observation that motivates such accounts of color is that surround effects (or *simultaneous contrast* effects) are ubiquitous in color perception. However, despite this fact, philosophers like Byrne, Hilbert, and Tye still strive for identifying colors with local reflectance properties of individual surfaces (see Tye, 2000, pp. 151-162). They do so because in their philosophical theory of color they want to honor the perception-based common-sense inutition that colors look to be local, intrinsic properties of surfaces. If colors look to be local properties, then that is what they are, they argue, since color perception had better be veridical, and not some sort of a pervasive illusion.

[22] With this move McLaughlin undertakes some risk, namely that if one day color science decisively shows that there is no surface property that is common to all surfaces that are disposed to look red (to perceivers of type P, in circumstances C) then redness does not exist. In other words, McLaughlin holds that colors cannot be identified with disjunctive properties (or disjunctions of properties), and he admits that currently it is an open question whether there exist physical properties that respond to his characterization of object color. If there turn out to be no such properties, then eliminativism about color has to be true (McLaughlin, 2002).

[23] For instance because the laws of nature in our world might have been different from what they actually are.

[24] And both of us can have perfectly normal color vision. This means that both of us passes the standard tests for color discrimination, since currently that is the only measure of the normality of one’s color perception.

[25] Not extremely rarely, such variation shows up in everyday verbal communication. Think of the debate between, say, husband and wife whether a particular fabric is predominantly green, or predominantly blue (Byrne and Hilbert, 1997, p. 272). Also, think of a customer’s consternation when a just-purchased expensive dress looks noticeably different in color in the street than it did in the store.

[26] One reason why we often do not notice changes in the perceived color of objects in everyday situations is that quite often these changes are temporal. We enter from the street into a store illuminated with

fluorescent tubes (an abrupt change); or we stay in the street while the sunlight gradually changes. In both these cases there is a slight change in our color perceptions due to the illuminant change. However, since we have little capacity to memorize exact shades (we can only memorize broader color categories like red, green, scarlet, teal, etc.: Raffman, 1995, pp. 294-295; Tye, 2000, pp. 11-13), such changes typically go unnoticed. In cases of spatial variation of illumination within one scene, color constancy obtains if we can discern the illuminant change and our color vision can estimate its magnitude (Maloney, 1999, 2002). Shadows are the most typical examples of this phenomenon. When half of a uniformly colored surface is in shadow whereas the other half is lit by direct sunlight, we do not attribute the apparent change in color to the object – even though in this case we can notice it. In cases where there is a spatial variation of illumination within a scene and we cannot discern it, color and lightness constancy tend to break down (Katona, 1929; Kardos, 1934; Gilchrist and Annan, 2002).

[27] Color scientists also introduced the notion of a standard observer. The standard observer is a theoretical construct: simply the average of the color-matching functions (and correspondingly, the cone sensitivity profiles) of a number of different color-normal subjects (see Hardin, 1988, pp. 76-82). Just as a vanishingly small number of people (if any) earn the average salary in any society, similarly, a vanishingly small number of color-normals (if any) will be standard observers. The advantage of the notion of a standard observer for color science is that most color-normals are fairly close to the standard observer in terms of color-matching, and this allows useful predictions for the industrial production of colorants and colored objects. But the emphasis is on color-matching: two subjects who make the same color matches may well differ in their color experience. Color-matching does not uniquely determine color experience (Wandell, 1995, Ch. 4).

[28] By ‘color’ I mean *narrow shade* in this context. The main problem for color absolutism is to secure specifically the narrow shades against the threat of relativization. I agree with McLaughlin (2002) and Hardin (1988) that this cannot be done in any remotely plausible way. It is no good as an argument to switch to broad color categories and note that riple lemons look yellow to virtually everyone (Stroud, 2000, pp. 173-174) since the maximum we can achieve by such a move is color absolutism about a small number of broader color categories or determinables, and the leftover relativism about the thousands of narrow shades that we actually perceive on particular occasions. Buy it if you like it.

[29] For instance, he says on p. 147: “Grass in the early morning looks to have the same color as it does at midday or late in the afternoon, even though the light is very different.” This claim is true if, by ‘color’, we mean *broad color category* like green, but false if ‘color’ means *narrow shade*.

[30] Blue and yellow are perceptually incompatible (no surface ever looks simultaneously bluish and yellowish all over), therefore if color perception veridically representes some physical properties that are the colors, then it should follow that blue and yellow as object colors are objectively incompatible. This incompatibility is also inherent in Byrne and Hilbert’s characterization of colors in terms of surface reflectances (1997, pp. 265-266; see also Tye, 2000, pp. 159-165).

[31] Using Sternheim and Boynton’s terms (Sternheim and Boynton, 1966; Byrne and Hilbert, 2003), if a particular surface looks to me 90 per cent greenish and 10 per cent yellowish, whereas it is, *objectively*, 100 per cent greenish (i.e., unique green), then I simply misidentified the narrow shade in question. I took it to be *the one* that is 90 percent greenish and 10 per cent yellowish (lightness and saturation specified) whereas it is, objectively, the one that is 100 per cent greenish or unique green.

[32] The opthalmological condition that brings about this sort of perceptual error is called *aniseikonia* (Byrne and Hilbert, 2003, section 3.4.).

[33] According to Byrne and Hilbert (2003, section 3.4.) another reason for the lack of an independent test is that we have no acceptable naturalistic theory of the content of color experience. Byrne (2002) and Byrne & Hilbert (2003) propose an account of color content to solve this problem.

[34] Or directions, like pigeons do (see Matthen, 1999).

[35] Davies’s examples are about visual perception of shape, and he does not go into details about how to generalize his argument to other perceptual modalities or perceivable aspects of the environment. Still it is clear that his goal is to establish externalism about experience in general.

[36] Tye (1995, 2000) uses a third criterion to distinguish perceptual content from belief content, namely that the former is *poised*. On Tye’s account, perceptual content attaches to the maplike (spatio-temporally organized) output patterns of sensory or perceptual modules, such that these contentful output patterns in turn stand in a position to influence the belief/desire system (Tye, 1995, p138; 2000, p. 62). In his externalist argument Davies does not make use of this criterion, but he mentions it (p. 311). He cites Evans

(1982, pp. 155, 158) who says that this is the feature that distinguishes unconscious non-conceptual states with content (e.g., states involved in an account of tacit knowledge of rules) from perceptual experiences. In his account of perceptual content, Dretske also notes that for sensory states to be experiences (i.e., for them to actually acquire phenomenal properties), the organism's cognitive machinery has to have a conceptual system on top of the perceptual (and behavioral) one (Dretske, 1995, pp. 19-20; note 17 on p172). Dretske, just like Tye, refers to Evans (1982, Ch. 7) who makes the same claim: in order for a sensory state to qualify as conscious experience, it has to be available, as input, for a conceptual processing system.

[37] For a detailed example see Davies, 1997, pp. 313-314.

[38] Of course, there is a sense in which inferential role depends on externalist content. For instance, the inferenctial role of symbols A and B can differ depending on their externalist content. The complex symbol 'A&B' means one thing if A covaries with bread and B with butter, and quite another if A tracks dogs and B tracks cats. The point is, however, that the internal computational structure and processes of the organism's brain impose constraints on what content its internal states can assume. One version of individualism or internalism about content is that such computational roles unambiguously determine the representational content of the system's internal states. Putnam (1981, Ch. 2 and Appendix **Reason Truth and History**, 1989**Rep&Real.**) offers a deep-running critique of this view. However, with respect to non-object-involving perceptual content, it appears more plausible. Putnam (1975 Meaning of 'Meaning') argues that though (linguistic) meaning is not in the head, certain factors or components of it are, notably perceptual stereotypes. Fodor's notion of narrow content (Fodor, 1987, Ch. 2, 1991; Davies, 1997, pp. 314-315) is a more detailed elaboration if this idea.

[39] Of course this misrepresentation does not matter for survival: color vision still endows us with a powerful means of surface discrimination.

[40] For support of this idea, and a reply to Peacocke's examples of allegedly nonrepresentational phenomenal differences that Davies cites in his paper, see Tye, 2000, Ch. 4. Even though I reject Tye's externalist approach to phenomenal character, I accept his idea that phenomenal character is representational in every detail.

[41] Tye (2000, Ch. 6) has a reply to Block's Inverted Earth argument. That reply shows that Block's version of the argument, in and of itself, cannot refute Tye's externalist claim about phenomenal character. But Tye's reply does not show that Block's position, as it is, is inconsistent. Furthermore, there might be a slightly modified version of the Inverted Earth argument against which Tye's preferred reply (Tye, 2000, pp. 136-140) does not apply. Here is a sketch of that version, admittedly inspired by Davies's paper.

Tye's preferred reply to Block's original Inverted Earth argument goes as follows. Some people, unbeknownst to them, are transferred from Earth to Inverted Earth, the transfer accompanied by implanting a color-inverting lens in the subjects' eye. In this case, the world continues to look phenomenally the same way to the subjects as it did on Earth (object colors are inverted on Inverted Earth, and the lens implant reinverts them), but the stimulus correlates of all color percepts are rearranged. The sky looks blue (due to the reinverting lens) but it is in fact yellow. Block claims that, according to Tye's account of perceptual content, this amounts to a change in the content of the subjects' color percepts. But this content change is not accompanied by any phenomenal change. Tye's preferred reply is that implanting the reinverting lens amounts to a tampering with the visual system, turning it into an abnormally or non-optimally functioning one, and so leaving unaffected the content-bestowing stimulus correlates, the ones that the subjects' color percepts *would* track, were they part of a normally functioning color-vision system.

Now imagine the following scenario. On Inverted Earth, if one looks at the sky, it is predominantly light between 500-700 nm ("yellow light") that hits one's retina (a difference in stimuli), but due to some nomological difference in the interaction of light and the retinal photopigments, 500-700 nm light predominantly stimulates (in our duplicates) what would be the S-cones in our eyes on Earth. Egg yolk on Inverted Earth reflects light predominantly in the 400-500 nm range ("blue light"), but due to the same sort of nomic difference, it stimulates, in our duplicates, what would be L- and M-cones in our eyes. (Cones and photopigments are identified by their biochemical composition). From the cones up in color processing, everything is physiologically the same in our duplicates as in us. Furthermore, this has also been part of the evolutionary history of our Inverted Earth duplicates. In this version there is no tampering, so Tye's preferred reply to the Inverted Earth argument seems not to work. Tye might still claim that, *in this modified case, the sky would look yellow to Inverted Earth denizens* (i.e., if the colors are inverted but there is a reinverting mechanism arising from nomological differences), just as it does in Block's version (where the colors are inverted and there is no reinverting mechanism – remember, only the Earthling visitors had reinverting lenses in Block's story), but this assertion seems to be in need of further argument. But unless Tye succeeds in supporting this idea, the modified Inverted Earth scenario apparently amounts to a phenomenal character inversion.