Color experience: empirical evidence against representational externalism

by
Zoltán Jakab, M.A.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

Department of Cognitive Science

at Carleton University
Ottawa, Ontario
September 17, 2001

Abstract

Contrary to some well-known views in cognitive science and the philosophy of mind, in general it is not the case that the felt character (phenomenal character, qualitative content) of sensory experiences is determined by the information that these experiences pick up, or represent, about the world. In this dissertation I shall focus on a particular sensory modality, namely color vision, to support this thesis.

Recently there has arisen a strong and popular view of phenomenal consciousness according to which the two fundamental problems about the mind: intentionality and phenomenal experience, can be traced back to just one: intentionality. On this view, the phenomenal aspect of experience is a special case of intentionality, or our mental states' carrying information about the external world. For instance, when we see the colors of objects, we see, in a direct and transparent way, exactly those kinds of properties that the external objects have. Not only are the colors of objects causally responsible for our experiences as of color, object colors crucially determine what is called the phenomenal character of color experience.

In this dissertation I shall argue that this view of color experience – the view called representational externalism – cannot be correct. I shall argue that from the empirical facts about object color and color vision we need not conclude that object colors do not exist, hence color vision is a grand illusion; however, we do have to conclude from these facts that though object colors are the causes of our color experience, what it is like to see the colors is not, in any theoretically interesting sense, determined by the colors themselves. To the contrary, what it is like to see the colors is crucially determined by how our color vision systems are constructed. In this dissertation I offer two independent arguments to support this claim.

Acknowledgements

I would like to express my very special thanks to my thesis supervisor Andy Brook who, during the last five years, shaped my work and my thinking in countless ways. I learned from him a lot, and he is largely responsible for making me a more mature cognitive scientist.

I would also like to thank the other members of my committee, John Logan, and Rob Stainton, and my external examiner Brian McLaughlin, for their enthusiastic support of my work, and for their critical observations that improved the quality of this dissertation so much.

This thesis would not be what it is without Rejean Baribeau giving me an opportunity to acquire the basics of color science in a hands-on way, and carrying out the key experiment in it. This allowed me to turn my initial ideas into a serious project. I would also like to say thank you to Jessica Cox for helping me with my experiment.

I am indebted to Andrew Bailey, Liam Dempsey, Eric Little, Jillian McIntosh, Dan Ryder, and William Seager for thought-provoking discussions of my views on color and color experience, and to Michael Tye for giving me valuable feedback on a first-paper version of this dissertation.

Finally, I am most grateful to my wife Anikó Lévay for devotedly supporting me in my studies.

Contents

A note on chapters and sections: chapters two to five constitute what is called below the first argument. Since these chapters belong together strongly, they are all numbered as section 2.

Chapter One: The concepts and the problem

0. Introduction: shape and color

The philosophical problems about color and color vision pretty well concentrate around the following question: what is color in objects? Alternatively, one might ask: are external objects colored? Or simply: what do we mean when we talk about color? To approach these problems, let me start by looking at some contrasts between color (and color perception) and shape (and shape perception).

The first observation one might make is that, it seems, there is no parallel problem for the notion of shape and shape perception. When we ask: 'What are the shapes of objects?' we can reply: well, 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.

A second observation might be the following. Shape perception plus intellectual reflection seems sufficient to form a conception of shapes that does not make reference to our perceptual experience of shapes. I have just given such a characterization, albeit in a crude form (I will add a little more detail to it below). Ancient Greeks obviously possessed a similar notion of shapes, and all they had as a means for acquiring it was shape perception and intellectual reflection.

The contrast with object color may well be obvious at this point. It seems that in order to form a conception of object color that does not make reference to our perceptual experience of color, we do need empirical science. Science has already taught us a lot about object color: for instance, we have learned that one crucial factor of color in most ordinary physical objects is their surface reflectance. However, we have also learned that in many cases surface reflectance does not figure as a key factor in object color. Light sources are the most obvious example of this. To learn such things, perception by the naked eye plus intellectual reflection are not enough. Ancient Greeks who did not have empirical science did not know these facts about object color.

Here is a third observation we might make, in the light of contemporary science. D. Marr's, I. Biederman's or S. Kosslyn's theories of visual perception (i.e., computational stories of how analog, maplike representations of spatial layouts are formed in the visual system) give us a plausible idea of how shape properties play an important role in determining the subjective (i.e., visual) quality of *shape appearances* – what it is like to see a particular shape. The visual system reconstructs contours, textures, surfaces, and out of some hierarchical set of primitive symbols (like symbol-filled arrays, shape primitives) it eventually builds up representations of complex spatial layouts. What it is like to see shapes is obviously crucially determined by such computational processes. For instance, what it is like to see a circular object is quite different from what it is like to touch one, even though the circular object in the two cases might be one and the same. But since in these two cases sensory information about the same shape is picked up via different proximal stimuli (incident light versus patterns of mechanical stimulation) and processed in radically different ways, the tactile and visual experience of the circular

experience will, understandably, be quite different. The other side of the coin is that, due to the systematic spatial mapping in vision, differences in shapes are readily picked up and are adequately represented. Moreover, and more importantly, the output of shape perception toward the rest of the cognitive system is rich in these details. Our conceptual representations have access to the details of analog representations of spatial layouts. For instance, we can produce a detailed verbal report on the contents of the visual buffer, say, in an imagination task.

Again, there is a contrast here with color perception. Even though color vision reliably tracks those stimulus properties that we call colors, this system does not give us nearly as much information about these stimulus properties as does shape perception about shapes. The information delivered by shape perception to the rest of the cognitive system about particular shapes is, it seems, immensely richer than the information delivered by color vision about particular colors. This may be one reason why color perception plus intellectual reflection alone could never lead us to the insight that object color is, in many cases, surface reflectance (or is quite closely related to surface reflectance).

Our knowledge of physics has suggested for a long time (perhaps ever since Democritus) that colors are not fundamental physical properties – they are not properties that figure in theories and explanations of physics. However, shapes aren't fundamental physical properties either, so in this respect there is no difference between color (a classical paradigm example of primary qualities) and shape (a classical paradigm example of secondary qualities). Another discovery about color that, I think, played an important role in facilitating recent philosophical debates over the nature of color is that

whereas shapes are physical types of some sort, object colors appear to be physically very heterogeneous.

For instance, what characterizes all and only spherical objects is that each and every one of them is a spatial distribution of matter that approximates sufficiently well, at some grain of spatial resolution, the geometrical notion of being a sphere. In other words, all (and only) spherical objects are physical realizations, or instantiations, of the abstract notion of a sphere. Of course there is a lot of vagueness here, and what counts as sufficient approximation is often, though not always, determined by what we would *see*, or accept on the basis of perception, as spherical. For beach balls the key criterion is perceptual; however, for ball bearings in high-precision machines, the criterion is much stricter.[1] Of course, any particular piece of matter has its own spatial distribution, hence its own (often irregular) shape, and this shape will vary depending on the spatial resolution at which it is characterized. Pizzas are round at some coarse-grain resolution (e.g., when viewed from two meters of distance), but they are not round at more fine-grained resolutions (e.g., when viewed from a smaller distance). So, given some coarse-grain spatial resolution R_C, and a finer-grain one R_F, *being circular at R_C* (i.e., approximating the circular shape sufficiently well at resolution R_C) is satisfied by some objects that are not circular at resolution R_F, and all objects that are circular at resolution R_F. In sum, shapes are types of spatial distribution; in other words, they are high-level physical types.

In contrast with shapes, colors, or the stimulus properties that are causally responsible for our color perceptions, seem, on evidence, not to be high-level physical types at all. When we examine the whole variety of objects and surfaces that look red to

us in ordinary circumstances of perception, no high-level physical type – no physical property common to all and only red objects – pops out. There are red reflecting surfaces, red transparent volumes, red light sources, red fluorescent objects, red phosphorescent objects, and so on. Even though, for our perception, these objects appear to have a striking distinctive and common property (i.e., redness as we perceive it), it seems that in terms of determinate physical properties, or high-level physical types, we cannot find any unique physical correlate for this perceptual attribute. This finding raises important problems both about the nature of object color and that of our perceptual experience of color. These are the problems that I shall discuss in this dissertation. I begin with a concise overview of the most important philosophical theories of object color. This will be followed by two independent arguments that attempt to support the view that (1) colors are not high-level physical types, that is, there are no unique correlates, in terms of high-level physical types, of our experiences of color, (2) what it is like to see the colors is crucially determined by how our visual systems operate, and not in any theoretically interesting sense determined by the environmental stimulus properties that are the systematic causes of our color experience. This view is consistent with the broader philosophical view that many call color realism; it is even consistent with the more specific claim that colors are causally effective physical properties of objects and surfaces. Simply from discovering that the stimulus properties that normally cause our color experience are heterogeneous *at any level of description* we need not conclude that objects are not colored. We may also conclude that object colors exist and are physically heterogeneous (see Section 1.4 for more detail on this point). This admission will slightly increase the theoretical distance between object color and color appearance in the sense

that the former does not significantly figure in the explanation of the latter. Object colors still cause our color experiences, but they do not determine the internal characteristics of those experiences – most importantly their *phenomenal character* (or what I take to be the same as phenomenal character, *what it is like to see the colors*).[2] I think this increased theoretical distance between color and color experience also follows from the observation that different species with color vision different from ours see the colors (i.e., largely the same colors as we do, since we pretty much share our environment with them) in significantly different ways than we do, and this shows up in the ways they categorize colors in experiments (Thompson et al., 1992, 1995, 148-155; Matthen, 1999).[3] Abandoning a characteristically species-chauvinistic attitude to color vision also helps us to understand that object colors are one thing, and what it is like to see them is quite another. So much so that, as I will argue, the characteristics of color experience cannot be understood from the assumption that we perceptually represent the colors. We perceptually represent the colors, yes, but the fact that it is the colors that are perceptually represented does not amount to an explanation of how the phenomenal character of our color experience arises. To explain that, we need to turn to other mechanisms – mechanisms that are internal to our brains, unlike the relations upon which the aboutness of mental states supervenes.

1. Theories and concepts in the focus of this dissertation

1.1. Theories of object color

In this section I introduce five different and well known philosophical theories of object color: eliminativism, dispositionalism, disjunctive physicalism, type physicalism,

and the so-called Simple View of color. The last three of these views belong to the family of physicalist theories of color. Since a key concern of this dissertation is with physicalist theories of color, the introduction of such views will be somewhat more detailed than that of eliminativism and dispositionalism.

1.1.1. Eliminativism (Subjectivism)

Colors as we perceive them have certain relational and intrinsic attributes that we reasonably call essential ones. For instance, purple is perceptually more similar to red than it is to green. In addition, there seems to be strong reason to assume that this relational property is essential to being purple. Purple could not be perceptually more similar to green than it is to red, yet stay the same, namely purple – at least we have no idea what this could mean. Take a typical purple surface S: it looks more similar to any red surface than to green surfaces. If we changed these similarity relations of S making it look more similar to green surfaces than to red ones, then, it seems, we would necessarily change S's purple look. Similarly for certain, more intrinsic attributes of perceived colors: for instance, the color orange is called a binary one because it is, perceptually, a mixture of red and yellow. To the contrary, red is called a unique hue because it is, perceptually, not a mixture of different colors. Moreover, it seems, a surface could not stay orange and yet not look both reddish and yellowish to some extent. It seems to be essential to being orange that orange is, *perceptually*, a mixture of red and yellow.

Now, if we describe the object colors in perception-independent terms, that is, in terms of surface reflectances or relative energy distributions of emitted light, then we will find no network of systematic similarity relations between them that parallels the

similarities between perceived colors. For instance, a surface S1 that emits pure 577 nm light will look, to most trichromat perceivers, unique yellow (or a color quite close to unique yellow). Another surface S2 emitting pure 590 nm light will, to trichromat humans, look orange, that is, binary. However, pure 590 nm light is no more a mixture of other lights than is pure 577 nm light; the same applies to the relative energy distributions of lights emitted by S1 and S2. Even though orange is a perceptual mixture of two other hues whereas yellow is not, when described in perception-independent terms, the object colors yellow and orange do not exhibit any corresponding, systematic structural difference. For all we know about color vision, it seems to be an empirical fact that our color vision systematically distorts the measurable similarity relations that obtain between object color stimuli (in perception-independent terms like surface reflectance), and it is this distortion that results in the perceived similarity relations of the colors (Thompson et al. 1992; Thompson, 1995, pp. 122-133; Matthen, 1999, pp. 64-69, 76).

To summarize, colors as we perceive them exhibit a characteristic pattern of similarity relations (we call this property *unity*), and a difference between what we call unique and binary hues (this is called the *unique-binary distinction*), but in perception-independent terms, nothing in the objects' color properties parallels these perceptual patterns. In addition, as I mentioned in Section 0 above, the object colors, in perception-independent terms, appear not to be physical types of any sort – rather, each object color appears to be, on evidence, a collection of various different physical properties.[4]

Another observation is that one and the same surface will look different in color to the same observer in different circumstances; similarly, the same surface in the same circumstances of perception will still look different to different perceivers (i.e., a

trichromat human and a pigeon, or a "normal" trichromat and an anomalous trichromat human). That is, one might infer, there is a one-to-many mapping between color stimuli (the relevant surface properties) and perceived colors.

Given such observations, a number of philosophers have concluded that since (1) no unique stimulus properties correspond to our color percepts, and (2) the stimulus properties themselves that cause our color experiences do not exhibit any measurable similarity relations that parallel the (essential) similarities between perceived colors, there simply are no such things as object colors (Hardin, 1988, 1995; McGilvray, 1994; Maund, 1995; see also Matthen, 1999, Section 4). Objects of course look colored, but this is a grand illusion: perceived color is a product of our brains, mistakenly attributed to external objects by perception. Perhaps the case of color is somewhat similar to what happened when the notion of gravity arose in Galileo's time. Between roughly Aristotle and Galileo, it was thought that the reason why objects fall is that they have an intrinsic inclination to fall. When the notion of gravity took over, it turned out that, even though objects seem to have an intrinsic inclination to fall, there is, as a matter of fact, no such property. So much so that the discovery of gravity later changed the common sense view of free fall as well: today not even parents teach their children that objects fall because they have an intrinsic inclination to do so (rather, what today's parents say to their children is something like that the earth attracts small objects, that's why they fall – a clear application of the theory of gravity).

This view of color is called eliminativism, or subjectivism about color (Hardin, 1988; Boghossian and Velleman, 1997a; McGilvray, 1994). However, to many philosophers, eliminativism seems to be rather an implausible view. Perhaps the most

unattractive feature of it is the idea that color perception is a grand error, a ubiquitous misrepresentation right at the bottom of visual organization. It is by colors that we see almost all other attributes of objects (surfaces, shapes, and so on) – and now it turns out that there are no colors. What there are instead are only color experiences, and color experiences are primarily products, or states, of our brains. Still, in color vision, certain attributes of some of our brain states (i.e., the phenomenal characters of color experiences) are "projected onto" external objects that we perceive – properties of our brain states are perceived as properties of external objects. Some philosophers (Tye, 2000; Shoemaker, 1994) have suggested that it is quite difficult, if at all possible, to make sense of this view.

1.1.2. Dispositionalism

Given this controversy with regard to eliminativism, what could be an escape route – that is, a philosophical theory of color that accommodates the empirical findings yet does not end up concluding that colors do not exist? One possible way out is to observe that though object colors may be, at any level of physical organization, very heterogeneous, still, all and only those objects that look a particular color C to trichromat human observers in ordinary circumstances of perception have a common functional, or dispositional property. This property is *the disposition to look color C* (to trichromat observers, in ordinary circumstances of perception). In general, talk about dispositions is acceptable if there is a counterfactual link between some causally effective agent A, and some effect E. It is also desirable to have a causal explanatory story about how A causes E on particular occasions, and we have to be able to characterize the specific

circumstances in which A causes E to occur. If these conditions are satisfied, we can say that A has a disposition to produce E. These conditions are satisfied in the case of color. For instance, a Granny Smith apple would look lime green if a trichromat human looked at it in daylight; moreover, we do have an explanation of how this effect arises: the apple reflects light that affects our retina in a certain way, and so on. Therefore, we can reasonably say that Granny Smith apples have a disposition to look lime green. Now, since other, physically quite different objects that nevertheless look lime green (for instance, some area of a color TV monitor) have this disposition too, we can propose the following account of object color: the object color lime green is the disposition of physical objects and surfaces to look lime green (to trichromat perceivers, in ordinary circumstances of perception). *Mutatis mutandis* for other colors. By this move we can endorse what is called the Principle of Charity (Davidson, 1984, p. 27; Shoemaker, 1996, p. 98): since, as a matter of fact, we apply color predicates to physical objects (and not to sensations), we had better provide an account of the semantics of these predicates that preserves such color attributions as veridical.

This theory of color raises an immediate question: on this account, to be red is simply to look red; but how are we to analyze the phrase 'looks red'? To say, for instance, that 'looks red' means simply *evoking a color experience that represents its object as red* would be hopelessly circular. That is, if we explain *being red* by reference to *looking red*, then we cannot also explain *looking red* in terms of *being red* (or being represented as red). Such an account would be vacuous, because it would miss exactly the meaning of the term 'red'. As far as this version of the dispositionalist account goes, any perceivable property could stand in as the meaning of 'red'. For instance, to look circular

is, plausibly, to be visually represented as circular. Moreover, circular objects do have a disposition to look circular to us in ordinary circumstances of perception. So as far as the above schema goes, the term 'red' could just mean *circular* (Lewis, 1997).

However, we can save dispositionalism if we do not explicate the concept of looking red in terms of being red. That is, being red is still understood by making reference to looking red; but looking red is now understood as a sensory quality, or phenomenal experience, of some sort. For instance, we look at a stoplight, and say: 'to be red is simply to be disposed to look this way', where the indexical 'this' refers to the color experience we undergo, and are aware of, on looking at the stoplight. Science may help us to go beyond such indexicals in specifying types of color experience. The core idea in dispositionalism is that the concept of being colored is explicated in terms of phenomenal color experience. (For more discussion of dispositionalism along these lines see Peacocke, 1997; Johnston, 1997; Boghossian and Velleman, 1997a).

On more thorough scrutiny, dispositionalism has some more philosophical virtues; but it also has a counterintuitive consequence. As we saw, shapes are physical properties of objects that are causally effective. However, dispositions aren't causally effective, so if colors are dispositions, then colors are not the things that cause our experiences of color.[5] In general, it is the bases of dispositions that are causally effective. These bases can be either dispositional or non-dispositional properties. A piece of hot iron is disposed to burn my hand; the basis of the disposition here is the high temperature of the metal, and that is not a dispositional property.[6] On the other hand, surface reflectances, themselves dispositional properties, are regular causes of our color experiences. More exactly, it is the manifestation of reflectance – the actual physical

event of reflecting light – that causes our color experience, not the disposition of reflectance itself.

1.1.3. Physicalist theories

Physicalism about color in general is a family of views rather than a single view. Different physicalist theories of object color differ substantially from each other. Still, all such views hold that colors are physical properties of objects, and that particular instances of object colors are causally responsible for eliciting our color experiences. Different physicalist views differ with respect to what kind of physical properties object colors are, according to them. A key divide between physicalists, and the one I will focus on throughout this dissertation, is the following question. Are colors like red, purple, lime green, and so on physical types, kinds (even natural kind essences); in other words, are they universals of some relatively strong standing – universals that scientific realism would recognize? Or, alternatively, is the color red just a heterogeneous collection of widely different physical properties? In the rest of this section I introduce three different physicalist views of color that answer this question in different ways.

1.1.3.1. Disjunctive physicalism

We left off the discussion of dispositionalism by mentioning a problem, namely that dispositions are not causes, so if colors are dispositions, then colors are not the causes of our color experience. In order to circumvent this problem, and retain most or all philosophical advantages of dispositionalism, one could endorse the so-called disjunctive physicalist view of color (e.g., Jackson and Pargetter, 1997). On this view, colors are the

categorical bases of the dispositions to elicit color experience. These bases remain heterogeneous, or disjunctive (i.e., the color scarlet is a disjunction of a large number of quite different physical properties – surface reflectance types, light emission profiles, and so on). Moreover, disjunctive properties as such do not cause anything – but their disjuncts, i.e., particular instances of a particular object color, certainly do. On disjunctive physicalism, the disposition to elicit color experience is still essential for being a color of some sort; however, now object colors have a perception-independent characterization, in terms of ordinary, high-level physical properties.

Disjunctive physicalism gives many empirical findings their due: it readily acknowledges that one and the same stimulus in different circumstances looks different in color to different perceivers, and that one and the same stimulus in the same circumstances looks different in color to perceivers with different color vision systems. Jackson and Pargetter (1997, pp. 75-76; McLaughlin, 2001, pp. 25-26) accommodate these features by *relativizing* colors to perceivers and circumstances. Relativization means that color ascriptions in general include a perceiver and a specific circumstance of perception. There is no such property as *green, full stop*: what we should say instead is that, for instance, grass is green *for trichromat human subjects in daylight*. Illuminated by violet light, grass would look black to trichromat humans. So, after relativizing the colors we say that grass *is* black for trichromat humans in violet light. Similarly, we know from experiments that pigeons make a sharp color category border in the range of colors that we would classify as green (Thompson et al., 1995, pp. 148-155). So it is arguable that there are at least some objects that look green to us and that do not look the same way colorwise to pigeons (i.e., in the same circumstances). Therefore, grass in daylight is not

green in the same way for pigeons as it is for us, or so is the consequence of relativization.

Another result of relativization is this: for any object O in circumstance C, if O looks color Q to perceiver P, then O *is* Q for P in C – that is, O is *veridically* perceived by P as Q in C, no matter how unusual circumstance C is. After relativizing the colors, there is no such thing as a non-veridical color perception, or color illusion. Still, color hallucinations remain possible. Furthermore, there is another move to relieve the possible feelings of discomfort arising from the impossibility of color illusion. Even though different color perceptions of the same object in changing circumstances are all veridical, some of them may be more *typical* than the other, and this idea helps us to construct a secondary notion of *color simpliciter* (Cohen, 2000). For instance, the typical colors of objects for us are those that arise when these objects are illuminated by white light. The typical color of a steak is reddish brown, even though it can look to us, and hence be, orange if the illumination and other circumstances in the restaurant are unusual. In such a case we can say that the typical color of the steak is its color in white light, but it has untypical colors like orange in other circumstances (i.e., for us trichromat humans).

There is another, somewhat strange consequence of relativization, namely that if we relativize, there remain no empirical constraints on which surface property counts as what color (Hoffman, 2001, p. 76). Almost any surface property now *is* almost any color, if we fill out the relativization formula in the appropriate way. For instance, the uniformly high reflectance of a patch of snow, that many theorists are inclined to identify with its *white* color (see below), becomes the color orange in sunset. The snow is red for trichromat humans in red light, green for trichromat humans in green light, and so on.

The reply by the relativist to this observation could be that by relativizing we abandoned the "absolute" notion of color. Saying that almost any surface now *is* almost any color implies a subtle mistake, since we have already embraced the notion of relative color. And it is utterly false to say that almost any surface is almost any *relativized* color. A patch of snow is *orange for us in sunset*, but not *orange for us in noon daylight*. That is, in terms of relativized color there are a lot of exclusions in the case of any particular object.

Still, there are physicalists about color who reject the idea of relativized color and adhere to the notion of absolute color. I will discuss these views in the following two sections. The idea in the "absolutist" version of physicalism is this: there are optimal and non-optimal conditions for color perception. For us humans, illumination by white (or yellowish) light is part of optimal conditions. Furthermore, in optimal conditions the colors of objects play a key role in determining what it is like to see the colors. In other words, object colors help to *explain* why the experience of seeing particular colors is like it is. As two authors, Michael Tye (2000, pp. 54-60) and John Campbell (1993) put this point, object color properties are transparent to color perception (see the following two sections). For those who relativize the colors, the intuition of transparency is not in any way central. Relativists can and often do deny that the stimulus properties that are the systematic causes of our color experience would in any interesting sense shape, or determine, what it is like to see the colors (see McLaughlin, 2001, esp. pp. 35-41).

1.1.3.2. Type physicalism

Type physicalism about color is the view that object colors are high-level physical types, similar to shapes. On this view, all and only objects that look to us the same in color in normal circumstances of perception (that is, objects that *are* the same in color) have some non-disjunctive, causally effective property that is specifically causally responsible for our color perception of them in normal circumstances. This principle applies to broad color categories like red, and also to the object colors that correspond to perceptually maximally determinate shades like unique green of a particular saturation and lightness. As I said in the previous section, type physicalists reject the relativization of colors out of hand, and hold that which property of objects and surfaces is their redness does not depend on the circumstances of perception. Redness is one and the same surface property no matter what the circumstances – neither the circumstances, nor the type of perceiver enter color ascriptions. Type physicalists hold that in non-normal, or non-optimal circumstances of perception we misperceive the colors of objects – we undergo color illusions.

Colors on type physicalism may be perceiver-relative properties in the sense that different organisms with different visual systems pick out more or less different variants of the stimulus properties that are the colors. This can explain, for instance, the difference between the ways dichromat and trichromat humans see colors. Dichromats and trichromats see colors differently because their visual systems carve up the realm of color properties in different ways. However, if two organisms pick out exactly the same color properties, then they see the same objective colors, and this entails that *they see colors*

the same way. Such a view of color is proposed by Hilbert (1987) and Byrne and Hilbert (1997), and it is also implied in Tye, 2000 (see esp. 4.2., case 7, pp. 89-93).

To be sure, objects and surfaces with the same color are definitely heterogeneous in microphysical terms. For instance two surfaces that look the same shade of blue can be microphysically very different. However, in terms of surface reflectance, they are not so different. Statistically, the vast majority of actual blue reflecting surfaces has a characteristic type of surface spectral reflectance (SSR). *Mutatis mutandis* for at least the other broad color categories: red, green, yellow, orange, purple, yellowish green and bluish green. That is, the theory that object colors are high-level physical types can perhaps be maintained if we formulate it in terms of surface reflectance. Therefore the particular form that type physicalism currently takes is the reflectance theory. The reflectance theory identifies object colors with types of surface reflectance. This account enjoys a remarkable popularity (Hilbert, 1987; Byrne and Hilbert, 1997; Matthen, 1988, 2001; Dretske, 1995; Tye, 1995, 2000). On Hilbert's and Byrne's view (Hilbert, 1987; Byrne and Hilbert, 1997) maximally determinate colors are particular SSR profiles. We humans of course cannot see, that is, discriminate, all maximally determinate colors from one another, because we have only three different types of spectrally sensitive receptors in our retinae. The trichromat human retina acts like a fairly crude spectrophotometer (Hilbert, 1987, pp. 103-106). Still, the colors that we trichromat humans can perceive (colors for trichromat humans) are *types* of surface reflectance. Under each such type there belongs a set of particular SSR profiles. Tye (1995, 2000) does not mention the idea of maximally determinate colors as particular SSRs, but he too seems to hold that both the colors (i.e., color categories like red, purple, lime green, cadmium red, etc.) and the

narrowest shades that do not comprise any further (perceptually discriminable) shades are types of surface reflectance (Tye, 2000, pp. 89-93).

This proposal has faced an immediate objection by its opponents – an objection that proponents of the reflectance theory are well aware of (Hilbert, 1987, Ch. 5, Tye, 1995, pp. 146-147; Matthen, 1988, 2001). The objection cites metamerism, the phenomenon that surfaces with quite different SSRs can look the same in color in ordinary circumstances of perception (for instance, under the same daylight illuminant, or under a number of different illuminants). That is, even in terms of surface reflectance, redness is a widely heterogeneous, or disjunctive, property, the objector claims. Not so fast, the type physicalist replies. Take all those metameric reflectances that look to us the same in color, say, blue, or a particular shade of blue, say, b_{13}. It can be shown, type physicalists argue, that there is some surface reflectance type to which all and only blue (or b_{13}) objects belong. That is, contrary to first appearance, blue (or b_{13}) objects are not disjunctive in terms of reflectance. (This move is made by Hilbert, 1987, p. 111; Matthen, 1988, pp. 24-25, 2001; Byrne and Hilbert, 1997, pp. 265-266; Tye, 1995, pp. 146-147; 2000, pp. 159-161). *Metamer sets* are sets of different surface reflectances that all look to us the same in color under normal illumination (and against the same background).

In what follows, I will take it that the type physicalist proposal has the following implications. Redness (*mutatis mutandis* for other colors and narrow shades) is a stimulus property that is (1) non-disjunctive (Hilbert, 1987, pp. 110-111; Tye, 2000, pp, 149-150, note 4 on p. 167), (2) causally effective (Tye, 2000, pp. 148-149) (3) characterizes *all and only* objects that look to us red in normal circumstances of perception (Byrne and Hilbert, 1997, p. 265), (4) specifically causally responsible for our red sensations (Tye, 2000, pp.

148-149). Furthermore, (5) redness is an *inherent*[7] property of the distal objects of perception: surfaces and volumes (Tye, 2000, pp. 147, 153). Color is not a property of the proximal stimulus of color perception, namely the light that passes between the objects perceived and our eyes. This is a very plausible idea since (a) light is itself invisible, only the objects that interact with incident light are visible (Hilbert, 1987, p. 133), (b) it is *the objects* that look to us colored – color looks to be an inherent attribute of objects, just like shape (see Tye, 2000, p. 153).[8] Furthermore, due to our perceptual color constancy, objects look to us to retain their color despite considerable changes in illumination (Hilbert, 1987, pp. 61-65; McCann et al., 1976; Land, 1977). This observation gave rise to the idea that object colors are illumination-independent, and in this sense, invariant properties of surfaces. It also motivated the rejection of the wavelength conception of color (Hilbert, 1987, p. 64). The illumination-invariance of object color is a key claim of the reflectance theory: color perception reveals surface properties that are not directly dependent on illumination – i.e., they do not immediately change as illumination changes. Instead of adding a new condition on object colors, in what follows I will understand 'inherent' as implying illumination-independence.

To these conditions I will add, in Section 2.2.4.1, the following: (6) colors are *perceiver-independent* properties – properties that continue to be physically instantiated in worlds where perceivers who can pick them up are not instantiated. This characterization of object colors in type physicalism is somewhat redundant since, as it will become clear in Section 2.2.4, perceiver-dependent properties cannot be inherent properties of surfaces or volumes. In what follows I will sometimes call properties of objects and volumes (i.e., those of distal objects of color perception) that satisfy

conditions (2) to (6) C-properties, and ask the question whether, on the evidence, the C-properties turn out to be disjunctive or non-disjunctive. I will also take it that being a non-disjunctive C-property suffices for being a natural kind essence of some sort – at least some vague, "anthropocentric", derivative natural kind essence (Hilbert, 1987, 13-5; 115; 119-120; Tye, 2000, pp. 124-125, 159-161). That is, if redness is a non-disjunctive C-property, then red objects, by virtue of being red, belong to one and the same (anthropocentric) natural kind.[9]

The reflectance theory has it that it is certain types of surface reflectance that satisfy this general description. This deserves a few introductory remarks. First question, *exactly what type* of surface reflectance is the color red – or green, blue, yellow, and so on? I will give an abundant treatment to this issue below. For now, here is a crude characterization: redness is a surface reflectance that is a long-pass cutoff filter between 600 and 650 nm. A long-pass cutoff filter between 600 and 650 nm is a reflectance profile that is quite low (i.e., less than 6 per cent) between 400 and 600 nm, rises sharply somewhere between 600 and 650 nm and stays high until 700 nm. Though not completely general, this crude characterization already captures a great number of metamers of different shades of red. Statistically, the vast majority of natural and artificial red reflecting surfaces satisfies this characterization.

Second question: colors are reflectances. Reflectances in turn are dispositional properties. On most philosophical views of dispositions, dispositions themselves are not causally effective properties. Then how can colors be the causes of our color experience? Reply: the manifestations of dispositions, which are typically physical events, are causally effective. Therefore, even though colors are dispositions (i.e., of surfaces to

reflect light), they are causally effective in a derivative sense, namely that their manifestations (actual events of reflecting light) are causally effective. Moreover, events of reflecting light are the ones that are specifically causally responsible for our color perceptions. I will give a more detailed characterization of dispositions later.

Third, one might ask, how about non-reflecting color stimuli like transparent volumes, colored films, or active color TV screens? To accommodate these cases, the reflectance theory has to be extended somewhat. I shall discuss this problem at length in what follows.

1.1.3.3. Campbell's Simple View

John Campbell (1993) proposed a view of color that, according to him, is not necessarily (or need not be) physicalist, nevertheless it seems quite close to physicalism (see Smith, 1993, and Tye, 2000, p. 149 for discussion). On Campbell's approach, (1) object colors are the grounds (bases) of the dispositions to elicit color experience, (2) they are not (need not be?) physical properties, nevertheless they are mind-independent. Moreover, (3) facts about colors are supervenient on the microphysical facts in the sense that two possible worlds that share all their physical characteristics cannot be differently colored (Campbell, 1993, p. 258). Next, (4) object color is transparent to us, that is, color vision is enough for us to know which property redness is, for example (Campbell, 1993, p. 265). Transparency, as Campbell conceives it, implies a straightforward link of determination: on this view, object color determines the phenomenal character of color experience. As Campbell puts it, the qualitative character of a color experience is inherited from the qualitative character of the color (Campbell, 1993, p. 268). Smith

(1993, p. 272) takes transparency to imply that ordinary perception is supposed to be enough to reveal everything there is to know about the nature of object color. Finally, (5) it is argued at length that this view allows for the idea that colors are the canonical causes of color experience.

Campbell gives us some suggestions of what kinds of properties object colors are, even though he is not explicit about the point. What he says (Campbell, 1993, p. 263; see also Smith, 1993, p. 271) suggests that object colors are akin to high-level physical types like shapes or sizes. Shapes and sizes supervene on microphysics; so do colors in Campbells' view. High-level physical types are causally effective and they can also be the grounds of dispositions to elicit color experience. Note that such a solution amounts to physicalism, despite Campbell's hesitation to categorize his view this way. Tye (2000, p. 167n3) also thinks that Campbell's view is correctly classified as physicalist.

Another interesting feature of the Simple View is its notion of mind-independence. Campbell criticizes the notion of mind-independence on which mind-independent properties are those that figure in an "absolute" or "objective" description of the world. There is general agreement that colors do not figure in any such description of the world, and this remains true in the Simple View as well. This view also recognizes the fact that to understand color ascriptions, one must have, or must have had, color experience. Campbell claims that there is no way of understanding what a particular color property is, other than undergoing color experience. On the other hand, the character of color properties is transparent to color experience (Campbell, 1993, pp. 258-259). This sounds like a powerful argument for the mind-dependence of color.

However, as Campbell argues, along these lines it can be proven that even particularity is mind-dependent. For in identifying particulars we inevitably have to make reference to spatio-temporal attributes that are relational and contingent. (The possibility of duplication prevents the identification of particulars merely by their "intrinsic" properties.) For physical particulars, there is a distinction between numerical and qualitative identity. Just the opposite for abstract objects: there spatio-temporal coordinates do not apply, and the sameness of all other relevant properties guarantees identity for abstracta. Now, in referring to particulars, Campbell argues, we use demonstratives, and the very use of demonstratives inevitably introduces the subject. The suggestion is that the location of the subject and that of the objects in her surrounding are interdefined: the subject's location is not absolute but rather, it is interpreted within the framework of surrounding objects; this framework also serves to identify the locations of parts of the surrounding. One might conclude from this that what makes a physical entity the particular it is, is (or includes) its relation to a mind (or a subject at least).[10] But this is clearly too much: particularity is *not* mind-dependent, so we need a better notion of mind-dependence that renders particulars mind-independent – and hopefully does so with colors as well. This notion has to work without assuming an "absolute" or "objective" way of identifying particulars.

Campbell's positive account of mind-independence goes like this. We have to appreciate the importance in our thinking of a view of perception, namely that perceptions are caused by a pair of factors. This pair consists of (a) the way things are in the environment, (b) suitable perceivers situated in suitable circumstances. Within this framework, the mind-independence of perceived objects can be understood thus. The two

factors (a) and (b) are independent of each other. The perception of particulars requires both (a) and (b), but the very existence of particulars requires only (a). Perhaps it is true that we can only individuate particulars via making reference to perceiving subjects,[11] but identification (description, etc.) is one thing and existence is quite another. The concession that there is no such thing as the description of the world from no point of view (i.e., that any understanding of the world inevitably includes a perspective of the cognizer) is consistent with the realist intuition that the way things are in the world is independent of how we cognize about the world. So it sounds reasonable to understand mind-independence on an ontological basis (i.e., that things are the way they are in the environment regardless of whether there is any perceiver around) rather than an epistemological basis (what figures in an "absolute" or "objective" description of the world). Note that this interpretation of Campbell's notion of mind-independence differs from the more critical stand adopted by Smith (1993, pp. 275-276).

However, there is a more problematic aspect of the Simple View. As we saw above, this view assumes that (1) color experience exhaustively reveals the character of color properties, whereas (2) there is no other way to reveal color properties. In addition, however, it is also assumed that color properties are mind-independent ones, not denizens of some irreducibly subjective, mental realm. This leads to a strange agnosticism about color. It turns out that, though colors are features of the world that stay in their place when all observers go, there is no independent conceptual grasp of object color properties themselves – in contrast with shapes, for instance (see Smith, 1993, pp. 272-273). We are told that colors are specifically causally responsible for our color experience, and

suggested that they are high-level physical types, still these types are transparent *only to color vision*, not to higher cognition.

Note that if we make the same claim about phenomenal color character, i.e., about the what-it-is-like-to-undergo-it aspect of color experience, it is not so implausible. For what it is like to see colors (undergo color experience) plausibly includes, as an essential component, a relation between certain states of the subject's visual system and the rest of her brain. Very crudely, the relation is that, say, firing pattern F_{496} (the one that specifically correlates with red sensations) *occurs in the appropriate way* in her brain, and this is a key element in making this state a phenomenal one *for her* – but not for anyone else. Moreover, no higher cognitive activity like concept formation *alone* (i.e., without perceptual support) could ever result in the occurrence of F_{496} in the visual system – that's just how our brains work (for more details of this view see Jakab, 1999, 2000). So why phenomenal color character is available to beings with color vision, but not to intelligent cognitive agents without it, can be understood. The trouble arises when we claim that certain entities or properties that exist mind-independently are transparent only to color vision, but not to higher cognition. *That* version of the claim is much more difficult to swallow, especially in light of the suggestion that the properties in question are high-level physical types on a par with shapes, direct current generators, and elephants. This is probably the Achilles's heel of the Simple View, and is criticized by both Smith (1993, pp. 272-273) and Tye (2000, p. 149). It seems to me that these two authors formulate a critique that is quite close to the one I have just given, and will extend in 2.5.2.4. below.

1.2. Some basic concepts of color science

In this section I introduce some basic concepts used to characterize color stimuli, illuminants, observer sensitivity, and the notions of color matching and metamerism. I am going to use these concepts in later sections. A more abundant introduction to this field is found in Wandell, 1995, Ch. 4, and MacAdam, 1997. However, this section itself contains a little more than is necessary to understand the later sections.

Surface spectral reflectance (SSR)

The spectral reflectance of a surface is its wavelength-dependent disposition to reflect a certain proportion of the incident light. The SSR of surfaces is a function whose domain is the 400-700 nm interval of electromagnetic radiation (visible light), and its range is the 0-1 (0-100 per cent) interval. SSR functions specify the proportion of incoming light that is reflected by a surface at any particular wavelength between 400 and 700 nm. Notation: $S(\lambda)$, where λ is wavelength.

Mathematically, any function with this domain (400-700) and range (0-1) is an SSR function. However, there are further limitations on what can be a *naturally occurring* SSR function. SSR functions are continuous, and they are smooth functions, with reflectance varying slowly with wavelength.

Spectral power distribution (SPD) functions.

SPD functions characterize the energy distribution of illuminants, or light emitted by a surface. They express the distribution of the total energy of the illuminant light over the wavelength spectrum. SPD curves are also typically continuous functions, varying

slowly over wavelength, though the SPDs of some typical light sources (like fluorescent tubes or color TV screens) show more abrupt variations. Notation: E(λ), λ is wavelength.

Color signal.

The color signal is the product of surface reflectance and illuminant SPD: it characterizes the light actually reflected by a particular surface under a particular illuminant. The color signal is the SPD of light coming from a reflecting surface to the perceiver's eye. Its unit of measure is the same as that of SPD: (relative) energy (of the illuminant) times a proportion (reflectance – a value between 0 and 1) gives an energy measure.

Color matching functions of the (standard) trichromat observer.

Three different functions that describe the sensitivities at different wavelengths of the three types of human retinal cones and the corresponding processing channels. The standard color-matching functions (CMFs) are obtained by averaging the individual color-matching functions of a large number of subjects. As a result, very few individuals have CMFs that are exactly at the average, and there is a significant between-subject variation in this respect. Notation: x(λ) is the spectral sensitivity function of the long-wavelength channel; y(λ) is the sensitivity function of the medium-wavelength channel; z(λ) is that of the short-wavelength channel. Color-matching functions in general are linear transforms of the spectral sensitivities of the three retinal cone types (Wandell, 1995, Ch. 4, esp. pp. 85-86, 95-96). The sensitivity curves of the cones are color-matching functions themselves, but there are other color-matching functions as well. The

CIE (Commission Internationale de l'Éclairage) standard color-matching functions represent one particular choice of the possible color-matching functions – a choice with some practical advantages (see Wandell, 1995 pp. 87-88).

Color-matching: some theoretical background

Color matching is the phenomenon that any SPD of light coming from an object to the eye can be perceptually matched by mixing just three separate wavelengths. In typical color matching experiments subjects see a vertically halved circular area, one side of which emits a test light with some specific SPD; the other half emits light which is a mixture of three specific wavelengths (primary lights). The task is to adjust the intensities of the primaries – by turning three knobs – so that the two halves of the circular area appear indistinguishable in color. With certain limitations, such a match can always be achieved.[12] Similarly, by adjusting the RGB system of a computer monitor, we can achieve a perceptual match between a reflecting surface (e.g., our T-shirt) as it appears in a certain illumination, and the color of a monitor area. In this case the color signal arising at the reflecting surface is perceptually matched by the SPD of light emitted by the screen.

Color matching is represented mathematically as a matrix transformation: a linear mapping between the test light spectral power distribution and the intensity of the three primary lights (Wandell, 1995, pp. 82-83). For instance, the SPD of the test light is characterized by an n-dimensional vector (a 1-column matrix); the matching mixture of the three primaries is characterized by three intensity values. Then there exists a 3 x n system matrix such that multiplying the test light SPD by the system matrix yields the

primary intensities required for the perceptual match. The rows of the system matrix are the color-matching functions. The existence of such a system matrix is an empirical fact: it can be found by different methods (all of which give equivalent results). One such method is to match monochromatic test lights by a triplet of primary wavelengths. Since the vector representing a monochromatic test light is zero at each entry except one, the product of the system matrix and the monochromatic test light vector equals a single column of the system matrix. Thus, by matching a series of unit-intensity monochromatic lights, the system matrix can be obtained.

Tristimulus values of color stimuli

The color-coordinates, or tristimulus values of a reflecting surface S are calculated as follows:

$$X = \int_{400}^{700} E(\lambda) * \mathbf{x}(\lambda) * S(\lambda) \, d\lambda$$

$$Y = \int_{400}^{700} E(\lambda) * \mathbf{y}(\lambda) * S(\lambda) \, d\lambda$$

$$Z = \int_{400}^{700} E(\lambda) * \mathbf{z}(\lambda) * S(\lambda) \, d\lambda$$

Sometimes these integrals are multiplied by a normalizing factor, K, where

$$K = \frac{100}{\int_{400}^{700} E(\lambda) * \mathbf{y}(\lambda) \, d\lambda}.$$

The point in this normalization is the following. For a perfect white diffuser (a perfectly white surface), whose reflectance $S(\lambda)$ is a constant function of wavelength, with all its values being 1 (or 100 %), the value of Y, normalized, will be 100. The Y color coordinate roughly corresponds to achromatic lightness[13], thus it is theoretically adequate to say that a perfectly white reflecting surface has a lightness value of 100. Mathematically, if $S(\lambda)$ is constant with each of its values being 1, then

$$\mathtt{E}(\lambda) * \mathbf{y}(\lambda) * S(\lambda) \equiv \mathtt{E}(\lambda) * \mathbf{y}(\lambda)$$

that is, $\mathtt{E}(\lambda) * \mathbf{y}(\lambda) * S(\lambda)$ and $\mathtt{E}(\lambda) * \mathbf{y}(\lambda)$ will be identical. That is, under this assumption,

$$\int_{400}^{700} \mathtt{E}(\lambda) * \mathbf{y}(\lambda) * S(\lambda) \; \mathrm{d}\lambda = \int_{400}^{700} \mathtt{E}(\lambda) * \mathbf{y}(\lambda) \; \mathrm{d}\lambda$$

will obtain. Hence, for perfectly white surfaces,

$$\mathtt{K} \; * \; \int_{400}^{700} \mathtt{E}(\lambda) * \mathbf{y}(\lambda) * S(\lambda) \; \mathrm{d}\lambda = 100.$$

It is obvious that the tristimulus values of a surface are dependent on the illumination, that is, on the $E(\lambda)$ function. The tristimulus values of emitting surfaces (light sources) are obtained simply by multiplying the SPD of emitted light by the color-matching functions. If emission and reflection combine (e.g., when an active color TV

monitor is illuminated by an external light source), then the SPDs of the two components (emitted and reflected) are added up and multiplied by the color matching functions.

Metamers

Take two surfaces S_1 and S_2, with different SSRs, $S_1(\lambda)$ and $S_2(\lambda)$ respectively. S_1 and S_2 are metameric (constitute a metameric pair) if and only if they have the same tristimulus values, that is, if

$$\int_{400}^{700} E(\lambda) * \mathbf{x}(\lambda) * S_1(\lambda) \; d\lambda = \int_{400}^{700} E(\lambda) * \mathbf{x}(\lambda) * S_2(\lambda) \; d\lambda$$

$$\int_{400}^{700} E(\lambda) * \mathbf{y}(\lambda) * S_1(\lambda) \; d\lambda = \int_{400}^{700} E(\lambda) * \mathbf{y}(\lambda) * S_2(\lambda) \; d\lambda$$

$$\int_{400}^{700} E(\lambda) * \mathbf{z}(\lambda) * S_1(\lambda) \; d\lambda = \int_{400}^{700} E(\lambda) * \mathbf{z}(\lambda) * S_2(\lambda) \; d\lambda$$

(The normalizing factor K would drop out at this stage, so using it or not does not make a difference for present purposes.)

A set Q of different $S(\lambda)$ functions is a *metamer set* if and only if each pair formed of Q's members constitutes a metameric pair. A *natural metamer set* W is a set with naturally occurring metameric reflectances – all members of W are metameric and all of them satisfy the above-specified features (continuous and slowly varying).

Metamerism is, by definition, illumination-dependent. Changing the illuminant causes the metamer sets to rearrange.

In the above-mentioned case where one adjusts a computer monitor to look the same in color as one's T-shirt, the perceptual match is achieved at the point where the tristimulus values of the T-shirt (in the given illuminant) are equal to the tristimulus values arising from multiplying the SPD of light emitted by the screen by the CMFs.

A simplified method of calculation. In practice, surface spectral reflectance is most often measured at 10 nm intervals between 400 and 700 nm. That means that an S(λ) function is specified by 31 discrete values (i.e., as a 31-dimendional vector). As a consequence, integration in the key calculation method becomes a discrete summation:

$$\sum_{i=1}^{31} E(i)*S(i)*\mathbf{x}(i) = X$$

$$\sum_{i=1}^{31} E(i)*S(i)*\mathbf{y}(i) = Y$$

$$\sum_{i=1}^{31} E(i)*S(i)*\mathbf{z}(i) = Z$$

Chromaticity coordinates, color coordinates

There are a number of different transformations of the tristimulus values that are important for color science. These transformations correspond to different variants of color space. Some of these variants have, as their dimensions, hue, saturation, and

lightness, whereas others do not. Of the tristimulus values, X roughly corresponds to the redness-greenness dimension, and Z to the yellowness-blueness dimension (Kuehni, 2000, p. 56), and the Y value corresponds closely with the perceptual lightness dimension (Wandell, 1995, p. 87).

One of the widely used transformations is the calculation method for obtaining the so-called *chromaticity coordinates*, x, y, and z. (The notation for tristimulus values is capital X, Y, and Z, that for chromaticity coordinates is lowercase x, y, and z.) The chromaticity coordinates are obtains as follows:

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

Since only two of the chromaticity coordinates are independent of each other (i.e., x+y+z=1), the chromaticity coordinates do not constitute an appropriate three-dimensional color space. Instead, color scientists widely use the Yxy color space whose dimensions are the Y of the tristimulus values plus x and y of the chromaticity coordinates. These dimensions retain all the information contained by the tristimulus values. (Whereas from x, y and z one cannot retrieve the tristimulus values.) The x and y dimensions do not correspond to the perceptual dimensions of hue and saturation, but there are other color spaces (transformations of the Yxy color space) whose dimensions correspond to the perceptual attributes of hue, saturation and lightness. Regarding the

Yxy color space, the achromatic and unsaturated colors are found in its center, and highly saturated colors are found close to its periphery. The chromaticity coordinates of monochromatic lights (the most saturated chromatic stimuli) are found right at the boundary of the Yxy color space. (For more on the Yxy color space see Section 2.4.2.2 and Figure 2 below.) In order to characterize approximate hue and saturation dimensions within the Yxy color space, the method of calculating *dominant wavelength* and *excitation purity* is used (Wyszecki and Stiles, 1967, pp. 321-333; MacAdam, 1997, pp. 53-59). Very roughly, the dominant wavelength of a color stimulus K is the wavelength of visible light that, when seen in pure form emitted by a surface, matches *the hue* of K (i.e., has the same chromatic shade, but not necessarily the same saturation and lightness, as K). The purple colors that have no chromatic match within the range of pure wavelengths of visible light are characterized by *complementary wavelengths*. The complementary wavelength of a purple color P is the pure wavelength that, when mixed with P in a certain proportion, cancels the purple look of P turning it into achromatic gray. Excitation purity is an approximate measure of saturation. Given an achromatic reference point in the center of the Yxy space (e.g., the chromaticity of noon daylight), the chromaticity point of a color stimulus K, and the chromaticity point of K's dominant wavelength, the method of calculation assures that these three points fall on a straight line, with K's chromaticity point being between the other two. The closer K's chromaticity point is to the periphery (i.e., to the chromaticity point of K's dominant wavelength), the higher K's excitation purity is, that is, the more saturated K is.

1.3. Representational theories of phenomenal character

According to philosophical tradition, there are two distinct aspects of the mind that are especially puzzling. The first is intentionality: the idea that mental states are somehow "about" states of affairs in the external world. To our thoughts and perceptions there correspond specific neurological events in our brain. At least for materialists, psychological events are to be explained in terms of such neurological events: for instance, their physiological, biochemical, or abstract computational properties. A key question is: how can such states stand for, or represent, entities, or states of affairs in the environment? It is now widely agreed that this representational capacity of states of the mind/brain can only be understood in terms of certain *relations* between the relevant brain states and the environmental items that they represent. Intrinsic (biochemical, computational, etc.) properties of mental (brain) states cannot alone explain their representational capacity.[14] In order for a mental state Q to become the representation of chipmunks, it is crucial for Q to acquire some sort of causal relation to chipmunks (e.g., be reliably activated by the occurrence of chipmunks in the subject's visual field). In general, the *representational content* of mental (brain) states – the information their occurrence carries about the environment – arises from such causal relations between states of the mind/brain and those of the environment.

The second deep problem of minds is their phenomenal aspect: for many psychological states (perceptions, emotions, perhaps thoughts) there is something it is like to undergo them. All conscious perceptions, for instance, come with such phenomenal character. In seeing, hearing, or tasting we undergo perceptual experiences that we can identify (discriminate from other experiences, or recognize) on the basis of what it is like to have them.[15]

Philosophical tradition has it that phenomenal properties are not properly explained by representational (intentional) ones – for there are sensory experiences with phenomenal character that do not represent anything (pain is the paradigm example of such an experience). This tradition has recently been attacked by a number of theorists (Dretske, 1995; Tye, 1995, 2000; W. Lycan, 1996; M. Matthen, 1988) who contend that not only does every single phenomenal experience have representational content, representational content straightforwardly determines the phenomenal character of sensory experiences. In other words, what has seemed as two distinct puzzles about the mind is, in reality, just one: if we thoroughly understand the different levels of intentionality in human minds, we thereby understand how our minds exhibit phenomenal consciousness. There are different representational accounts of phenomenal character. In this dissertation my primary focus will be F. Dretske's and M. Tye's recent theories (Dretske, 1995; Tye, 1995, 2000). Both these accounts address in detail the problems that color experience raises for externalism about phenomenal character. What I say below about Dretske's and Tye's views of phenomenal character I think generalizes relatively easily to other phenomenal externalist views (e.g., Lycan, 1996; Ross, 2000a, 2000b). In the rest of this section I introduce Dretske's and Tye's account of phenomenal character in general and color experience in particular.

1.3.1. Tye's account

Tye (1995, 2000) presents a comprehensive theory of phenomenal consciousness that offers an account of basically all aspects and well-known problems of conscious experience. Tye's theory accommodates a great number of empirical findings related to consciousness. He claims that phenomenal character in general is identical with *a certain sort of* representational content. To support this claim Tye first argues that, contrary to philosophical orthodoxy, *every* kind of sensory or perceptual experience has representational content (Tye, 1995, Ch. 4). Pains are no exception: they are sensory representations of bodily damage or disturbance (Tye, 1995, p.113). In his book Tye endorses a covariation theory of representational content (Tye, 1995, pp. 100-101; 2000, pp. 60-66, 118-122), though he sometimes includes evolutionary history as well, as a possible mediator of content (Tye, 1995, p. 153; 2000, p. 56). For him, a particular kind of pain *lawfully (i.e., predictably, under normal circumstances) covaries with*, *therefore represents that,* there is a such-and such disturbance present at such-and-such a bodily location.[16] Another key example is color experience. The activations of different physiological states of the color-vision system covary with different types of surface reflectance, thereby informing the organism about the presence of such types. This information-bearing relation provides for the representational content of the states of the color vision system. These contents have, as their crucial element, the corresponding reflectance types – they are contents *that* such-and-such a reflectance type is present. These contents are then identified with the phenomenal characters of color experiences. Since the relevant stimulus properties enter the contents of color experiences (and become the key element of these contents), object colors *determine* the representational content, or what is the same, the phenomenal character, of color experiences. This,

according to the theory, makes object colors *transparent* to us: what it is like to see the colors is crucially determined by what the colors themselves are like. We just see *those* properties – the colors – directly, and that explains why the colors look to us, perceptually, the way they do (Tye, 2000, Ch. 3; pp. 54-60). This notion of transparency is very similar to Campbell's one (see Section 1.1.3.3. above), except for the idea that object colors now have an independent characterization in terms of surface reflectance. Tye (2000, Section 3.3) formulates his transparency claim by denying that the non-epistemic 'looks' context is hyperintensional.[17] As he argues, if redness is surface reflectance of type R, and a particular surface looks (non-epistemically) red to someone, then that surface does look (non-epistemically), to that person, a surface reflectance of type R (Tye, 2000, p. 55). He defends this admittedly strange claim at length without mentioning the difference between color and shape perception that I introduced in Section 0 above. In this dissertation I shall not discuss Tye's notion of transparency any further, even though I think it is a problematic one.[18]

A further complication that needs to be addressed is that not every kind of representational content is, at the same time, some phenomenal character. Belief content, for instance, is not phenomenal. Phenomenal character is *Poised Abstract Nonconceptual Intentional Content* (that is, PANIC). The term 'poised' means that this sort of 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). The term 'abstract' (Tye, 1995, p. 138; 2000, p. 62) means essentially the same as 'not object-involving' (Davies, 1997, 310; 313-314): numerical identity of the objects does not play a role in the

identity 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 (hence the phenomenal character) they give rise to. This feature is obviously not true of belief content: quantitative identity of the object of belief plays a role in determining belief content. 'Nonconceptual' (Tye, 1995, p. 139; 2000, p. 62; see also Davies, 1997, 310-311) means that the properties that enter into these contents need not be such that the subject possesses matching concepts for them. That is, perceptual states carrying PANICs do not, all by themselves, constitute concepts.

The relation between PANIC and phenomenal character is that of metaphysical necessity. Being PANIC is not a contingent, superficial attribute of phenomenal character, but rather an essential one (Tye, 1995, Sections 7.1, 7.2, 7.3). Phenomenal character is not a multiply realizable abstract kind, one of whose realizations is PANIC. If Tye's theory is right then it is metaphysically necessary that phenomenal character *is* PANIC (just like water is H_2O) (Tye, 1995, p. 184; pp. 188-191). An important difference between the water-H_2O case and the phenomenal experience-PANIC one is that in the former case it is metaphysically possible that something with the superficial appearance of water is not H_2O whereas in the latter case there is no parallel possibility. Phenomenal character is an essential property of experiences, not a superficial one. Anything that feels like a pain *is* a pain (Chalmers, 1996, pp. 146-147; Tye, 1995, p. 188). In contrast, *being the waterish stuff in our environment* is a contingent property of H_2O. Therefore in the phenomenal character case the only relevant possibility that we can claim to *really* imagine (i.e., entertain as a possibility) is that the PANIC theory is wrong. If the PANIC theory is right, then both 'PANIC R' and 'phenomenal character red' are

rigid designators that pick out the same thing in every possible world. Hence there remains no way to imagine that, even though phenomenal character is PANIC in our world, it is something else in another possible world.

Tye's account of phenomenal character is externalist: as we saw, he claims that phenomenal character is one and the same thing as perceptual content. Since perceptual content is a relational property of perceptual states, so is their phenomenal character. On this account, stimulus properties like surface reflectances, bodily damages, the chemical properties of foods we taste and the like determine the perceptual content, hence the phenomenal character, of perceptual states they reliably elicit. So comes one key consequence of phenomenal externalism: stimulus properties (or the information represented about stimulus properties) play a key role in determining the phenomenal aspect of perceptual experience. That is, phenomenal characters are not intrinsic properties of our brains; they are not determined by the neurological or computational properties of our brains.

The antithesis of phenomenal externalism is phenomenal internalism: a general view, or category of views, under which a number of different theories of phenomenal experience belong. On this view, phenomenal characters are products of our brains, somewhat like firing patterns in the neural tissue. No doubt, sensory experiences with their phenomenal character are reliably elicited by environmental stimuli, but the phenomenal characters are *determined by*, in any theoretically interesting sense of the term, properties of the nervous system. This determination relation is sometimes expressed by the notion of supervenience. This version of the claim is that phenomenal color character supervenes on the internal constitution of the organism. A given color-

perceiving organism (or any of its molecule-by-molecule duplicates) would be capable of undergoing the same phenomenal color experiences (provided that the same physiological activity patterns occurred in its visual brain) in any possible world in which it is capable of biological functioning (Davies, 1997, pp. 312, 313, 323-324). That is, no matter how the environment in which the organism is found (hence the relevant, content-bestowing organism-environment relations) varied, given invariance in internal constitution, invariance in phenomenal color characters would be the result. As internalists contend, the perceptual similarity relations of object colors (unity) and the unique-binary distinction do not derive from those stimulus properties that cause our color experiences. In order to explain such attributes of perceived colors, we have to turn to the opponent-processing model of color perception. By the same coin, what the surface reflectance of ripe tomatoes is like has no interesting role in explaining what it is like to see ripe tomatoes – this is the internalist intuition. And this is precisely the view that Tye and other phenomenal externalists want to deny.

1.3.2. Dretske's account

Similarly to Tye, Dretske (1995) makes a distinction between sensory and conceptual representations. However, there are some differences between the two authors' general views of representation. Dretske, unlike Tye, contends that the notion of function, in addition to information, is a key to understanding the relation of representation.[19] In his view, what distinguishes meaning from mere information, and makes room for misrepresentation, is the notion of function (Dretske, 1995, pp. 3-4, 77). Function in turn is analyzed by Dretske mostly in terms of causal history: selection

history or, in the case of human artifacts, the history that captures what a system was designed for.[20]

Dretske distinguishes between *systemic* and *acquired representations* (Dretske, 1995, pp. 12-13). Sensory or perceptual states are systemic representations; the semantic content of these representations arises from their *systemic indicator function*. Systemic indicator function can be derived: for example, a speedometer is devised by humans to indicate speed, or carry information about speed, therefore this is its function. Living organisms have non-derived systemic indicator functions: this function arises as a product of evolution. Color vision was selected for carrying information about surface reflectance; therefore carrying information about surface reflectance is the non-derived systemic indicator function of color vision (Dretske, 1995, pp. 2-6, 11-15). An object or system can carry information about its environment without this being its systemic indicator function. For instance, the volume of a spoon reliably covaries with the temperature that obtains in its environment, hence the volume of the spoon carries information about temperature. However, the spoon does not have the function to carry information about temperature: it was not designed, not to mention selected, for carrying information about temperature.

Acquired representational function arises not from the properties of sensory systems or gauges, but rather, from additional contextual features that obtain in the current environment, or in the larger system of which the sensory system (gauge) is a part (Dretske, 1995, pp. 12-13). In the case of human artifacts like a gauge, the relevant inner states can be assigned representational functions quite independent of the systemic function. A typical example of this is calibration. For instance, if a speedometer estimates

the speed of the car on the basis of axle rotation, then for different tire sizes, the same needle position will indicate different speeds. In such a case the number scale on the number plate of the gauge has to be repainted (or a digital speedometer has to be readjusted) whenever the tires are replaced by new ones of a different size. In the case of natural systems like animals or humans, the paradigm source of acquired representational functions is learning (Dretske, 1995, pp. 14-15). In simple cases learning results in the associative linking of behavioral responses to sensory or perceptual states (one such example is conditioning). Instances of word learning by ostension, where perceptual categories representing types or classes of stimuli are associatively linked to auditory plus articulatory signals, constitute another example. On Dretske's view, experiences are states whose representational properties are systemic; thought and conceptual states are states whose representational properties are acquired. The representational content of experiences is fixed by the functions of the sensory systems of which they are states (Dretske, 1995, p. 15).

The attributes of Dretske's systemic representational states are similar to those by which Tye characterizes perceptual representations. Here is a list of the similarities.

(1) Dretske 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, par. 4) 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.

(2) Dretske also says that systemic representations are analog (Dretske, 1995, p. 172, note 16); this is not far from Tye's claim that perceptual representations are nonconceptual, or maplike (Tye, 1995, pp138-139).

(3) Dretske also assumes that systemic representations are not object-involving (Dretske, 1995, pp. 23-27; pp. 79-80). As he formulates the point, representations in general have a sense, and often a reference as well, but their sense does not determine their reference (Dretske, 1995, p. 23). (Reference in this context is the object whose properties are represented; sense is the property of the object that is indicated by the representation – the property by which the object is picked out.) Representations in general are such that they can only represent *properties* of whatever object their host system is connected to. For instance, a speedometer can represent the speed of the car in which it is installed; were it relocated in another car, it could represent the same speeds without giving a hint that now it is a different object the properties of which it is representing.

Dretske's identity thesis differs somewhat from Tye's. As he formulates his point (Dretske, 1995, p. 73): "In accordance with the Representational Thesis, I continue to identify qualia with phenomenal properties – those properties that (according to the thesis) an object is sensuously represented as having." In this formulation, the term 'qualia' refers to the same thing as Tye's term 'phenomenal character': both terms mean the what-it-is-like-to-undergo-it aspect of sensory experiences. The term 'phenomenal properties' refers to physical properties of distal physical objects – properties that are represented by sensory states. The thesis says that the phenomenal character of the sensory experience is one and the same thing as the stimulus property that is represented

by the sensory state. This view also seems to imply a straightforward determination (i.e., of phenomenal character by stimulus properties) and a strong sense of transparency.

Thus, Dretske's proposal is that qualia *are* external stimulus properties. On p. 83 he reinforces this view: "If you know what it is to be 18 °C, you know how the host feels to the parasite. You know what the parasite's experience [of temperature] is like as it 'senses' the host." On p. 84 he continues: "…(second fact) if things ever *are* the way they seem, it follows that **qualia**, the properties that define what it is like to have that experience, **are exactly the properties the object being perceived *has*** when the representation is veridical." (His italics; my boldface.) As he continues on the same page, the physical temperature of the host simply is the heat quale of the parasite's experience, whether or not the host actually has this property, that is, whether or not the parasite is perceiving the host veridically.

This view immediately raises some questions about the possibility of hallucination (and misrepresentation in sensation). Hallucinations are, by definition, phenomenal experiences in the absence of the corresponding stimuli. When I vividly hallucinate a big blue patch, the phenomenal character of my hallucinatory experience is distal blueness, the object color – but there is no blueness present because I am hallucinating. So it might occur to someone that in the absence of actual stimuli with perceivable property P we cannot undergo *phenomenal* experiences as of P, since the phenomenal character, that is, the stimulus property P, is absent. As a consequence, hallucination and misrepresentation in sensation are impossible, or so it might occur to someone. Dretske's answer: it is not particular physical instantiations of stimulus properties that constitute phenomenal characters, but rather, *property universals*.

Universals are properties (and kinds) in the metaphysical realist view: properties that exist in the world independently of how they are conceived of. On Dretske's view, for each sensory state type there exists exactly one stimulus property (i.e., universal) that that state indicates, hence represents. Such stimulus property universals partially constitute the representational content of sensory states.[21] The question about hallucinations occurs now in a different light. When we hallucinate, we undergo a phenomenal experience, but there is no corresponding, *instantiated* stimulus property (instantiated universal) present that causally affects our perception. However, in the case of hallucination, the stimulus property represented by the hallucinatory perceptual state *is still present as an uninstantiated universal* – a property P that nothing happens to have *in that particular instance* – but in many other instances, particular objects do have P. This uninstantiated universal enters the representational content of the sensory state, thereby providing for its phenomenal character.

1.4. The two lines of argument I will follow

Following Thompson (1995, pp. 122-133) we can say that there are two key problems that have to be solved in order to establish a credible *objectivist* account of color.[22] Correspondingly, there are two strategies of argumentation against such views. As Thompson puts it, the minimum requirement for objectivism is that the candidate physical properties for color be distal ones that the visual system can track or detect. The further requirement is that the well-known phenomena of color appearance (like unity, the unique-binary division, and the characterization of perceived color in terms of the three dimensions of hue, saturation, and brightness) should have a robust mapping onto

the color-candidate properties. This robust mapping is understood in a strong sense of isomorphism – that is, isomorphism with some further restrictions, not just isomorphism in the philosophically cheap sense in which everything is isomorphic with just anything else. For instance, the perceptual dimensions of hue, saturation, and brightness should correspond to measurable physical dimensions of the stimuli such that there is a strong *linear* correlation[23] between perceived hue, saturation and lightness on the one hand, and the corresponding measurable stimulus properties on the other. In addition, this linear correlation should correspond to an explanatorily relevant causal link: the measured hue property should be the one that is specifically causally responsible for our perceptions of hue – similarly for the other two dimensions. Such a linear equivalence of the dimensions would entail that the perceptual similarity relations of the colors that are expressed as Euclidean distances in color space are a linear transformation away from the measurable similarities of the color stimuli along the physical stimulus property dimensions that are the candidates for being recognized as objective hue, saturation, and lightness.

From the work of other authors it should be obvious that no such linear isomorphism obtains between stimulus properties like surface reflectance and perceived colors. Two strong arguments to this effect are found in Thompson, 1995 (pp. 122-133) and Matthen, 1999. I am not going to pursue this line of argument in the present dissertation.

The failure of isomorphism still leaves open the possibility that the first requirement of objectivism about color is satisfied: there are stimulus properties that color vision tracks or detects. Without further qualification, this constraint is obviously satisfied. In the rest of this dissertation I shall add some qualifications to this requirement.

I shall argue that the stimulus properties that correspond to perceptions as of a particular color (e.g., red) do not turn out to be physical types at any level of description. Redness, I shall argue, is a genuinely heterogeneous, or disjunctive, stimulus property. I shall discuss in detail two reasons to support this view. The first reason is that there is no physical type that unites different instances of redness like the redness of ripe tomatoes and that of hot iron (*mutatis mutandis* for other colors). The second reason lies in the individual differences in color perception: the fact that when different, equally normal trichromat human perceivers look at the same stimulus in the same circumstances of perception, they often see a slightly different color (Byrne and Hilbert, 1997, p. 272; Tye, 2000, pp. 89-93; Block, 1999, pp. 41-47; Kuehni, 2001, pp. 63, 65).

However, this admission need not lead us to abandon color realism (or color objectivism). We can just put on file the data that object colors are not physical types; we need not infer from this that there are not such things as object colors. To this different authors (e.g., Hardin, 1998; McGilvray, 1994) would reply that the failure of isomorphism plus the fact that color stimuli are not in any way physical types *together* constitute a strong reason to abandon color realism and subscribe to the view that colors exist only in the realm of perceptual experience. In response, Matthen (1999, pp. 64-69) argues that the perceptual color spaces of different species are probably very different (as suggested by discrimination data), and so it sounds like species chauvinism to hold that the properties that we trichromat humans perceive object colors as having (binary-unique division, unity, etc.) are *essential* to their being object colors.[24] The perceptual color space of every color-perceiving species dissects, *and distorts*, the corresponding stimulus property space in some systematic way (where 'distortion' means the application, in

perceptual processing, of some rather complex nonlinear transformation), still, what all these species distort in their own idiosyncratic ways are perfectly real stimulus properties. So why not call them colors? Well, because the nature of color stimuli is not revealed in color perception in a way in which the nature of shapes *is* revealed in shape perception, the subjectivist might reply (see Boghossian and Velleman, 1997b, for an argument along these lines). At this point intuitions might divide, some admitting subjectivist inclinations whereas others insisting on color realism. I do not see a compelling reason here to abandon color realism. I also think that this issue between color realists and color subjectivists is a genuinely philosophical one – it is a debate about what the most reasonable way of thinking and talking about object color is.

There is, however, another consequence of the empirical findings about object color. This is the one I shall be concerned with in the rest of this dissertation. I shall argue that our phenomenal experience as of color is not, in any theoretically interesting way, determined by the stimulus properties that are the standard causes of color experience, therefore phenomenal externalist views of color experience fail. In particular, I shall argue that (1) since object colors are not in any way physical types, we cannot assign representational content of the sort required by these views to color experience (I will call this line of thought **the first argument**), and (2) given whatever sort of representational content color experiences can have, the phenomenal character of color experiences varies independently of their content, hence externalist representational content and phenomenal character cannot be the same (this line of reasoning will be called **the second argument**). This forces us to endorse internalism about color experience: it is factors within the nervous system that crucially determine what it is like to see the colors. The internalism-

externalism debate I see as a hardcore scientific one in the sense that empirical data on color and color vision are directly relevant to this debate. While the available empirical data about color are compatible with at least some versions of color realism (e.g., disjunctive physicalism), they seem not compatible with the idea that the phenomenal character of color experiences is in any theoretically interesting sense determined by its distal causes, the color stimuli. Of course, an empirically sensitive philosophy of mind can do a lot of groundwork in clarifying the issues of the externalism-internalism debate about phenomenal character. This is what I attempt in the present work.

Chapter Two: Type physicalism about color and the first argument

2. First argument: Content, natural kinds, and phenomenal externalism

Here is the logical structure of the first argument that was introduced two pages ago. On the first level, we have:

[**P1**] If type physicalism about object color is wrong, then phenomenal externalism about color experience is wrong too.

[**P2**] Type physicalism about object color is wrong.

[**C1**] Phenomenal externalism about color experience is wrong.

I shall defend **P2** by analyzing existing versions of type physicalism and the problems they have. Then I will present a key problem that, to my knowledge, no existing type physicalist proposal addresses adequately – the problem of generalizing the reflectance theory of color to non-reflective color stimuli. I will then argue that to this problem no plausible solution can be found, because the case of non-reflective colors shows conclusively that object colors, characterized in terms of stimulus properties, are truly heterogeneous – they are disjunctive properties, or disjunctions of properties.

P1 I will defend in the following way. Having rejected type physicalism, disjunctive physicalism still remains a plausible account of object color, and I will endorse that view. This gives us the following link to the consequent of **P1**:

[P3] Any representational externalist view about color experience has to be able to maintain that object colors play a key role in determining the phenomenal character of color experience.

[P4] Disjunctive physicalism has no resource to support the claim that object colors play any important role in determining the phenomenal character of color experience.

[C2] Disjunctive physicalism cannot, in any plausible way, go externalist about color experience. (In other words, disjunctive physicalism and phenomenal externalism are incompatible.)

I will defend **P3** by mentioning some theoretical considerations about representational externalism. Finally, I will defend **P4** by arguing in two ways. First, the disjunctive physical properties that are the colors cannot subserve representational content of the kind that is identified with phenomenal character in phenomenal externalist theories. I will explain how such an argument stands up against two leading versions of phenomenal externalism: Dretske's and Tye's theories (Dretske, 1995; Tye, 1995, 2000). That is, I shall argue that color experiences do not have representational content of the kind that, in Dretske's and Tye's account, is identified with their phenomenal character. The second direction of my argument for **P4** will be the following. The fact that there is no physical substrate of perceived color similarity makes it virtually impossible to attribute any role to object colors in *causally explaining* how the phenomenal character of color experience arises.

In order to lend full support to **P1**, **C2** has to be supplemented by the following auxiliary premise:

[P5] No theory of object color other than type physicalism would make phenomenal externalism a coherent and plausible view.

In what follows, I organize my material in such a way that the argument for this auxiliary premise will be included in the defense of **P4**. That is, in effect, I will defend the following argument:

[P4'] Only type physicalism about color has the resources to support the claim that object colors play an important role in determining the phenomenal character of color experience.

I will argue that **P4'** holds true at least as long as we endorse either Dretske's or Tye's accounts of representational content – the notions that these authors use to explain phenomenal character (Dretske, 1995; Tye, 1995, 2000). I shall supplement this picture (in 2.5.2.2.) with an argument to the effect that no notion of disjunctive representational content (i.e., content that could arise from a disjunctive physicalist theory of color) would do for purposes of phenomenal externalism. In sum, I shall argue that as long as some causal theory of representation is assumed, only type physicalism about color can fill the phenomenal externalist bill. Moreover, since all current versions of representational externalism explain representation by some causal relation between states of the

mind/brain and entities in the environment (i.e., all current externalist theories of representation are causal), they all require the correctness of type physicalism about color.

The premises **P3** and **P4'** together support the following conclusion:

[C2'] Phenomenal externalism about color experience is not compatible with any view of object color other than type physicalism.

As I just said, this holds true for the currently available phenomenal externalist views that rely on a causal theory of representation. To summarize, my defense of **P1** will consist in supporting **C2'**, which is just a paraphrase of **P1**. This way I reach **C1**, the defense of which is my aim in this dissertation.

2.1. Defending the second premise: Type physicalism and the reflectance theory of color

In sections 2.1 to 2.4 I shall consider a number of options to defend type physicalism about color. As I will argue, such a defense has to consist of two steps. First, one has to fill the reflectance theory with empirical content: one has to tell precisely which perceptual color categories correspond to which types of reflectance. This of course need not take the form of an explicit list consisting of a million different descriptions of SSR types (roughly the number of shades that an average trichromat human can discriminate), but it has to be some sort of a schema that yields empirically testable type descriptions for at least the most basic color categories. The general claim that colors are non-disjunctive types of reflectance sounds vacuous without such a schema. Furthermore, the truth of this claim is an empirical matter, despite the *prima facie* impressions of plausibility that might arise from armchair reasoning. (Remember from Section 1.1.3 that all type descriptions derived from such a schema must be non-disjunctive.) As I will argue later, it is very likely possible to give such a reflectance schema for broader color categories like red, green and so on. However, due to individual differences in color perception, such a schema very likely cannot be given for narrow shades.[25]

The second step in establishing type physicalism is to generalize the reflectance theory of color to color stimuli that are not reflectances. (One such example is emitting surfaces like a color TV monitor.) The generalization that would support type physicalism should hold that color is some stimulus property P that is not just reflectance, but a more general property of which reflectance is a special case. I shall argue that there

is no plausible way to make this generalization, that is, to find a non-disjunctive, causally effective property that is true of all and only red objects and that is specifically causally responsible for our red sensations (*mutatis mutandis* for other colors). This will be my main reason for rejecting type physicalism.

2.2. Colors are types of reflectance, but which colors are what types of reflectance?

As I said in section 2.1 above, the reflectance theory needs to give us a schema for characterizing particular colors (at least the broadest color categories) in terms of surface reflectance, in order to have any plausibility. In other words, the reflectance theory is at least in part a theory of empirical science, and not a "pure philosophical theory of object color", to support which armchair reasoning alone is sufficient. In what follows I consider some proposals to solve this problem available in the literature. Note that each one of these proposals is automatically a proposal to solve the problem of metamerism, namely to specify the reflectance property that is common to all and only those reflectances that give rise to a particular perceptual look (for trichromat humans in normal circumstances of perception).

2.2.1 Hilbert's proposal: triplets of integrated reflectances

Hilbert (1987, p. 111) claims that colors are *triplets of integrated reflectances* (TIRs), without offering a characterization of any particular color in terms of such triplets of reflectances. The TIR of a surface (i.e., the TIR that is relevant to human color perception) can be obtained in the following way. Take the ranges of sensitivity of the three kinds of retinal cones: it is approximately from 400 to 525 nm for short-wave cones,

from 435 to 640 nm for middle-wave cones, and from 450 to 680 nm for long-wave cones (see e.g. DeValois and DeValois, 1997, fig. 4.1 on p99). Take a colored surface; *integrate its surface reflectance above the sensitivity range* of the short wave cones (400-525nm). (I.e., add up the reflectances obtained for every adjacent, very narrow band of wavelength within that range.) This gives the first member of the TIR of the colored surface under examination. Repeat the same procedure for the remaining two sensitivity ranges, thereby obtaining the other two members of the triplet. It is claimed that basically every member of a given metamer set has the same TIR whereas members of different metamer sets have different TIRs.[26] TIRs are the same kind of properties as individual, determinate reflectances – they can be regarded as some very crude characterizations of surface reflectances. Due to the limitations in spectral discrimination of our color vision, different objective colors that have the same TIR are perceived as the same color.[27]

TIRs nicely illustrate the idea of perceiver-relativity, and the fact that object colors in the type physicalist view are highly derivative, "uninteresting" properties, or *anthropocentric natural kinds* (Hilbert, 1987, 13-5; 115; 119-120, Gibbard, 1996; Tye, 2000, p. 161). Integrals of reflectance over the three mentioned ranges of wavelength are picked up by human color vision, but not by the color vision of very many other species. For those other species, *our colors* do not make much difference. The colors of other species (i.e., the reflectance properties that the color vision of other species track) are pairs, triplets or quadruples of reflectance integrals over more or less different ranges of electromagnetic wavelength.

Hilbert's proposal is based on Edwin Land's experiments on color constancy (Land, 1977; see also Land 1997). An experimental test of the predictions of Land's

59

Retinex Theory was conducted by McCann et al. (1976). In Part II of their paper, the authors examine the hypothesis that color sensations are determined by triplets of lightnesses and that these lightnesses in turn correspond to integrated triplets of reflectances (McCann et al., 1976, p. 448). As they summarize their results: 'Our results show that the color sensations are very highly correlated with the triplets of reflectance' (McCann et al., 1976, p. 446).

However McCann et al. did not even examine the hypothesis that TIRs, that is, *integrals of reflectance as such* of the surfaces used in Land's color constancy experiments correspond to perceived color. They began their examination by examining the correlation between subjects' color perceptions and *integrals of reflectance, weighted by the spectral sensitivity distributions of the retinal cones* (McCann et al., 1976, p. 449). Hilbert mentions this variation (1987, note 9 on p. 111), but he does not seem to find it in any way problematic. However, it *is* a highly problematic modification: I will explain why this is so in Section 2.2.4. It is also important to repeat that in his 1987 book, Hilbert does not give us any example of how a particular color or color category could be characterized in terms of TIRs.

2.2.2. Tye's schema: surface reflectance and opponent processing

Byrne and Hilbert (1997, pp. 265-266) offer a vague clue as to how a schema of colors in terms of surface reflectances could be made more specific. Tye (2000, 159-161) works out this schema in more detail to characterize the eight broad chromatic categories (red, green, yellow, blue, orange, purple, yellowish green and bluish green), plus achromatic grays in terms of surface reflectance. Tye proposes to obtain the particular,

highly derivative and uninteresting physical types (or *anthropocentric natural kinds*: see Hilbert, 1987, 13-5; 115; 119-120) that are the colors we trichromat humans can see directly from the opponent process theory of color perception (Tye, 2000, 159-161). He admits that he uses an oversimplified version of this theory (taken from Hardin, 1988), but assumes (presumably) that such a version will do as a rough approximation to define at least the nine broad color categories mentioned above. Capturing these color categories is the maximum that we can expect from Tye's schema, even if it works well enough. However, if the schema does *not* work well enough, then, according to Tye, it can be corrected by parameters characterizing trichromat human observers (Tye, 2000, pp. 160, 161). In addition to its limitations (i.e., to just the eight most basic color categories) the definitions contain vague elements that need to be sharpened if one really wants to see how the schema works in classifying surface reflectances.

So here's the schema. Let S^* be the percentage of "short-wavelength light" (a vagueness that needs to be properly disambiguated for purposes of empirical testing) that a particular surface **S** tends to reflect (or reflects actually when illuminated by white light). Similarly, M^* is the percentage of "medium-wavelength light" that **S** tends to reflect from white light. L^* is the proportion of "long-wavelength light" reflected by **S**. According to Tye (see also Byrne and Hilbert, 1997, pp. 265-266, and note 9 on p. 282) here's how broad color categories can be defined in terms of surface reflectance. Red surfaces are those for which $M^* <_s L^*$ & $S^* =_a M^*+L^*$ ('$<_s$' means *significantly smaller than*, whereas '$=_a$' means *approximately equal to*). Green surfaces are defined thus: $M^* >_s L^*$ & $S^* =_a M^*+L^*$. Yellow ones are those for which $M^* =_a L^*$ & $S^* <_s M^*+L^*$. Blueness is defined as $M^* =_a L^*$ & $S^* >_s M^*+L^*$; being orange is $M^* <_s L^*$ & $S^* <_s$

M*+L*; purpleness is $M^* <_s L^*$ & $S^* >_s M^*+L^*$. A surface is bluish green if $M^* >_s L^*$ & $S^* >_s M^* +L^*$ is true of it; yellowish green is defined as $M^* >_s L^*$ & $S^* <_s M^*+L^*$. If $L^* =_a M^*$ and $S^* =_a L^*+M^*$ obtain for a surface, then it is classified as achromatic: white, black, or gray. To this Tye adds (2000, p. 159) that typical achromatic white, gray and black surfaces have SSRs that are approximately constant functions of wavelength. For instance, a typical white surface reflects between 90-100% of the incident light at any wavelength between 400 and 700 nm; a typical black surface reflects around 4% at any wavelength. Hereafter by a typical achromatic surface I will mean a surface whose SSR is an approximately constant function of wavelength.[28]

2.2.2.1 Colors, reflectances, and the opponent process model of color perception

It is important to briefly explain why and how this schema is based on the opponent process theory of color vision. In what follows, I will use the following notation: 'L' will mean the output of cones sensitive to long wave light (in effect, 450-680 nm, with peak sensitivity being around 590 nm). 'M' will mean the output of the cones sensitive to mid-wavelength light (435-640 nm, peak sensitivity around 550 nm). 'S' will mean output of the short wave cones (sensitivity range 400-525 nm, peak sensitivity around 445 nm). Sometimes I will also use '**S**' (boldface) as a variable to refer to some hypothetical surface.

The opponent process theory is an abstract, mathematical model of human color vision that explains a large number of findings in color vision, and hence is very well supported empirically. However, the ways in which this abstract model is realized by neuronal mechanisms is not yet entirely clarified. Still, evidence shows that key steps in

calculating the so-called opponent signals from the activities of the three retinal cone types occur as the cone signals, via the ganglion cells, reach the LGN of the Thalamus (DeValois and DeValois, 1997). Spectrally opponent cells are frequent in the LGNs. Spectrally opponent cells are cells that show response to one wavelength (i.e., to signals of cones sensitive to that wavelength) and inhibit to signals from cones sensitive to another wavelength. Correspondingly, evidence shows that all LGN cells receive input from at least two cone types (DeValois and DeValois, 1997, pp. 104-107). Chromatic adaptation studies with single-cell recording showed six different, chromatically opponent cell types in the LGN. These are: +B-Y cells show elevated activation to S-cone inputs, and inhibition to L-cone inputs. +Y-B cells show inhibition to S-cone inputs and excitation to L-cone inputs. +G-R cells are excited by M-cone input and inhibited by L-cone input. +R-G cells are inhibited by M-cone input and excited by L-cone input. +Wh-Bl cells are excited by both L-cone input and M-cone input. Finally, +Bl-Wh cells are inhibited by both L- and M-cones (DeValois and DeValois, 1997, pp. 104-107). From these data we can already derive the key characteristics of color appearance. The activity of the LGN cell types forms three axes of the (human) perceptual similarity space for colors. The first is an achromatic lightness, or white-black channel. The second is the red-green channel, and the third is the blue-yellow channel. As a result of the opponent coding, and in entire accordance with the phenomenology of color experience, we can never experience colors like reddish green or bluish yellow. If the red-green channel gives positive input, the result is a perception of something reddish; if the red-green channel gives negative output, the result is the perception, or experience, of green. (Note that the signs like "positive is red, negative is green" are merely a matter of convention.)

If the blue-yellow channel gives positive output, the result is the experience of yellow whereas negative output from this channel signals blueness. The lightness or brightness of color stimuli is coded by the activity level of the white-black channel; saturation is coded by the intensity of the chromatic signals as related to the white-black signal. That is, from the abstract model of opponent processing and its neurological underpinnings we can start to understand the organization of the color space – the trichromat human perceptual similarity space for colors.

The distinction between unique and binary hues can also be understood from this model. If the red-green channel outputs zero (shows baseline activity, no bias to either direction), and the blue-yellow channel outputs a positive value, the result is the perception of unique yellow (of some saturation and lightness). If the red-green channel outputs a positive value, and the blue-yellow channel outputs a negative one, the result is the perceptual mixture of redness and blueness, that is a color in the purple range. A color in the purple range might be purple, magenta, pink, or purplish blue, depending on [i] the ratio of the absolute values of activity of the red-green channel and the blue-yellow channel, and [ii] saturation, i.e., the ratio of activity of the chromatic channels and that of the white-black channel. For instance, a strong red signal with a relatively weak blue signal can result in a perception of magenta if saturation is high; if saturation is low, the result is a perception of pink. A strong blue signal plus a weak red signal with high saturation results in a purplish blue perception. If both chromatic channels are in balance (i.e., give baseline, or "zero" output), then the result is an achromatic perception: that of white, black, or gray.

The mathematical details of a simplified version of the opponent process theory (Hardin, 1988, pp. 34-35) are slightly different from what we could directly infer from the neurological data presented above. This is because neither the neurological findings nor their interpretation is completely free of controversy. Details of the opponent process theory are under development. In the model presented by Hardin, L+M (the sum of L-cone inputs and M-cone inputs) results in the achromatic white-black response, and L=M is neutral mid-gray (sometimes called "brain gray"). L-M is the red-green response – so far in entire agreement with the physiological data presented above. However, the yellow-blue response results from L+M-S, and this formula cannot be derived from just those physiological data I presented above. The case of the yellow-blue channel seems a little more complicated than that of the other two. As Hardin remarks (1988, p. 35; he refers to Werner and Wooten, 1979 and Hunt, 1982), there is a nonlinear component in the opponent process coding, and this is most prevalent in the operation of the blue-yellow channel. Obviously, the simplified version of the model I just presented omits these nonlinearities. However, this feature is of interest for my purposes. As I shall point out below, the problems I found with Tye's schema when putting it to an empirical test, concentrate around blue-yellow (bluish-yellowish) color assignments. I will also explain how such a finding relates to the fact that the blue-yellow channel is nonlinear.

It now becomes clear how Tye's schema derives from the opponent process theory. For the sake of hypothesis, make the following two assumptions. First, assume that a linear relation obtains between [1] the difference of light reflection in the long-wavelength range and that in the middle-wavelength range (of the surface perceived) and [2] the red-green opponent signal. That is, if the former is multiplied by some number **k**,

then the latter is multiplied by **k** as well – similarly for the other two channels. Second, assume that no differential weighting (multiplicative transformation) of the color signal occurs before it is converted into opponent process signals. The assumption here is that the quantities S, M, and L – the ones that get added up and subtracted in the color vision system as modeled by the simplified opponent process formulas – are still directly proportional to the integrals of reflectance (reflection under white light). If such a case obtains, then, *and only then*, we can reach, from the fact that there is balance, say, in the red-green channel, the conclusion that the surface being perceived reflects roughly as much light in the middle wavelength range as in the long-wavelength one (Tye, 2000, pp. 160-161). Given the two assumptions, the fact that unique blue and unique yellow perceptions correspond to L-M=0 opponent process response (baseline activity in the red-green channel) entails that, for any surface **S**, if $L^*-M^* =_a 0$ for **S**, then **S** is neither reddish nor greenish. However, if there is nonlinearity or differential weighting in the processing of color stimuli, then any such inference is unwarranted, therefore we might suspect that the predictions it results in are incorrect.

2.2.2.2 Variations on Tye's schema: Matthen and Kuehni

Matthen (1999, pp. 49, 65-66; 2001, pp. 119-120) offers a very similar scheme for defining colors in terms of reflectances. His schema too is based on the opponent process theory of color vision. For some unspecified reason, Matthen assumes that the red-green signal arises from computing the L+S-M function: he takes the red-green channel to compare the relative strength of the color signal at the two ends of the visible spectrum as compared to the middle (Matthen, 1999, p. 49). The blue-yellow signal is taken to be

L+M-S: it compares the strength of the color signal in the middle-plus-long wavelength range to that in the short wavelength one. Matthen outlines only four chromatic color categories that derive from this model (Matthen, 2001, p. 120; in his 1999, pp. 49-50 we find even vaguer characterizations). If L+S>M (i.e., the two ends of the spectrum dominate over the middle – this corresponds, in his view, to perceived reddishness) and L+M > S (the long-plus-middle wavelength ranges dominate over the short one – this corresponds to yellowishness) obtain, the result is the perception of orange.[29] If L+S<M (mid-wavelength range dominates over the ends – the result is greenish) and L+M>S obtains, then the result is green or yellowish green. If L+S>M and L+M<S obtain (the short-wavelength range dominates over the long-middle one resulting in bluish) then the result will be some purple color like violet. Finally, if L+S<M and L+M<S then the result is bluish green.

This being Matthen's proposal, it should be immediately obvious that it contains a weird contradiction, and for that reason it is quite unusable. The condition that he offers for bluish greens is impossible to satisfy: L+S<M and L+M<S cannot be true together. This is because (speaking in terms of light reflection or emission) L*, M*, and S* are greater than or equal to zero: there is no such thing as less-than-zero emission or less-than-zero-percent reflection. Similarly, L, M, and S, the cone responses, are either positive or zero – negative cone responses do not figure in the opponent process theory (only negative opponent channel responses do).[30] If so, then the condition for greenishness (L+S<M) entails that S<M; however, the condition for bluishness (L+M<S) entails that S>M. So, Matthen's schema predicts that no stimulus can ever look bluish green, and so there are no such things as bluish green objects – simply because no surface

can ever reflect, or emit, more light in the M* range than in the S* range and at the same time, reflect, or emit, more light in the S* range than in the M* range. My guess is that Matthen simply overlooked something here. Apparently he is not using the simplified version of the opponent process theory presented in Hardin (1988), and he gives no indication where he takes his version from. The justification for his choice seems to be that, in terms of lights reflected or emitted, both ends of the visible spectrum appear reddish: the short end appears violet whereas the long end appears red or orange (Matthen, 1999, pp. 49-50, 65-66). This is correct as far as it goes; however, the two formulas he offers are incompatible.

Rolf Kuehni (personal communication[31]) proposed that the simplified version of the opponent process theory that gives rise to roughly correct characterizations of color in terms of reflectance consists of the following two functions:

`a=L-M+0.5*b;`

`b=L+M-S.`

Of these, L+M-S, or **b**, still corresponds to the blue-yellow response, whereas **a**, that is,

`1.5*L-0.5*M-0.5*S,`

corresponds to the red-green response. Happily, this modified schema is not contradictory: both a>0 and a<0 are consistent with either b>0 or b<0. In what follows I will assess this version of the schema, in addition to Tye's. Due to its inconsistency, I will not test Matthen's version in detail, but I will add a few remarks on it.

2.2.3. Empirical assessment of the opponent processing schema for characterizing colors in terms of reflectances

2.2.3.1 Tye's version

Tye's proposal for identifying the reflectance types that correlate with our color perceptions is radically wrong. This is so despite the fact that he derived his method from the otherwise-well-confirmed opponent processing model of color perception. As it stands, Tye's schema gives entirely incorrect predictions as to the perceived color of surface reflectances. I applied his schema, disambiguating its vague parts in several different ways, to my own surface reflectance measurements of various natural and man-made surfaces taken by a Spectrogard II spectrophotometer.

Some details of this empirical test are as follows. I took the measurements at 10 nm intervals, in most of the cases with the interface-reflection excluded.[32] Surfaces measured included the Macbeth Color Checker, all samples of the Optical Society of America Uniform Color Scales, colored papers, plastics, clothes, paints (watercolor, food color), minerals, wood samples, autumn tree leaves, fruits, vegetables, and miscellaneous items like soap bars, chocolate bars, and coins. I used the Macbeth Color Checker to check the correctness of my measurements. I compared my own data obtained from this color rendition chart to the measurements of it published in Meyer, 1988. Matches were pretty good: visually, my reflectance curves and Meyer's ones matched to minute detail. Also, I checked the actual percentage values at selected points of reflectance curves obtained from the same sample by Meyer and me, and nowhere did I find a difference exceeding two per cent.[33]

As a means of checking Tye's schema, I calculated integrals of different color signals obtained from a given surface. For calculating color signals (i.e., to characterize actual reflection as well, not just reflectance) I used the spectral power distributions (SPDs; relative energy distributions) of four CIE standard illuminants: A, C, D50, and D65. I also used the non-standard fluorescent illuminant F11 (the data of these five illuminants I took from Hunt, 1991, Appendix 5). Finally, I used a theoretical uniform illuminant whose spectral power distribution was the same at all wavelengths (the value was 1), in order to assess integrals of reflectance directly.[34]

As short, medium, and long wavelength parts of the spectrum I used two different triplets of intervals. First, the sensitivity ranges of the human cone types: 400-525 nm (short wavelength range), 435-640 nm (medium wavelength range), and 450-700 nm (long wavelength range). Second, I used simply the 400-500, 500-600, and 600-700 nm intervals. In order to obtain numerical values that correspond to S*, M*, and L* for a given surface, I integrated the color signals arising from multiplying the reflectance of the surface by the data of the six illuminants mentioned above. Thus, for a given surface, I calculated twelve different triplets of values (two triplets of intervals times six illuminants). After calculating the twelve triplets of integrals for an SSR, I applied Tye's schema to each one of them to see whether the surface in question is categorized by them correctly – that is, as having the same color as their perceptual look.

By "significantly greater/smaller than" I meant that for two integrals, their proportion was outside the [0.8; 1.25] interval. Two integrals were taken to be approximately equal when their proportion remained in the [0.87; 1.15] interval. I tried out other values as well, but that did not make much difference to the classification

results. By correct categorization I meant the correct application of the eight color categories the schema was meant to capture: e.g., categorizing both scarlet and cadmium red simply as red. I repeated this procedure for a number of red, green, yellow, blue, orange, purple, yellowish green and bluish green surfaces.

The results were daunting: typical red surfaces were not classified as red by the schema; the same was true of green, blue, and purple ones. Yellow surfaces were sometimes classified as yellow (heavily depending on which standard illuminant was used); in the majority of the cases, yellow surfaces were classified as orange. Orange surfaces were correctly classified as orange, but the value of this finding is questionable since red and purple surfaces were also quite unambiguously classified as orange by the schema. The schema gets partly right yellowish green surfaces, but it also predominantly classifies bluish green ones as yellowish green (sometimes as yellow, depending on the illuminant used). Also, Tye's schema massively classifies typical achromatic white and gray surfaces (i.e., those kinds of SSRs that he himself identifies as achromatic white/gray/black on p. 159) as yellow or orange.

This latter result is easy to understand. Just because the medium-plus-long wavelength range is wider than the short wavelength range alone, if we add up (integrate) the amount of light reflected by a typical achromatic surface in the medium-plus-long wavelength range (under some normal, broadband illuminant), the result will be substantially greater than the amount of light reflected in the short wavelength range alone. Hence, even if $M^* =_a L^*$ stays true for typical achromatic surfaces (it need not, but it sometimes does) $S^* <_s M^*+L^*$ will be satisfied, and so a surface with a constant SSR will always be classified as yellow or yellowish.

In general, with respect to the relation between L* and M*, the schema is not so widely off the mark. It is true that for red (reddish) surfaces, $L^* >_s M^*$; the opposite is true of green (greenish) ones. However, for many reflectances that look *unique* blue, $L^* <_s M^*$ typically obtains, therefore most unique blues are classified as bluish green (some less saturated unique blues, for which $S^* =_a L^*+M^*$ obtains, are classified as green). The relation between S* and L*+M*, is, in reality, entirely different from what the schema assumes. For instance, for typical red, green, or even many purple surfaces, S* is significantly smaller that L*+M*. Tables 1 and 2 show some of the test results for Tye's schema.

Illuminant	Division	Red	Green	Yellow	Blue	Orange	Purple	Yell. green	Bluish green	White
Refl. Direct	D1	orange	yellow	orange	yellow	**orange**	orange	orange	yellow	orange
	D2	orange	yel-grn	orange	**blue**	**orange**	orange	**yel-grn**	yel-grn	yellow
A	D1	orange	orange	orange	orange	**orange**	orange	orange	orange	orange
	D2	orange	yel-grn	orange	orange	**orange**	orange	orange	yellow	orange
C	D1	orange	yellow	orange	yellow	**orange**	orange	orange	yellow	yellow
	D2	orange	yel-grn	**yellow**	blu-grn	**orange**	orange	**yel-grn**	yel-grn	yellow
D50	D1	orange	yellow	orange	yellow	**orange**	orange	orange	yellow	orange
	D2	orange	yel-grn	orange	**blue**	**orange**	orange	**yel-grn**	yel-grn	yellow
D65	D1	orange	yellow	orange	yellow	**orange**	orange	orange	yellow	yellow
	D2	orange	yel-grn	**yellow**	blu-grn	**orange**	orange	**yel-grn**	yel-grn	yel-grn
F11	D1	yellow	yellow	**yellow**	yel-grn	yellow	yellow	yellow	yellow	yellow
	D2	orange	yel-grn	**yellow**	blu-grn	**orange**	orange	**yel-grn**	yel-grn	yel-grn

Table 1. Tye's schema applied to nine samples of the Macbeth Color Checker. The names of the nine columns are the same as those of the corresponding Macbeth samples. The Macbeth Color Checker consists of 24 samples, the first 18 of which are chromatic colors, the last six are an achromatic series from white to black. Five different CIE illuminants were used (e.g., illuminant A corresponds to a tungsten bulb; illuminant C to average daylight; illuminant F11 represents fluorescent tubes, and so on). The first row shows the results obtained by using a theoretical illuminant whose spectral power distribution was constant over wavelength: the value was one (1) appropriately chosen unit of measure. This corresponds to

integrating the reflectances directly, without taking into account any modifying effect of the illumination. D1 and D2 are two different divisions of the visible spectrum into short, middle and long wavelength ranges. D1 corresponds to the sensitivity ranges of the three cone types: 400-525 nm for short wavelength range, 435-640 nm for middle wavelength range, and 450-700 nm for long wavelength range (450-680 nm was also tried, but no significant difference occurred due to this modification). D2 corresponds to division into three equal intervals: 400-500 nm for short wavelength, 500-600 nm for middle wavelength, and 600-700 nm for long wavelength range. The cells contain the output of the classification algorithm that implements Tye's simplified opponent process schema, with the vague elements clarified as stated in the main text. ('yel-grn' means yellowish green, 'blu-grn' means bluish green.) Hits are typed in boldface. Note the strong illuminant-dependence of the schema that is entirely different from the invariant classification of the same color samples under the same illuminants made by any trichromat human subject.

Illumi-nant	Divi-sion	Red Lego block	Red plastic boat	Green Lego block	Green plastic boat	Yellow Lego block	Yellow plastic boat	Blue Lego block	Blue plastic boat	White ceramic tile
Refl. Direct	D1	orange	orange	yellow	yellow	orange	orange	yellow	yellow	orange
	D2	orange	orange	yel-grn	yel-grn	orange	**yellow**	blu-grn	blu-grn	yellow
A	D1	orange	orange	yellow	orange	orange	orange	yellow	orange	orange
	D2	orange	orange	yel-grn	yel-grn	orange	orange	green	orange	orange
C	D1	orange	orange	yellow	yellow	orange	orange	yellow	yellow	yellow
	D2	orange	orange	yel-grn	yel-grn	**yellow**	**yellow**	blu-grn	blu-grn	yellow
D50	D1	orange	orange	yellow	yellow	orange	orange	yellow	yellow	orange
	D2	orange	orange	yel-grn	yel-grn	orange	**yellow**	blu-grn	blu-grn	yellow
D65	D1	orange	orange	yellow	yellow	orange	orange	yellow	yellow	yellow
	D2	orange	orange	yel-grn	yel-grn	**yellow**	**yellow**	blu-grn	blu-grn	yel-grn
F11	D1	orange	yellow	yellow	yellow	**yellow**	**yellow**	yel-grn	yel-grn	yellow
	D2	orange	orange	yel-grn	yel-grn	**yellow**	yel-grn	blu-grn	blu-grn	yel-grn

Table 2. Tye's schema applied to colored plastics and a white ceramic tile provided to calibrate the Spectrogard II spectrophotometer. Note that though both blue plastic surfaces are neither obviously yellowish nor obviously greenish (i.e., they are as close to unique blue as one could require from an industrial product), they are never classified blue, contrary to the Macbeth Blue. This shows that, in addition to being unreasonably illumination-dependent (see Table 1 above), the schema is also not robust against small variations in surface reflectance that do not strongly affect trichromat color perception. The colored plastics used were Lego building blocks and plastic toy boats. As in Table 1 above, hits are typed in boldface; all other notations are the same as well.

On looking at the tables, notice three things. First, it makes a big difference which wavelength range division we use. The [400-525nm; 435-640nm; 450-700nm] division is theoretically motivated, as these intervals roughly correspond to the sensitivity ranges of the three human cone types. The [400-500nm; 500-600nm; 600-700nm] division is not theoretically motivated. Ironically, the latter division gives somewhat better results (that

are still far from acceptable): for instance, in Table 1, all blue and yellowish green hits are under this division – with the [400-525nm; 435-640nm; 450-700nm] division the schema never gets right the Macbeth Blue and Yellowish green. In general, results are closer to correct under the [400-500nm; 500-600nm; 600-700nm] division. Second, though both blue plastic surfaces in Table 2 are neither obviously yellowish nor obviously greenish (i.e., they are as close to unique blue as one could require from an industrial product), they are never classified blue, contrary to the Macbeth Blue. This suggests that the schema is not robust against small variations in surface reflectance that do not strongly affect trichromat color perception. Third, the schema seems to be too illumination-dependent in some cases (especially illuminants A and F11 are apt to produce results that deviate from those with the other illuminants); still, in other cases it is fairly illumination-independent, both in its errors and in its correct classifications (e.g., orange in Table 1).

What went wrong? As we saw, Tye based his schema on the opponent processing model of color perception. That model is quite well supported by empirical data; then how could Tye's schema, based on it, be so badly mistaken? On this point, Tye (2000, p. 161) says: "…any actual counterexample to this proposal will also be a counterexample to the opponent processing model…". But of course, no such link exists. Tye's schema faces countless actual counterexamples, but this does not affect the correctness status of the opponent process theory at all. Recall what was said at the end of Section 2.2.2.1. Tye fails to realize that [1] there is no simple linear relation between surface reflectance (or light, actually reflected by illuminated surfaces) in the short-, medium-, and long wavelength ranges on the one hand, and the opponent process response elicited by

reflecting surfaces on the other (e.g., Hardin, 1988 p. 35), but [2] his schema assumes exactly such a relation. In addition, differential weighting of the three cone signals also occurs in color processing. And that is probably the reason why the schema goes wrong.

2.2.3.2. Matthen's version

In addition to its inconsistency, another problem with Matthen's proposal is that he does not spell out explicitly the definitions of unique hues, as opposed to binary ones.[35] Had he done so, the problems with his schema would have turned up immediately. For instance, unique blue is a chromatic color that is neither reddish, nor greenish, and it is not yellow either. But if a color is neither reddish nor greenish, then, on perceptually processing it, no bias in the L+M-S (red-green) channel should occur – this channel should exhibit zero (baseline) activity on processing a stimulus in the unique yellow range. This follows from the opponent processing theory of color vision. However, what does *not* follow from the opponent process theory is that when the red-green channel is in balance, then, and only then, *the color signal* hitting the retina is itself balanced – i.e., that it has roughly as much energy in the short-plus-long wavelength range as in the middle one. In the case of color signals (or reflectances) that correspond to unique blue perceptions, the latter is radically false. Unique blue surfaces give rise to much more radiation in the short plus long wavelength ranges together than in the middle one, under normal illuminants. So Matthen's schema predicts that stimuli that, as a matter of fact, look unique blue should look strongly reddish. Similar problems arise with regard to unique red and unique green. When the blue-yellow channel is predicted to be balanced by the opponent process theory, the color signal itself that gives rise to the

corresponding unique green or unique red perceptions contains *much* more energy in the long-plus middle range than in the short one (see 2.2.3.1 above). This again suggests that we cannot directly turn the formulas of the simplified opponent processing model into type specifications of object color.

2.2.3.3 Kuehni's version

Without further testing, we can immediately say the following about Kuehni's proposal. Since it leaves the blue-yellow channel function unchanged, the errors that occur in Tye's and Matthen's schema due to the L+M-S formula will also occur in this modified version. In particular, this version will, just like Tye's, classify red, orange and purple surfaces as yellowish – that is, it will not be able to properly distinguish between these three color categories along the blue-yellow dimension. Table 3 shows some classification results obtained by this version of the schema.

Illumi-nant	Divi-sion	Red	Green	Yellow	Blue	Orange	Purple	Yell. green	Bluish green	White
Refl. Direct	D1	orange	orange	orange	orange	orange	orange	orange	orange	orange
	D2	orange	yel-grn	orange	blu-grn	orange	orange	orange	yel-grn	orange
A	D1	orange	orange	orange	orange	orange	orange	orange	orange	orange
	D2	orange	orange	orange	orange	orange	orange	orange	orange	orange
C	D1	orange	orange	orange	orange	orange	orange	orange	orange	orange
	D2	orange	yel-grn	orange	blu-grn	orange	orange	orange	yel-grn	orange
D50	D1	orange	orange	orange	orange	orange	orange	orange	orange	orange
	D2	orange	yel-grn	orange	blu-grn	orange	orange	orange	yel-grn	orange
D65	D1	orange	orange	orange	orange	orange	orange	orange	orange	orange
	D2	orange	yel-grn	orange	blu-grn	orange	orange	orange	yel-grn	orange
F11	D1	orange	orange	orange	orange	orange	orange	orange	orange	orange
	D2	orange	yel-grn	orange	blu-grn	orange	orange	orange	yel-grn	orange

Table 3. Classification data of the nine Macbeth samples shown in Table 1, using the schema suggested by Rolf Kuehni (**a**=L+M+0.5**b**; **b**=L+M-S).

I will further discuss Kuehni's relevant work in Section 2.2.4 below.

2.2.4. Corrections by observers' parameters

To the findings in the previous section Tye would reply (2000, pp. 160-161) that the oversimplified opponent process model (on which he based his color-definition schema) has to be corrected, or "qualified" by parameters of the trichromat human observer (among other things, by the spectral sensitivities of the three cone types). This would lead us to a more exact characterization of colors in terms of reflectances. As I shall argue in this and in the following section, this proposed remedy either leads us straight into a severe category mistake, or it leads us to a very difficult mathematical task, that, if solved, could possibly provide an exact characterization of colors (though only broad color categories, not narrow shades) in terms of surface reflectances. In this section I show how the proposal can easily run into a category mistake. This happened with Hilbert's 1987 account (see below for details), and since Tye's reflectance theory is obviously a descendant of Hilbert's one, his remarks about correction (they are very brief, and rather obscure: Tye, 2000, pp. 160-161) need careful scrutiny as well. In the next section I will show some ways in which one could possibly proceed in characterizing color categories in terms of surface reflectance. However, what I shall provide are just a few suggestions, not a complete solution.

If we measure the percentage of incoming light reflected by surface **S** in the short wavelength part of the spectrum (say, in the 400-525 nm range), then what we get is a reflectance property of **S**: integrated reflectance. Integrated reflectance is a perceiver-independent property of **S**. What percentage of the incoming light S is disposed to reflect back in the 400-525 nm range (or reflects back actually, when it is illuminated) is an empirical fact about **S**, or a property of **S** that is instantiated, or physically realized, in the

absence of any perceiver. Reflectance measured in a very narrow wavelength interval, and reflectance integrated over a wider wavelength interval both have the same unit of measure: per cent of incoming light, reflected (in the wavelength interval under consideration).

If, however, we measure certain reflectance properties of **S**, and then transform them by parameters characterizing some observer (e.g., multiply or divide them by coefficients, or raise them to different powers[36]), then the result of these transformations is not theoretically interpreted as a *reflectance* property of **S** – not at least if the coefficients or exponents used in the transformation are used to characterize the perceiving organism, (and not the surface **S** independently of its relation to the organism). Rather, the result of such transformations is theoretically interpreted as some characteristic response of the observer (i.e., as a perceiver-dependent, remote relational property of **S** – a property that does not inhere in **S**). In accordance with this observation, these latter transformations (unlike simply integrating reflectance) *will* change the unit of measure of their operands.

Here is a key example in more detail. When the disposition of surface **S** to reflect light in certain way manifests itself (i.e., **S** is illuminated), the *color signal* arises (e.g., Wandell, 1995, pp. 290-291). The color signal is the product of reflectance and illuminant spectral power distribution. It is the light actually reflected by **S** and can possibly reach the eye of an observer. The color signal is the key information source about surface reflectance for the color vision system. Also, the color signal is something that is instantiated, or physically realized, independently of any observer. In a world

where there are no observers, but there are illuminated reflecting surfaces, there is light reflected back from the surfaces, hence there is what we call a color signal.

Next step: weighting by the spectral sensitivity profiles of the three cone types is a correction that seems essential for deriving empirical correlates of color perceptions (see Hilbert, 1987, p. 111; McCann et al., 1976, pp. 449-451; Land, 1977; see also Maloney and Wandell, 1986). However, when we multiply the color signal by the spectral sensitivity profile of the cones (the unit of measure of the latter is probability of absorption per photon), we get a measure that expresses *amount of light (number of photons) absorbed by the cones*. Maloney and Wandell (1986, p. 29) call this measure *sensor quantum catches* (of the three cone types at particular retinal locations). The amount of light absorbed is transformed into a hyperpolarization of the cone membranes (this is the transducer function of the cones). Cone signals are then picked up by the ganglion cells. Both the absorption by the cones and the firing frequency that results upon this absorption belong to the response characteristics of the organism. These properties are not perceiver-independent: if no observers are instantiated, these characteristic responses aren't instantiated either. *Color signal, weighted by cone spectral sensitivity,* expresses (is interpreted in the relevant theory as) cone absorption. This property, unlike reflectance or integrated reflectance, is instantiated (physically realized) only when observers or appropriate measuring instruments (like telescopic photometers equipped with color filters: see McCann et al., 1976, Part II.) are instantiated.

If, as a result of the transformations and corrections, we get an empirical measure that (1) correlates well with reportable color perceptions and (2) characterizes some response characteristic of the observer, then type physicalists cannot lean back content in

their armchair, of course. For so far we haven't shown any perceiver-independent property of surfaces that correlates with color perceptions. Moreover, it is already well known that what underlies sameness of perceived color for us trichromat human perceivers is that somehow we perceptually respond the same way to all those stimuli that look to us the same in color. This is no news.

Now, in order to make Tye's schema work, we have to "qualify", or correct it by certain parameters of the observer (Tye, 2000, pp. 160-161). This is already obvious from the simple empirical test that I deployed to assess it. Such a correction will evidently include weighting the color signal arising from different surface reflectances by cone spectral sensitivities.[37] Cone spectral sensitivities constitute a color-matching function (Wandell, 1995, pp. 95-96). All color-matching functions are related to one another by a linear transformation (Wandell, 1995, pp. 85-86). Thus the unavoidable step of weighting by cone spectral sensitivities leads us straight back to the good old idea of CIE color standards: two surfaces perceptually match in color just in case they have the same tristimulus values – that is, they have the same effect on the three cone types. We slide back from Tye's schema to CIE's concept of color matching because in calculating the CIE tristimulus values from surface reflectance and illuminant SPD, we use the CIE standard color-matching function, which is again related to cone spectral sensitivities by a linear transformation. That is, weighting the color signal by cone spectral sensitivities amounts to essentially the same thing as weighting it by the CIE standard color matching function. Now, what is common, *in terms of reflectance*, to all and only those surfaces that have the same CIE tristimulus values, and look the same in color (under some "normal", or standard, illuminant), is a difficult question.[38] I will provide some details of

this issue in the next section – though I will not offer a complete solution to it. But since Tye's definitional schema does not work uncorrected, it fails to give an answer to this question – i.e., it fails to provide a solution to the problem of metamerism. Correction of Tye's schema by observers' parameters has to include weighting by cone spectral sensitivities; but after this step is made, the schema does not tell us any more about color *in terms of reflectance* than does the CIE color standard itself. Conclusion: Tye's proposal for characterizing colors in terms of surface reflectance works neither way: without corrections it is hopelessly incorrect, and the idea of correcting by observers' parameters, as a means of getting at more exact *reflectance* properties is a category mistake (i.e., a misunderstanding of what categories the properties in question belong to).

Hilbert's original schema for characterizing colors in terms of reflectances (triplets of integrated reflectances: Hilbert, 1987, p. 111) is mistaken for the same reason as Tye's more recent one.[39] Hilbert explicitly mentions that integrals of reflectance need to be corrected by cone spectral sensitivities in order to correlate well with color perceptions. He based his claim on McCann et al.'s (1976) study. In that study, the authors did not even examine whether integrated reflectance, unweighted by cone sensitivity, correlates with color perceptions. What they did find was that integrated reflectance weighted by cone spectral sensitivities correlated quite well with color perceptions.[40] Furthermore, when these weighted integrated reflectances were transformed by a power function (with exponent 1.3) to compensate for the fact that equal increments in reflectance do not represent equal increments in lightness sensation (McCann et al., 1976, pp. 449-450), the correlation of the thus-arising scaled-weighted-integrated triplets of reflectances with the subjects' color perceptions was excellent. But

apparently, this power function is a (nonlinear) sensory function whose exponent is there to characterize the sensitivity of the perceiver to the particular kind of stimulus in question. Once again, the measure that showed such a nice correlation with the subjects' conscious color perceptions in McCann et al.'s study expresses, or describes, a phenomenon that is a perceptual reaction of observers, not any *reflectance* property.

More recently, Rolf Kuehni (Kuehni, 2000a; see also Kuehni, 1999, 2000b) conducted a study to determine the relation between cone response functions and color-opponent responses. He found a simple opponent color model that derived the perceptual color space from simple subtractions of color-matching functions, in agreement with certain neuro-physiological findings about opponent processing. However, clarifying the relation between cone responses and the responses of the color-opponent channels leaves entirely undetermined the relation between surface reflectance (or color signal) and opponent response. The reason is that the very same cone response (or triplet of cone responses) can result from many different color signals. How should different reflectances or color signals be related in order to give rise to the same cone response ratio (i.e., the problem of metamerism) is left open in Kuehni's model.

2.2.4.1 Many-to-one mappings between stimulus properties and sensory states

There is a second stage of the problem just discussed that we have to consider. This key issue now is: do the different mental transforms of the color signal (sensory quantum catches, cone signals, or opponent neural responses) *necessarily* correlate with, or represent, some *non-disjunctive* stimulus property, or just a bunch of different, disjunctively related properties? Answer: it is by no means necessary that these states of

the color-vision system have non-disjunctive stimulus correlates. As I shall argue in a moment, sensory or perceptual states can in principle indicate disjunctive properties, just as they can non-disjunctive ones.

First notice that it makes a difference exactly how observers' parameters are used in correcting, or adjusting, measured stimulus properties. For instance, if we use the borders of sensitivity ranges of the cones as borders of integration, all that amounts to is a selection from the available facts about reflectance that obtain in our environment, independently of observers. What percentage of the incident light a surface **S** is disposed to reflect in the 400-525 nm range is a fact about **S** – a fact that is kept track of by human color vision. What percentage of the incident EM waves in the 265-335 nm range **S** is disposed to reflect is another fact about **S**, but a fact that is not kept track of by human vision.

Multiplication, or raising to powers (where the coefficients and exponents are derived from properties of the observer), are not so straightforwardly interpreted in terms of external properties – *it becomes an empirical question whether the resulting mental, or neural, transforms correlate with any non-disjunctive property (or natural kind essence).* They may or may not do so: there is no warranty that the outputs of sensory functions, especially those of nonlinear ones, have such correlates. This might need some explanation – here it is.

Step1: Abstract point and abstract example. Mathematical transformations are abundant. If we take some measurable attribute **q** of physical entities (e.g., reflectance, temperature, atomic weight, electric charge, etc.), and take an arbitrary set $\mathbf{H_1}$ of **q** values, we will always be able to find some abstract mathematical function that takes all (and

only) H_1 members into one and the same value. Now if we take another set of q values H_2 such that, the intersection of H_1 and H_2 is empty (i.e., there is no particular value that is a member of both sets), but the smallest closed interval in which all H_1's members are found *overlaps* with that in which all H_2 members are found, then there will always be a mathematical function which takes q values and gives a value V_1 to all (and only) H_1 members, whereas it gives another value V_2 to all and only members of H_2.

A simple example: let H_1 be {0, PI, 2*PI, 3*PI, 4*PI, 5*PI}. (PI = 3.1415926536…) Let H_2 be {1/2*PI, 3/2*PI, 5/2*PI, 7/2*PI, 9/2*PI}. Now, the abs(sin(x)) function[41], *defined only for the [0, 5*PI] closed interval*, will yield the value V_1=0 for all and only H_1 members, and the value V_2=1 for all and only for H_2 members.

Step 2: Analogy with sensation. In the above example, the members of H_1 and H_2 are analogous to stimuli for a q-sensor; limits of the domain (0 and 5*PI) are analogous to the sensitivity range of the sensor, and the values V_1 and V_2 are analogous to the outputs of the q-sensor ("sensations", or "sensory states"). That is, for this hypothetical q-sensor, all H_1 members are equivalent stimulus properties; so are all H_2 members. It can discriminate any H_1 member from any H_2 member, but cannot make any discrimination between two members of the same set. That is, this q-sensor now senses two "disjunctive" properties: one is *being either 0, or PI, or 2*PI, or 3*PI, or 4*PI, or 5*PI*, whereas the other is *being either 1/2*PI, or 3/2*PI, or 5/2*PI, or 7/2*PI, or 9/2*PI*. That is, there is a many-to-one mapping between stimulus properties (**q**-values) and sensory states (**V**-values). Whether such many-to-one sensory mappings are advantageous for an organism having a q-sensor depends on the evolutionary situation: there seems to be no reason to rule that they cannot, in principle, be advantageous. I will

return to this issue in a moment; for now, the key point is as follows. Sensory state types of a sensory system can correspond to, or signal, disjunctive stimulus properties, just as they can indicate non-disjunctive ones – even natural kind essences. The idea of a one-to-one mapping between (non-disjunctive) stimulus properties and sensory states, and that of a many-to-one mapping of the same kind are both perfectly consistent. There can be sensors, or gauges, whose particular needle positions indicate disjunctive properties. Such gauges can also be quite useful.

Step 3: Fending off an objection. Objection: members of H_1 are integral multiples of PI whereas H_2 members are odd-number multiples of PI/2 and this is a non-disjunctive common property in the two respective cases. That is, the stimulus properties corresponding to V_1 and V_2 are not disjunctive after all. Reply: wrong. First, another example could easily be constructed in which H_1 and H_2 members are randomly or irregularly distributed. Second, such a common mathematical property of H_1 and H_2 members does not make it the case that the different q-values in, say, H_1, constitute a non-disjunctive physical property (let alone a natural kind essence) rather than just a disjunction of different properties (q-values). Another intuition pump is in order to help understand this.

There surely exists (in the way in which abstracta exist in some Platonic realm of universals) some mathematical transformation T that takes the value of the atomic weight of lead and that of mercury into the same value k. That is, the T-transformed atomic weight of lead and mercury are the same: k. However, T is such that the T-transform of no value other than that of the atomic weight of lead and mercury will yield k as a result.[42] Does it follow *from this abstract mathematical fact* that all samples of lead and

mercury together constitute a single natural kind (i.e., that *lead-or-mercury* is a natural kind)? Or does it follow that lead and mercury have in common some measurable non-disjunctive property that they do not share with any other element? Hardly.

This is so even in the following imaginary situation. Imagine an organism called the HeavyMetalEater that has to regularly ingest, as part of its proper nutrition, small pieces of mercury and lead that can be found in its environment in pure form. (Say, because these elements are contained by some of its coenzymes.) The HeavyMetalEater recognizes lead and mercury on the basis of their relative density: it can sense the relative densities of samples in some way (e.g., by a sophisticated sensory mechanism that estimates the volume and weight of small pieces of metal). In order to distinguish lead and mercury from all other elements, the HeavyMetalEater's nervous system implements some computation equivalent to the T-transformation. That is, there is some neuronal state W whose occurrence (tokening) indicates to the organism that *either mercury or lead* is sensed/contacted. State W activated corresponds to value **k** as the result obtained from neurally running the T-transformation. Some other neuronal state, R, is the neuronal response to any other chemical element or compound contacted that is available in the HeavyMetalEater's environment. Of course, the mere fact that the HeavyMetalEater's nervous system implements some computation equivalent to the T-transformation, hence that its nervous system gives the same reply to mercury and lead, and some other reply R to any other solid substance available in its environment, does not make it the case that mercury and lead together constitute a single natural kind, nor that they share any measurable, non-disjunctive, "HeavyMetalEater-independent" property[43] that they do not share with any other element. Due to the HeavyMetalEater, mercury and lead share only

a remote relational, "HeavyMetalEater-dependent" property, namely that the HeavyMetalEater reacts to them the same way. If organisms like the HeavyMetalEater aren't instantiated, then this HeavyMetalEater-dependent relational property of lead and mercury (i.e., the HeavyMetalEater's neuronal response to samples of these two metals) isn't instantiated either, even though the T-transformation (and the notion of the HeavyMetalEater) still exist just like other abstracta.

To summarize, this imaginary case and the argument so far shows that many-to-one mappings between stimulus properties and sensory states is an entirely consistent idea. There is no guarantee that the outputs of sensory systems, or sensory states, correspond to, or are reliably correlated with, some non-disjunctive, causally effective stimulus property. They can be, but they need not be.

Step 4: Actual cases. Akins (1996) gives an illuminating analysis of how this insight raises problems for the standard representationalist accounts of sensation and perception. Akins discusses in detail the case of heat sensation. Evidence shows that, in human heat sensation, there is a many-to-many mapping between external temperatures and heat sensations. Many different temperatures can elicit the same heat sensation, and one and the same temperature can elicit many different heat sensations depending on which part of the body it is applied to, the temperature of the skin, and so on. As Akins argues, the function of human heat sensation can be best understood thus: we sense "narcissistic" temperature properties – that is, we cannot consistently discriminate stable, narrow ranges of temperatures, but rather, we sense properties like "too cold for my head". If we compare this with Dretske's account of sensation – of heat and other properties (Dretske, 1995, Ch. 3), the difference is more than obvious. Dretske likens the

heat sensation of biological organisms to thermometers or pressure gauges. In the case of such gauges it is obvious that particular needle positions correspond to narrow ranges of the property measured. Using this analogy, Dretske offers a picture according to which different heat sensations correspond, in a one-to-one fashion, to different external temperatures (temperature ranges) at the skin. A particular temperature is then identified with the phenomenal (qualitative) aspect of the corresponding heat sensation (Dretske, 1995, p. 84). It seems from Akins's paper that such an account of felt temperature is seriously challenged by the actual facts about heat sensation.

In addition to Akins' analysis of heat sensation, Matthen (1999, 63-64; 2001) argues that the properties that human color vision reliably detects are indeed heterogeneous, though this is not due to metamerism. Metamer sets can be characterized by corresponding non-disjunctive types of reflectance, Matthen contends (see above). Still, physically very heterogeneous properties often look to us the same in color: reflective and emitting surfaces, transparent volumes, holograms, diffraction gratings, volumes that scatter, rather than transmit light (as in the case of the sky and rainbows), and so on. That is, the idea that there must be a fixed number of determinate environmental qualities on which human color vision and color vision in different species converges is wrong (Matthen, 1999, pp. 73-76). In light of empirical data this idea appears simply as wishful thinking. Still, this need not lead us to color irrealism at all. Colors are physically quite heterogeneous properties, but they – the properties that are causally responsible for evoking our color sensations – are perfectly real, physical attributes.

Let me recapitulate what I have been arguing for. If we "correct" stimulus properties by observers' parameters, then, in the first round, we do not get stimulus properties as a result, but rather, we get observers' sensory response characteristics. Then, in the second round, it becomes an empirical question whether the thus-arrived-at sensory response characteristics correlate well with some non-disjunctive stimulus property (victory for the type physicalist) or only with a disjunction of a whole bunch of different properties (defeat for the *type* physicalist). Neither Byrne and Hilbert, nor Tye gave us any convincing reason to believe that sensory quantum catches, cone signals, opponent process responses, or some other neural transform of the color signal correlate well with some non-disjunctive *reflectance* property of surfaces.

2.2.4.2 A distinction: perceiver relativity versus perceiver dependence

Given the discussion so far, we can draw a useful distinction between what we can call perceiver-relative and perceiver-dependent properties. The reason why I wish to make this distinction explicit is that it has been contended that if we allow that object colors are perceiver-relative properties, then we can save type physicalism about color, simply because we can solve the problem arising from metamerism (Hilbert, 1987; Tye, 2000, p. 161; Matthen, 2001). My argument in this section (Section 2.2.4) has been that we must be careful with such proposals since assuming perceptual transformations of reflectance properties can easily lead us into the realm of perceiver-*dependent*, not just perceiver-*relative*, properties. However, perceiver-dependent properties are insufficient to establish type physicalism about color.

Perceiver-relative properties are those that are of interest only for organisms with perceptual systems of a specific kind – human trichromat color vision or bat ultrasound sensors for instance. The object colors that we humans can see are perceiver-relative properties – they are "anthropocentric", "uninteresting", highly derivative properties. However, perceiver-relative properties can be instantiated in the absence of perceivers. What percentage of light a given surface reflects in the 400-525 nm range is a fact about the surface that is entirely independent of any human being around. Such a property remains instantiated in the absence of any human being. In a world with no humans but in other respects like our world, objects have their trichromat-human-relative colors, *physically instantiated*.

Perceiver-dependent properties are importantly different. A perceiver-dependent property is one that cannot be physically instantiated in the absence of perceivers of a particular type. *In this sense*, perceiver-dependent properties are not stimulus properties, but rather, they are perceptual reactions of some sort. Cone absorptions, for instance, are perceiver-dependent properties (of reflecting surfaces).[44] In the absence of humans (or other color-perceiving organisms), cone absorptions and their ratios are not physically instantiated. If they exist in any sense, they exist only as abstracta in a Platonic realm of universals (*if* there exists such a realm). Even if the Platonic realm does exist, there remains the difference in instantiation between merely perceiver-relative and perceiver-dependent properties. In the absence of humans the values that characterize cone absorption ratios, and hence are the same for members of metamer sets[45], exist only as abstracta, whereas light reflection in broad wavelength bands remains physically instantiated.[46]

The key question for type physicalism comes down to this: do conscious color perceptions correspond to perceiver-relative, or only perceiver-dependent properties? Can a substrate for perceived color similarity be found in terms of perceiver-relative, or only in perceiver-dependent properties? In other words, can we find some perceiver-relative surface property (or non-disjunctive C-property: see Section 1.1.3) that characterizes all and only red surfaces? If any relevant, non-disjunctive, causally effective property that characterizes all and only red surfaces turns out to be a perceiver-dependent one, then type physicalism about color cannot be maintained in any form worthy of the name. This is so because there trivially are perceiver-dependent (non-disjunctive, causally effective, etc.) properties that characterize all and only[47] those objects that are reasonably regarded as red: our perceptions as of red (occurring in normal circumstances of perception) are such properties. Or if this is too tendentious, then the corresponding physiological response types of the color vision system (like "positive" activation in the red-green channel, and baseline activity in the blue-yellow channel) are the perfect examples. What makes type physicalism a strong claim is that according to it object colors are non-disjunctive, causally effective properties of distal objects *that are also perceiver-independent* – they are physically instantiated properties of stimuli, not reaction types of perceivers.

2.2.5. Can we save the reflectance theory?

Given the argument in the foregoing sections, reflectance theorists can reply in the following way. First, as I just said, there is a way to avoid category mistakes when using observers' parameters to characterize colors in terms of reflectance. Second, even if

one-to-one mapping between stimulus properties and sensory state types is not necessary, it might, in actuality, still obtain for reflective colors. Even if the relation between integrated reflectances and opponent process signals is nonlinear and highly complex, there may still be non-disjunctive types of reflectance corresponding to metamer sets.[48] This reply is correct as far as it goes, however, it is not enough to save type physicalism. In this section I will address what is correct about this move.

The problem that is being addressed is still metamerism. What I showed so far is that the particular type physicalist solutions offered by Tye, Matthen (and Kuehni) are mistaken. Colors are not *those* types of reflectance – i.e., the ones derived from the simplified opponent process schemas. True enough, colors (at least broad color categories) may still be some other types of reflectance. But what types? We are back to the question asked in the title of section 2.2.

Given what I said about color-matching functions above, we can reformulate our question in the following way, to avoid category mistakes. Given a particular triplet of CIE tristimulus values (say, $X=a_1$, $Y=b_1$, $Z=c_1$), and some standard illuminant with a specified SPD function $E(\lambda)$, *what properties do those surface reflectance curves have in common that, under the specified illuminant, give rise to the tristimulus values a_1, b_1, and c_1*? It is known from empirical data that there are limitations on the complexity of those SSR curves that occur in our natural and artificial environment. All such natural reflectances are continuous functions of wavelength, and they have "smooth" curves, which means that the reflectance changes slowly, never abruptly, with wavelength. These limitations on natural SSR curves are readily explained by microphysical processes that underlie light reflection (Maloney, 1986, pp.1677-1678). These limitations can also be

taken into account when looking for common reflectance properties of metamer sets. After answering the above question, we have to look for generalizations in two directions. First, what are the reflectance properties that render an SSR function into *a range* of tristimulus values (i.e., a particular local area of color space) like the one that corresponds to the color category red? Second, how are such types of reflectance affected by changes in the $E(\lambda)$ function?

As I said at the beginning of the previous section, this is a quite difficult mathematical problem, to solve which one needs to prove some theorems in linear algebra. I will not do this in the present dissertation, but at the end of this section I will briefly review some studies that made promising steps in this direction. Before doing that, however, I present a very simple method that already does a lot to support the idea that colors – at least broad color categories – can be successfully characterized in terms of surface reflectance.

To see the grain of truth in the reflectance theory it is enough simply to look at the SSR curves of some natural and artificial surfaces (e.g., MacAdam, 1997, figures on pp. 36, 38-40). Most natural and artificial red surfaces have a reflectance that is a long-pass cutoff filter between 600 and 650 nm. A long-pass cutoff filter between 600 and 650 nm is a reflectance that is quite low (i.e., less than 6 per cent) between 400 and 600 nm, rises sharply somewhere between 600 and 650 nm and stays high until 700 nm. Though not completely general, this very crude characterization already captures a great number of metamers of different shades of red. Statistically, the vast majority of natural and artificial red reflecting surfaces satisfies this characterization. Similar characterizations can be given for the other seven broad chromatic categories: green, yellow, blue, orange,

purple, yellowish green and bluish green. Orange is a long-pass cutoff filter around 550 nm; yellow is a long-pass cutoff filter around 500 nm. (Again, there can be orange and yellow metamers that don't satisfy this characterization, but statistically the overwhelming majority of actual orange and yellow surfaces, including both natural and artificial ones, does.) Blue is a short-pass cutoff filter around 550 nm. Greens reflect very few at both ends of the spectrum, and a lot in the middle. Just the contrary, purples reflect a lot at the two ends (in, roughly, the 400-500 and the 600-700 nm ranges), but much less in the mid-range (500-600 nm). In sum, there are interesting commonalties between SSRs that look broadly the same in color. These commonalties pop out to the naked eye. Figures 1A to 1G (following pages) show some examples of surface reflectances corresponding to broad color categories.[49]

Figure 1A: Surface reflectances of some red objects. From left to right along the series axis: Red IKEA watercolor, red plastic (Lego block), an autumn tree leaf, and two samples from the Macbeth Color Checker. Notice the common feature: each of the surfaces reflects very little light between roughly 400-600 nm, reflectance rises around 600 nm and stays high until 700 nm.
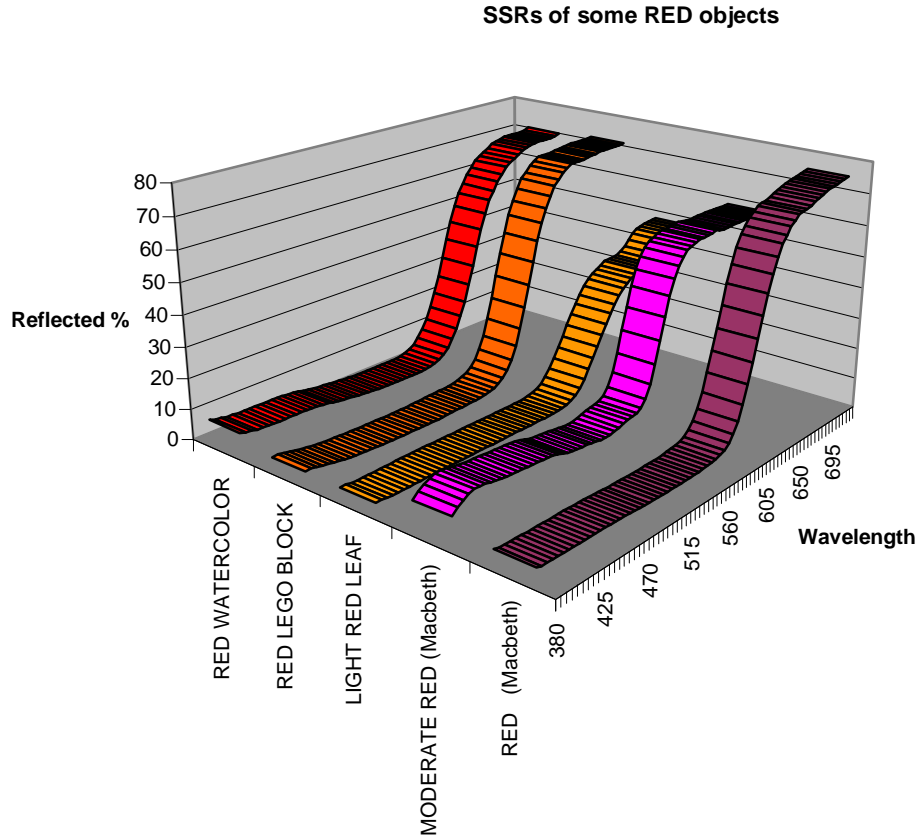
Figure 1B: Surface reflectances of some green objects. Left to right along the series axis: mixture of blue and yellow IKEA watercolor, green plastic (Lego block), a fading green autumn tree leaf, and the green sample of the Macbeth Color Checker. The common feature is that all these surfaces are highly reflective between 485 and 600 nm, and reflectance is low in the rest of the visible spectrum. In the case of the green leaf there is a second rise around 695 nm, but that does not significantly influence the perceived color because the sensitivity of the cones to the 695-710 nm range is very low. Subtractive mixing of blue and yellow gives green (as in the case of the watercolors), but this does *not* mean that green is a perceptual mixture of blue and yellow. Perceptually, (unique) green is neither bluish, nor yellowish – indeed, no color can look both bluish and yellowish at the same time.

SSRs of some GREEN objects

96
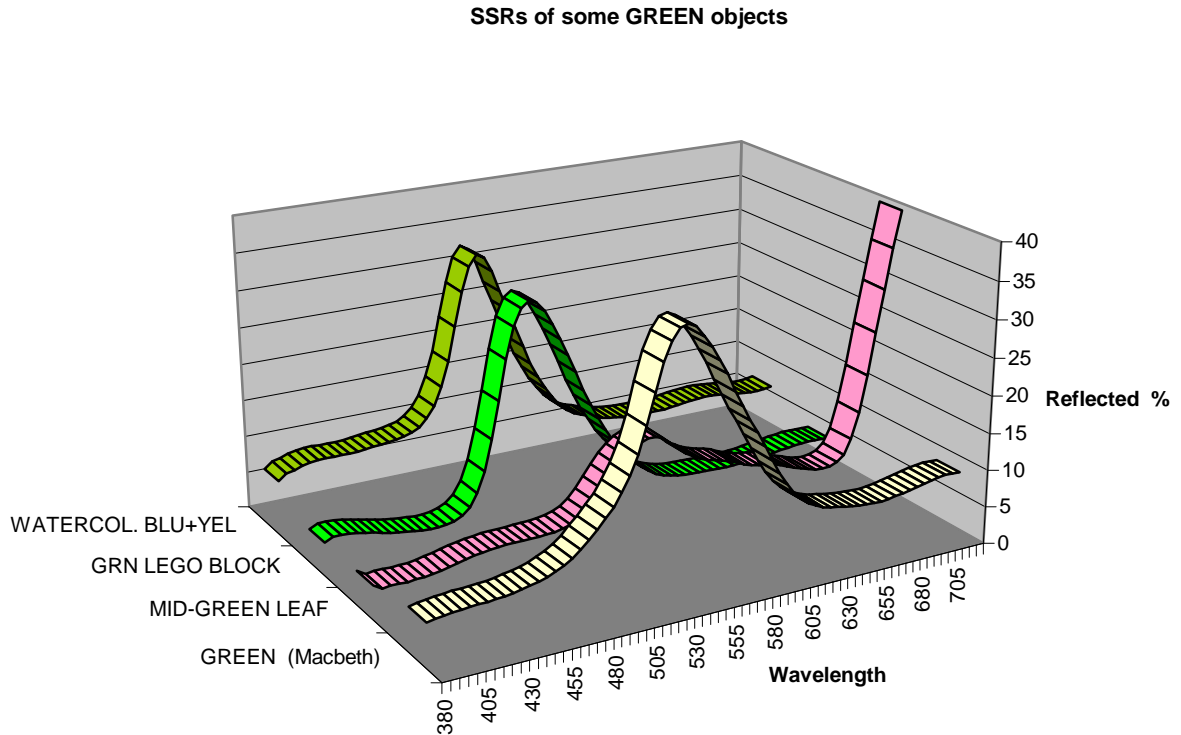
Figure 1C: Surface reflectances of some yellow objects. Left to right along the series axis: yellow IKEA watercolor, yellow plastic (Lego block), an autumn tree leaf, and the Macbeth yellow. The common feature is that all these surfaces are highly reflective between roughly 500 and 700 nm, and their reflectance is low between 400 and 500 nm. In the case of the tree leaf, the gap around 680 nm may have some influence on perceived color (but not much), resulting in a slightly greenish yellow look. If the gap were eliminated, the result would probably be closer to unique yellow.

SSRs of some YELLOW objects
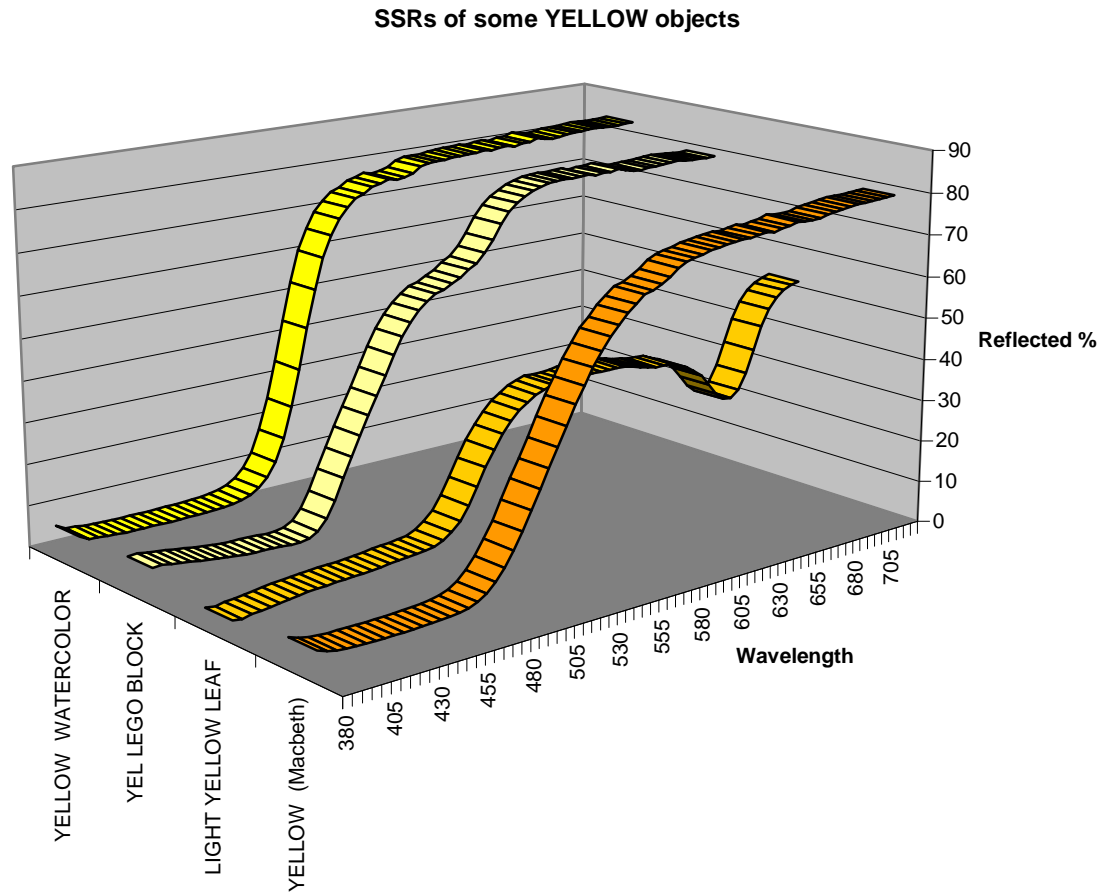
Figure 1D: Surface reflectances of some blue objects: three Macbeth samples from the blue range (blue, blue sky, and cyan), blue plastic (Lego block) and blue IKEA watercolor. The common feature is that reflectance is high roughly between 400 and 530 nm, and low between 550 and 700 nm. Again, the rise starting around 680 nm in the case of the Macbeth blue has virtually no effect on perceived color. If it did, the Macbeth blue would look slightly purplish, but that is not the typical impression of trichromat subjects. The Macbeth sample called purplish blue has a small second peak in reflectance at 665 nm going up to about 12 per cent (see Fig. 1F for this reflectance curve).

SSRs of some BLUE objects
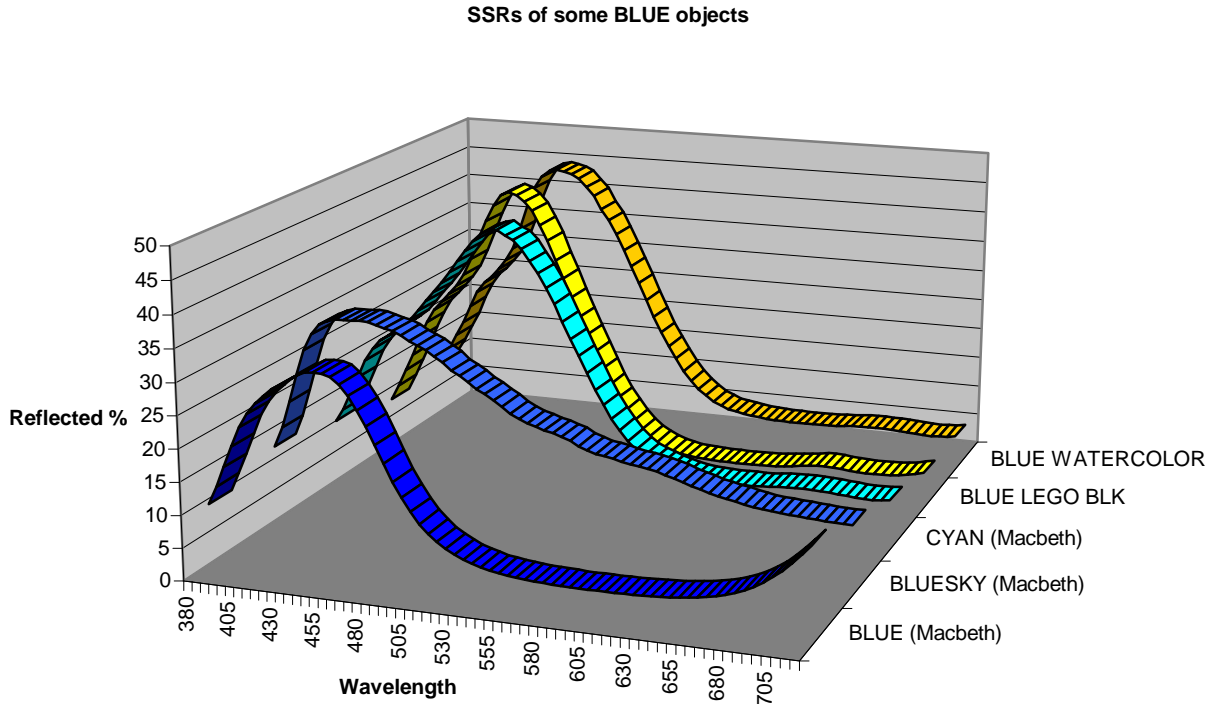
SSRs of some ORANGE objects
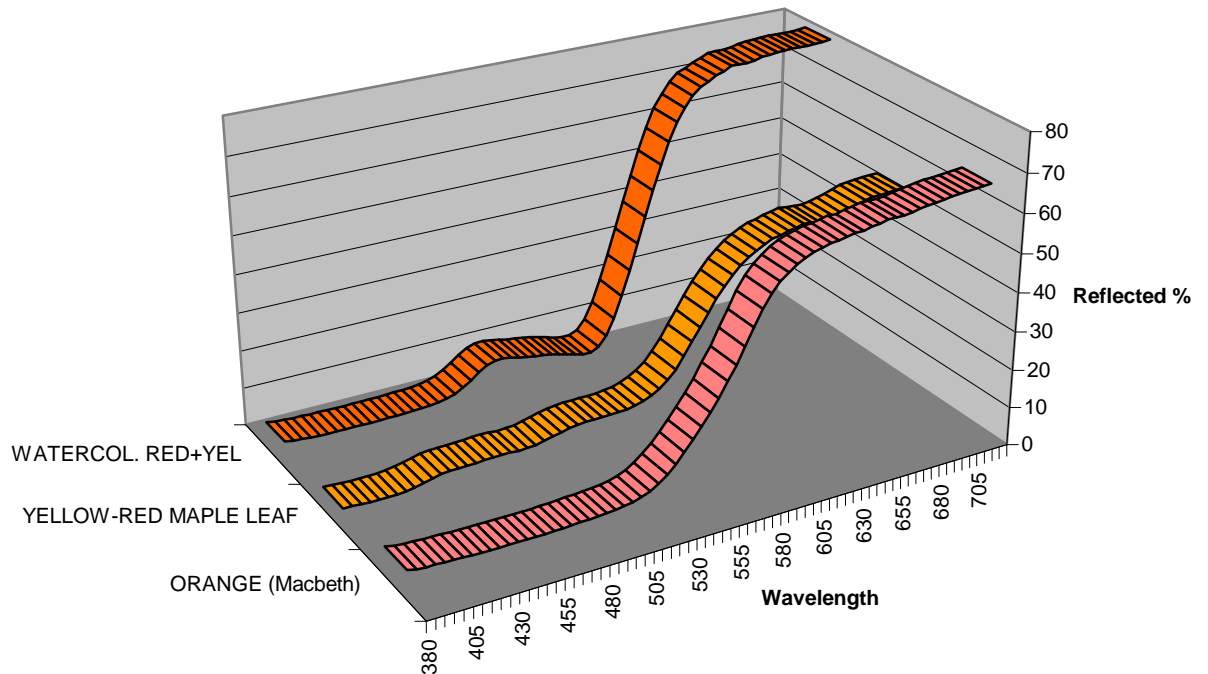
Figure 1E: Reflectances of some orange objects: mixture of red and yellow IKEA watercolors, an autumn maple leaf, and the Macbeth orange. The common feature of these surfaces is that their reflectance rises around 550 nm, stays high until 700 nm, but is low between 400 and 550 nm. Note the characteristic difference between the Macbeth orange and the watercolor mixture. The mixture has a secondary reflectance peak around 510 nm (12,93 %), then it goes down and does not start to rise before 570 nm. The Macbeth orange has no secondary peak but it starts to rise at 515 nm. Such subtle differences in reflectance are not captured in any salient way by trichromat color perception: both these surfaces look orange, though different shades of orange. However, a secondary peak of the same size can result in a change of perceptual color categorization, if it occurs in the right place. The Macbeth purplish blue (Fig. 1F) would not look purplish but just blue if the secondary peak in its reflectance (around 665 nm) were eliminated.

SSRs: PURPLE colors
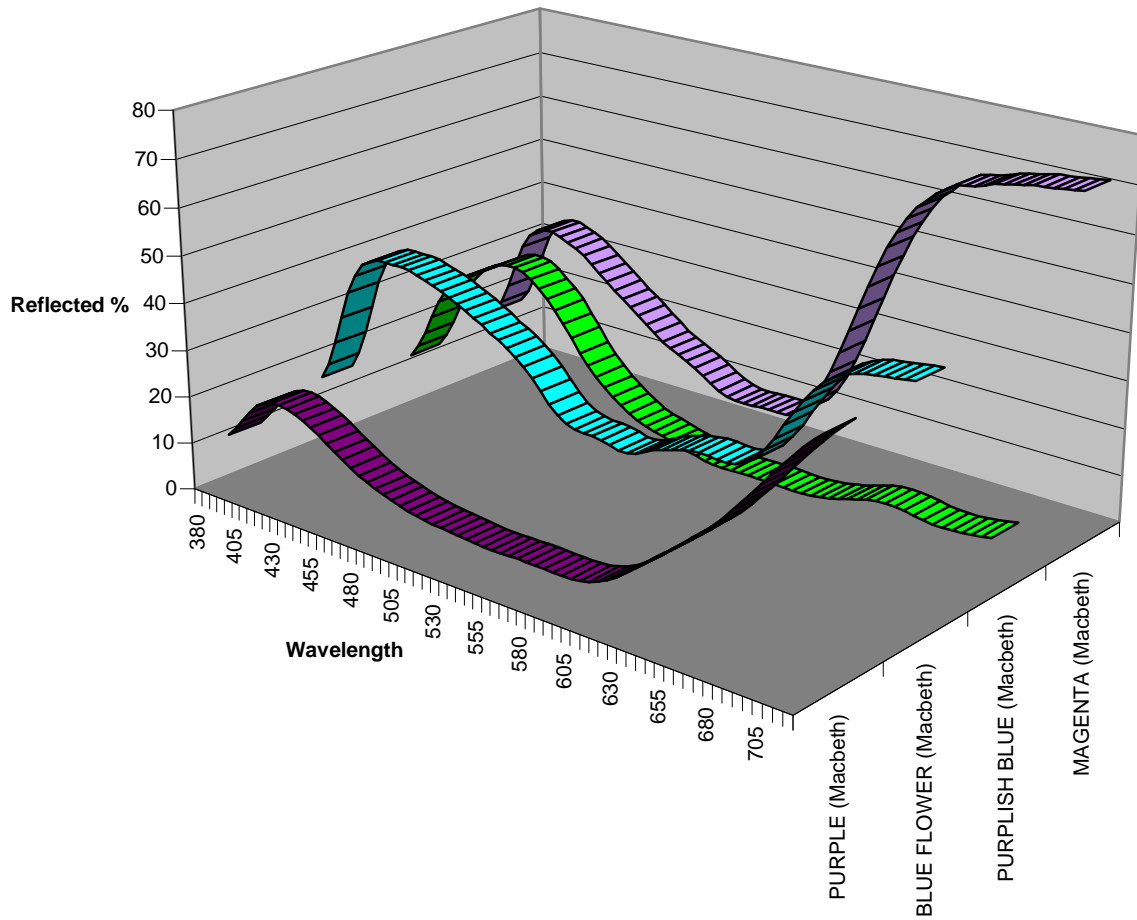
Figure 1F: Reflectances of four Macbeth samples from the purple range: purple, blue flower, purplish blue and magenta. The common feature is that reflectance is high at both ends of the spectrum, but low in the middle. Note the purplish blue curve (third from the front) that was referred to above.
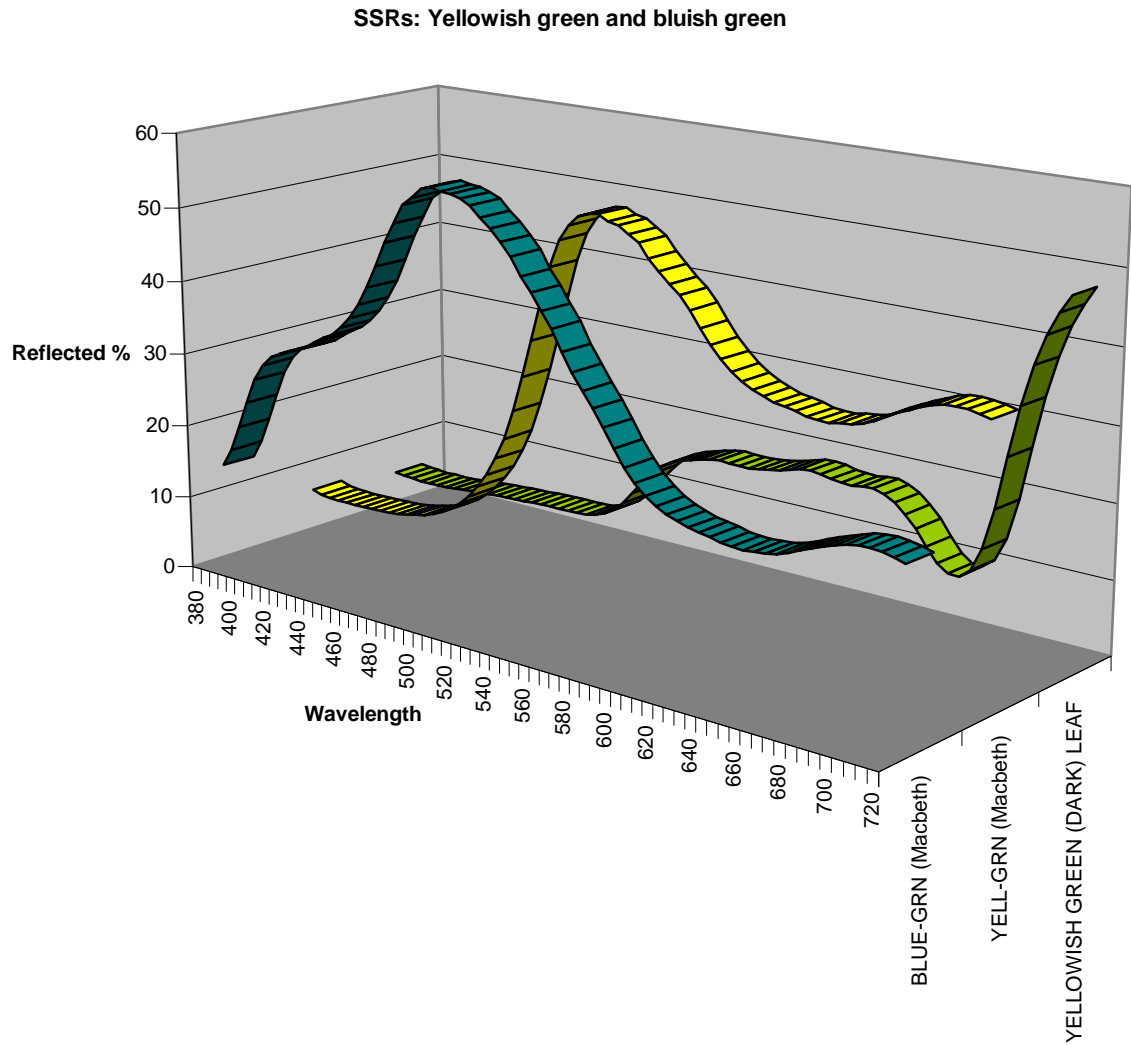
Figure 1G: reflectances of yellowish green and bluish green objects: two samples from the Macbeth Color Checker and an autumn tree leaf. For bluish green objects, the common feature is that their reflectance is high between 400 and 600 nm and relatively low above 600 nm. Yellowish green objects have reflectances that are relatively low between 400 and 500 nm, high between 500 and 700 nm (similarly to yellow surfaces), but typically, their reflectance is even higher in the middle wavelength range than at the long end.

On the other hand, how these commonalties in reflectance-families relate to the simplified opponent process model presented in Hardin (1988) is less than immediately obvious. For instance, as I mentioned in section 2.2.3.1 above, for unique red and unique green surfaces the blue-yellow channel should be in balance (L+M-S $=_a$ 0), but from this it does *not* follow (*pace* Byrne and Hilbert, 1997, and Tye, 2000) that red and green surfaces reflect approximately as much light in the short wavelength range as in the other two ranges together. Indeed, nothing is farther from the truth: both reds and greens reflect much less light in the short wavelength range than in the other two ranges together. The surface that Tye's schema would categorize as, say, red would be, in reality, purplish blue (predominantly blue, to some extent reddish). And this is far from being the only problem with the schema. Moral: the simplified opponent processing model is not at all useful to solve the problem of metamerism – probably because it is (too) simplified.

Still, the suggestion of the above perceptual test is correct as far as it goes. I think if we limit ourselves to reflecting surfaces and broad perceptual color categories, then we can specify non-disjunctive types of reflectance that uniquely correlate with our color perceptions. If, however, one wants to go beyond the first perceptual impression, the above-mentioned mathematical treatment needs to be embraced. Here are some steps in that direction.

Maloney (1986, p. 1680) suggests that the function of cone spectral sensitivity curves is to low-pass filter natural reflectances, thereby helping the color vision system to represent specifically some low-pass component of surface reflectances, ignoring information about higher-frequency variations in the SSR curves. If this suggestion gained further support, then at least one key transformation by observers' parameters

(weighting the color signal by cone spectral sensitivities) would be proven to actually give access to some external property (the slowly varying component of surface reflectances). Maloney's proposal is made in the context of color constancy: he seeks reflectance features that are recovered by color vision despite certain limited variations in lighting – especially variations in daylight.

It has been shown (Further, Finlayson and Morovic 2000a, 2000b) that for two reflectances to be metameric to the human visual system, they have to have at least three crossovers in the visible spectrum (sometimes more, but at least three). Further, Finlayson and Morovic showed that these three crossovers, *for all members of all metameric sets*, tend to occur (statistically) around 450, 540 and 610 nm respectively. The reason why metamer crossovers concentrate in these three narrow wavelength bands is that these bands correspond to the peak sensitivities of the three cone types (Finlayson and Morovic, 2000b, pp. 13, 14).

This result might give us a way to seek out distinctive features of *particular* metamer sets like red, green, etc. metamers. For instance, one could hypothesize that both red and green metamers have crossovers (as according to the proposal) at the three specified wavelengths, but the 540 nm crossover occurs at level p_1 for red metamers, and level $p_2 \neq p_1$ (presumably $p_2 > p_1$) for green metamers; similarly for the other two crossovers. That is, all red metamers have a reflectance of p_1*100 % at 540 nm, whereas all green metamers have a reflectance of p_2*100 % at 540 nm. Moreover, in this particular example most likely p_2 is greater than p_1, since at 540 nm (which is in the mid-wavelength range) all green surfaces reflect more light than all red surfaces – just the

opposite for 610 nm. This is a so far unexamined extension of the Finlayson and Morovic proposal, that could also be tested.

Finally, note that these promising accounts of the underlying commonality in metamer sets do *not* proceed from the opponent process theory. All they use is the model of the very first stages of color-information processing: the one that Wandell (1995, Ch. 4) calls "wavelength encoding". This level of color processing gives us an understanding of color-matching, but not that of color appearance (Wandell, 1995, pp.100-101). Color matching and color appearance are quite independent of each other: stimuli that match in color to a subject, can change color appearance while continuing to match (in *entirely* normal circumstances of perception). For instance, when a subject looks at two metameric red patches beside each other, each presented against the same mid-gray background, the two patches look the same in color. Leaving the illumination unchanged, but changing the background from gray to, say, bright yellow, will change the color appearance of the two red patches, still, the patches will continue to look *indistinguishable in color* to the subject. Simultaneous contrast effects and opponent process coding play a key role in determining color appearance, but they operate at higher levels of color processing than those responsible for color matching. Metamerism and color matching, as modeled in color science (i.e., as a product of the color matching functions alone, interacting with the color signal), are prior to, and can be treated independently of, those higher-level processes. This might be another reason why the simplified opponent process theory is so unhelpful in solving the problem of metameric plurality.

Chapter Three: Colors that are not reflectances

2.3. Beyond reflective stimuli: can we generalize the reflectance schema?

Once the problem of defining colors in terms of reflectances is solved, type physicalism is faced with the following problem. Can we extend the reflectance theory of color to emitting surfaces, volume and film colors, fluorescent surfaces and other color stimuli? In order to be able to maintain that *being red* is a non-disjunctive C-property and also maintain that stoplights, hot irons, strawberry juice, red fluorescent plastics, and so on are genuinely red (i.e., they are red just like ripe tomatoes[50]), what we have to say is that color is some physical stimulus property P that is not just reflectance, but a more general property of which reflectance is a special case. This more general property should apply to at least volume and film colors, emitting surfaces, and fluorescent ones. In this section I shall assess the prospects of such a generalization focusing on these four categories: reflecting, transmitting, fluorescent and emitting surfaces. As I shall argue below (in Section 2.4), these are basic kinds of color stimuli that any view of object color has to admit as *genuinely* colored, not as displaying only illusory colors.

It is typical among defenders of type physicalism to postpone this issue to some future paper, and formulate the theory exclusively for reflecting surfaces (Hilbert, 1987, though see pp. 132-134; Byrne and Hilbert, 1997, p265; Tye, 2000, 159-162). Proper generalization of the reflectance theory into a full-blown type physicalist account is, at the moment, an outstanding promise: no remotely plausible account has so far been proposed to solve this problem. I know of only one such attempt that I will discuss in detail in Section 2.3.3.

I think this reluctance on the part of type physicalists to address non-reflective colors is no accident. As I see the problem, it is possible to extend the type physicalist account to volumes and films. However, when it comes to fluorescence, metaphysical worries arise as the vagueness of the common property that applies to all and only objects with the same color (including reflecting, transmitting, and fluorescent objects) reach a remarkable level. When at last we hit emitting surfaces, it turns out that it is utterly hopeless to find a causally effective, non-disjunctive property had by all objects that look to us the same in color (in some broadly normal circumstances).

In this section, I start with what I see as the easy case: volumes and films. Then I continue with what I think is the ultimately hopeless case for type-generalization: emitting surfaces (or light sources). In discussing light sources I first provide a general argument why light emission and reflectance together will never fit into the type physicalist view of color. Then I consider in detail Hilbert's sketchy proposal to generalize the reflectance theory of color to emitting surfaces (Hilbert, 1987, pp. 132-134). Finally, I address the case of fluorescent and phosphorescent surfaces, and suggest that a Wittgensteinian family resemblance view of object color is much more plausible than a natural kind view.

2.3.1. Generalizing to transparent objects and filtering

As a first step of generalization to include volumes and films, one can try to say that color is not simply reflectance, but rather a disposition of objects to *filter* the incident light in certain ways. Filtering can take two different forms: reflective filtering and transmissive filtering (the latter applies to volumes and films; reflective filtering is

reflectance). This move gives one a broader concept that includes both reflecting surfaces and transmitting ones.

Note, however, that transmission and reflection are underlain by quite different microphysical processes, so the unification of the two under the concept of filtering is strictly functional, that is, abstract. For instance, metals act as reflecting surfaces. They absorb visible light of any wavelength, due to the almost continuous scale of excited states of their electrons. Still, metals are not black but gray or white (or shiny) because most free electrons that absorb a photon and jump to an excited state immediately reemit a photon of the same energy and return to their original energy level (Nassau, 1997, p. 19). In most metals such absorption and reemission is uniform at any visible wavelength. In the case of gold, copper, or alloys like brass, some wavelengths are absorbed and reemitted more efficiently than others. On the other hand, when light is transmitted through a solid or liquid medium, no absorption and reemission happens. The light waves that make their way through the medium do not get absorbed, but the ones that are filtered out do. The passing wavelengths are subject only to refraction – a change in speed and direction of the light transmitted. Refraction does not include absorption and reemission, just an interaction between the electromagnetic field of the light radiation and the electric charges of the electrons (Nassau, 1997, p. 24).

For the idea of filtering to work in the present context, two conditions have to be satisfied. First, one has to accept that the more general phenomenon described by the concept of filtering, just like reflectance alone, can constitute a natural kind essence of some sort. For objects with the same color to constitute, by virtue of this very fact, a natural kind, the key property that makes objects colored has to be a natural kind essence

of some sort. It is plausible to hold that **if** types of reflectance are natural kind essences of some sort, **then** types of transmissive light filtering are, by the same coin, natural kind essences as well. Perhaps the possibility of such an extension is the reason why Tye (2000, p. 147) says that films and transparent volumes are also colored. However, he does not even hint as to how to generalize the reflectance theory to transparent volumes.

Second, for such a generalization one needs some systematic, non-arbitrary correspondence between reflectance functions and transmittance functions in terms of the color perceptions they elicit. Mathematically, these functions are of the same kind: they both arise as a ratio between the SPD of some broadband illuminant and the SPD of the non-absorbed (reflected or transmitted) component. The SPD of reflected/transmitted light is the numerator, whereas the SPD of the illuminant is the denominator. As a consequence, both reflectance and transmittance functions map values in the 400-700 nm range to values in the 0-1 interval. The most straightforward correspondence relation between these two groups of functions would be identity: for instance, ripe tomatoes are red because they tend to reflect light dominantly in the long-wavelength range (and absorb the rest); strawberry juice is red because it tends to transmit light dominantly in the long-wavelength range (and absorb the rest). That is, the reflectance function of ripe tomatoes and the transmittance function of strawberry juice are roughly the same in shape. Such a correspondence relation is pretty much the case. In general, if a film **F** has a transmittance curve T, and a reflecting surface **S** has a reflectance curve R such that T and R are identical in shape (i.e., mathematically the transmittance function $T(\lambda)$ and the reflectance function $R(\lambda)$ are identical), then **F** and **S** will look essentially the same in color.[51] Film transmittance and reflectance are both measured by spectrophotometers, and

transmittance functions in these measurements arise from placing the colored film in front of a white standard that reflects back most of the incident light at any wavelength between 400 nm and 700 nm.

2.3.2 The problem of emitting surfaces

When characterizing the color of reflective surfaces, type physicalists identify colors with types of reflectance – a light-disposition (i.e., a disposition to reflect light in certain ways: see Johnston, 1992). Three favorable features of this identification are as follows. First, object colors are inherent, and also invariant properties of surfaces – properties that do not change with changes in illumination, in harmony with the phenomenon of color constancy, and our common-sense intuitions about color. Second, in a derivative, but still explanatorily interesting sense, colors thus construed are causally effective, simply because the manifestation of reflectance, namely the physical event of light reflection, is causally effective. Third, the colors of objects are retained in darkness and when they are not seen.

How about emitting surfaces (or light sources)? Do they have a property that satisfies these three requirements and that is also causally responsible for our color perceptions? Well, they do: *the actual physical event* of emitting light with a certain spectral power distribution (SPD) is such a property. The event of their emitting light with a specific SPD is an attribute (property) of light sources that is causally effective, it is retained in darkness (and when these objects are not seen). It is also an inherent characteristic of the objects at particular points in time, an attribute that is independent of changes in illumination.[52] Therefore, it seems reasonable to identify color, in the case of

emitting surfaces, with the physical event of emitting light. In a wide range of circumstances, this attribute correlates pretty well with our color perception: the same emission SPD results in approximately the same color perception in a wide range of circumstances. Think of stoplights, or turn signals of cars: they look red and yellow respectively, in a wide range of normal perceptual conditions (at noon, at sunset, at night, etc.). Also, shift from fluorescent tubes to illumination by daylight in your room, and the colors on your computer monitor will remain essentially unchanged. Slight differences in perceived color are present of course: a brake light might look slightly different in color in the noon daylight and around sunset. Headlights of cars also look slightly yellowish in daylight, and more whitish at night. However, this should not be a big problem as color constancy is only approximate even for reflecting surfaces (Wandell, 1995, pp. 314-315; Fairchild, 1998, pp. 156-157). There definitely are slight changes in the perceived color of one and the same reflective surface under different, broadly normal illuminants, but these changes are effectively masked by the vagueness of our color memory (Raffman, 1995, pp. 294-295; Tye, 2000, Ch. 1, p. 11).

Fine so far; emissive color is the physical event of surfaces' emitting light. Now, think of a ripe tomato, and a piece of red-hot iron. A perhaps better example is the case when one sits in front of a computer monitor and adjusts the RGB signal to match the color of the screen with one's T-shirt. If a match is achieved, then the screen looks the same in color as the T-shirt. In this case, do the surface of the T-shirt and that of the screen *share* an inherent, causally effective property that they retain in darkness? They seem not to. The inherent attribute of emitting surfaces that best correlates with their perceived color, and is also specifically causally responsible for evoking color

experience, is *the actual physical event of emitting light with a certain spectral power distribution (SPD)*. In other words, light emission is a process that takes place in space-time. As we saw above, the relevant color property of reflecting surfaces is their reflectance – a disposition. Dispositions are not physical events.[53] According to the now dominant, functionalist account of dispositions, a disposition **D** is a functional state, or a multiply realizable abstract causal role. To have **D** is to have some ordinary physical state that endows its bearers with the required causal role (i.e., the one that *is* **D**). On the accepted view, **D** is *not* identical with its base – the accidental physical property that endows its bearers with **D**. Moreover, even though the manifestation of **D** (some physical event) is essential to being **D**, **D** is not identical to its manifestation either. Brittleness is a disposition to break (on being struck), but it is not the actual event of breaking. Hence physical events and dispositions to produce such events belong to two different ontological categories. (For more on events, see Davidson, 1970; Goldman, 1970; Kim, 1976). For this reason, the "gap of disjunctivity" inevitably opens for object color: redness is *either* the event of light emission of some sort, *or* a reflectance of some sort. This seems to exclude views on which the property of being red is a natural kind essence.

In order to bridge this gap, one might try to define reflective color in terms of physical events, or, alternatively, emissive color in terms of dispositions. Alas, neither move works. Saying that reflective color is the actual physical event of light reflection leads to counterintuitive consequences. These are: (1) objects would not retain their color in darkness, (2) object color would not be an invariant, illumination-independent property of objects, but rather one that varies with any variation in illumination, (3) due to our

limited perceptual color constancy, actual light reflection (the SPD of light reflected by particular surfaces) would not correlate well with perceived color.

Trying to understand emissive color in terms of dispositions does not work either. Saying that emissive color is not the physical event of emitting light, but it is a *disposition* to emit light, or emittance, will not do because emittance in the case of hot iron or computer monitors is not a *light*-disposition (i.e., a disposition to reflect certain portion of the incident light), whereas the reflectance of the tomato is. Worse still, cold iron *is* disposed to emit long wavelength light, given that it is heated up to about 800 ºC – but cold iron is not red.

A last attempt to maintain the type physicalist view might be to note that the relevant wavelength range of light involved is a commonality in the case of, say, red objects. I.e., red objects *either are disposed to reflect, or actually emit,* **long wavelength light** (light between 600 and 700 nm) – but no shorter wavelengths. Since it is a commonality in all and only red objects that they interact – in one way or another – with light of long wavelength, the type physicalist proposal could be that the color red just is light of 600-700 nm wavelength.

To identify redness with some wavelength range of light would mean that object color is no longer a property that inheres in the distal object of perception – the one that interacts with incident light. Object color, on this view, becomes a property of the proximal stimulus – the light passing between perceived objects and our retina. However, Tye is right to say that colors appear to us to be inherent attributes of surfaces and volumes – the distal objects of perception (Tye, 2000, pp. 147, 153). Arguably this is also the common sense view of color. To say that it is the lights that have the colors is to go

against this intuitively very plausible view. Moreover, if it is the lights that are colored in the first place, then we simply do not see the things that are colored in the first place – simply because we never see lights themselves, only objects and surfaces that interact with light (Hilbert, 1987, p. 133). Another, already discussed problem is that perceived color is not variable in the way that the SPD of light, reflected from surfaces is (Wandell, 1995, pp. 314-315; Hilbert, 1987, pp. 64-65; Matthen, 1988, pp. 8-9; 1999, n33, pp. 64-65). Finally, the evolutionary function of color vision is to detect, discriminate, and recognize, distal stimuli like objects, surfaces – or heliocentric directions in the sky like pigeons do (Matthen, 1999, pp. 60-61), that is, all kinds of different things that interact with incident light. However, it is not the function of color vision to discriminate *lights themselves* by their wavelength. We do not know of any evolutionary advantage that such a discrimination could have conferred on different organisms, but we do have a clear idea of how discriminating distal visual objects by the aid of color vision enhanced fitness in our ancestors.

For these reasons color physicalists are united in repudiating the idea that object color is a property of the proximal stimulus (light coming from the object to the retina). Therefore the problem remains: hot iron and ripe tomatoes exhibit two entirely different relations to the same wavelength range of visible light (i.e., one actually *emits* it whereas the other *is disposed to predominantly reflect* it) – so again they seem not to have any inherent, invariant, non-disjunctive, causally effective surface property in common that could be identified with their redness.

2.3.3 Hilbert's proposed solution to the problem of emitting surfaces

2.3.3.1 The proposal

Hilbert (1987, pp. 132-134) attempts a generalization of his reflectance theory of color to emitting surfaces along the following lines. At the start, he makes two remarks. First, emitting surfaces may also reflect some light, not just emit it, but typically the reflected component is of little importance in determining the color of these surfaces. Second, when considering emissive color, it is important to note that color is still the property of the surfaces that emit light, not the light rays themselves that leave the object. Light rays are invisible; only objects or surfaces from which light arrives at the retina are visible. The core of Hilbert's proposal is that it is still possible to compute a ratio between the light leaving the emitting surface and the illuminant. The numerator here is the sum of emitted and reflected light; the denominator is the illuminant SPD. Of course, this ratio may exceed 1, or 100 per cent, at the wavelengths at which the surface emits light – this is a difference from ordinary reflecting surfaces.[54] Still, we have an extension of the reflectance concept to emitting surfaces, or so Hilbert claims.

There is an apparent problem with this extension. For emitting surfaces the emitted light is typically independent of the external illumination, and the emitted component in the numerator is typically much larger than the reflected one. Hence this ratio will vary heavily with variations in illumination. (Obviously not so for reflectance.) If this ratio describes the color of emitting surfaces, then emissive color, in contrast with reflective color, is also heavily illumination-dependent. But this observation in fact corresponds to how our color perception works, claims Hilbert: for instance, the flame of a gas stove appears bright blue in the light of an incandescent lamp, and nearly invisible

in daylight. This correspondence shows that the just proposed account of emitter color is on the right track, or so Hilbert concludes.

2.3.3.2 Critique

Now let us see if there is a genuine problem with Hilbert's attempt to capture emitting surfaces in a generalization of his reflectance theory of color. Unfortunately, there is. As we have seen, the proposal is that object color in general is emission (E) plus reflection (R), divided by the external illuminant (I). In other words, object color in general is the color signal divided by the external illuminant: (E+R)/I, or what is the same, E/I+R/I.

The first observation might be that whereas the division R/I reveals an inherent property of surfaces (namely their reflectance), the proportion E/I does not describe any such property. R/I is reflectance, and it is agreed that this is an important functional property of surfaces – an invariant attribute of them that can be measured. Here the denominator (I) varies independently of any surface property, and the numerator, for any particular surface, shows a variation that correlates with variation in I. Covariation results in a constant proportion, one that characterizes an invariant property of the surface in question, namely its reflectance. E/I, on the other hand, does not express any invariant (illumination-independent) property of surfaces, exactly because E does not covary at all with I. E/I will vary as I varies. As Hilbert himself admits, (1987, p. 134), on this account, *color in general* is not an invariant, illumination-independent property of surfaces. In some cases it is (e.g., in the case of reflecting surfaces), in others it is not (in the case of emitting surfaces). Since R/I expresses an invariant property of surfaces

whereas E/I does not, the sum of the two will not express any illumination-independent, invariant surface property either.

Now, if we return to the key question: "Is object color in general a natural kind essence of some sort?", the answer is this. Once again, we have failed to find any inherent, non-disjunctive, causally effective surface property that is reliably correlated with our perception of redness in broadly normal circumstances. Indeed, I think the mathematical addition in the R/I+E/I formula does nothing more than conceal true disjunctivity.[55] Saying that object color is the sum of reflectance plus *emission divided by the illuminant*, either of which can play no role in determining our color perception in any particular case, amounts to saying that object color is either reflectance, or emission – that is, a disjunction of these two factors.

One could try to reply to this objection in the following way. To the question "What is the theoretical interpretation of the (E+R)/I ratio, (i.e., the ratio of the color signal and the external illuminant)?" the answer should be, "Well, what this ratio expresses simply is object color, in full generality". For this answer to make sense, the (E+R)/I ratio should correlate reasonably well with our color perception of objects. If E=0, (i.e., for reflecting surfaces) then there is an interesting, though by no means perfect, correlation of this kind. If E>0 (emitting surfaces), the correlation is poorer, if it exists at all. Practical color science uses the R/I ratio and the corresponding concept of reflectance extensively, whereas the (E+R)/I ratio is ignored.[56] However, there are some cases that are, *prima facie*, well explained by the (E+R)/I ratio. One is Hilbert's example of the gas stove flame (see above and Hilbert, 1987, p. 134). Let me briefly consider this phenomenon, and at least another, well-known chromatic effect, that is also, *prima facie*,

well explained by Hilbert's proposal. Then I will point at another problem with the proposal which rules it out entirely.

When bright sunlight falls on an active color TV monitor, the colors in the picture tend to fade out, or become entirely invisible (the active monitor looking gray, similar to inactive ones). This is because the amount of light reflected by the monitor's surface far exceeds the amount of light emitted by it. Since, in such a case, our eye adapts to the bright sunlight, we will hardly notice the emissive color of the active monitor. The reflective component dominates, and its SPD is an even distribution of relative energy, typical of white or gray surfaces. This explains how the gray look arises. Moreover, obviously, this fadeout phenomenon is predicted by the $(E+R)/I$ formula. In the bright sunlight E becomes negligible, so $(E+R)/I$ becomes R/I, and it characterizes the reflectance of the monitor's typical gray reflecting surface.

There is also a chromatic effect that is, prima facie, well explained by Hilbert's proposal. This is the phenomenon that the light of tungsten bulbs looks yellowish in noon daylight, whereas at sunset it looks white, or even slightly bluish. In this case R is negligible (the tungsten filament emits light that is too strong to be overridden by reflection of the incident light), so $E+R/I$ becomes E/I. Very roughly, the SPD of noon daylight is an even distribution of relative energy, whereas the SPD of sunlight at sunset is strongly biased toward the long-wavelength part of the spectrum. The SPD of tungsten bulbs is also biased toward the long-wavelength part of the spectrum, but not as strongly as that of daylight at sunset. If we divide the SPD of the bulb by that of the noon daylight (i.e., take the E/I ratio in that case), then the resulting distribution will still be biased toward long wavelengths, and that explains the yellowish look. If, on the other hand, we

divide the SPD of the bulb with that of daylight at sunset, the resulting distribution will be either roughly even, or it will be slightly biased toward the short wavelengths. That explains the whitish or bluish look.

However, there is a bigger problem for Hilbert's proposal, one that is decisive. Object color cannot be (E+R)/I for the following reason. Take an emitting surface S with constant emission intensity E_0. When I, the external illumination, approaches zero, $(E_0+R)/I$ will approach infinity (at any wavelength at which S emits light). That is, if color in general is (E+R)/I, then, on Hilbert's proposal, the brightness of any emissive color is infinite in the absence of external illumination. But of course, when we perceive, say, a firefly at dusk, or at a very dark night, it does not look infinitely bright, nor is it infinitely bright in any sense of the word.[57] Therefore, the (E+R)/I formula, as a generalized descriptor of object color is hopelessly wrong. It fails in exactly the case that it was intended to capture: emitting surfaces. [58]

2.3.4. Fluorescent and phosphorescent objects

Arguably, fluorescent color is a third class of disjuncts of object color. Flourescence is an illumination-dependent disposition to emit light (i.e., to absorb light at one wavelength, and, as a result, emit light at another, typically longer wavelength). Some fluorescent surfaces absorb wavelengths in the ultraviolet range and then emit visible light, whereas others absorb light in the 400-500 nm range and emit light at a longer wavelength. This disposition comes mixed with ordinary reflectance (fluorescent surfaces also reflect light). As fluorescence is some combination of reflectance and a disposition to emit light, it does not fit into our ordinary concept of filtering. Some idea

of "negative filtering" – i.e., that, at some wavelength, and in terms of radiant energy, more light leaves the surface than what's present in the illuminant, even though true, does not capture the phenomenon of fluorescence properly. Fluorescence is *transforming* light energy: given input at one wavelength, output occurs at another (and adds to the reflected component). Note also that, similar to the case of reflection and transmission, such a "conceptual unification" of the reflective-plus-transmissive and fluorescent component is strictly functional, as in terms of microphysical processes the fluorescent and the reflective/transmissive processes are entirely different from one another – they are also completely separable (Nassau, 1997, pp. 11, 13). Rubies can be either fluorescent or non-fluorescent; emerald differs from ruby in the transmissive color component, but the two have the same fluorescent component. (Emerald has green or bluish green transmissive color, but if it has a fluorescent process, then that gives rise to red emission, just like in ruby.) In general, the fluorescent process is the generation of visible light that is not present in the incident light from some other source of energy (visible light of some other wavelength, or ultraviolet waves), whereas what happens in transmission is simply that the transmitted wavelengths are those that are present in the incident light and left unaffected[59] by the transmitting medium (see Nassau, 1997, pp. 10-13).

Temporal synchrony is an important feature as well in defining fluorescence: fluorescent surfaces emit light due to wavelength transformation *in exactly the time window in which they are illuminated*. That is, they do not retain emission in darkness, in contrast to phosphorescent surfaces. The latter accumulate energy from illumination, start emission, and maintain it even when external illumination ceases. So, phosphorescent surfaces are, in an important respect, similar to emitting surfaces. However, whereas

phosphorescent surfaces emit light as a result of being illuminated, many other emitting surfaces like hot iron, or computer monitors, emit light as a result of some other influence.

Many surfaces that we think of as reflective ones in fact produce some fluorescence as well. White shirts are a typical example: after being washed several times, white shirts turn yellowish as they reflect slightly less light in the short-wavelength range than in the rest of the spectrum. Whiteners counterbalance this by adding a little flourescence in the short wavelength range (Nassau, 1997, p. 18). Fluorescence is similar to reflectance in that both the component of light that is emitted by fluorescent surfaces and the one that is reflected are dependent on the illumination. This makes the division of *emitted plus reflected* light by the illuminant a sensible transformation. (Remember section 2.3.3 above: just the opposite is true of emitting surfaces.) Another feature worth noting is that, in the case of fluorescent surfaces, the emitted component, as compared to the reflected one, is not as predominant as in the case of most emitting surfaces (e.g., a light source in an otherwise dim room, or an active computer monitor).

2.3.5. Conclusion: family resemblance rather than natural kinds

What emerges from these cases is a Wittgensteinian family resemblance picture of object color rather than a natural kind view. Redness in objects is more similar to games as Wittgenstein sees them than to water or gold as Kripke sees them. There is a whole variety of different physical properties that all give rise, in ordinary circumstances of perception, to the same color appearance (for more examples see Hardin, 1988, Thompson et al. 1992, Nassau, 1997, and Matthen, 1999, pp. 63-64).

I am sure that there are type physicalists who will want to contest this result. Here is a little more reflection on what such a contention would amount to. Let us return to the problem of generalization for a moment: how could we generalize the filtering concept of object color (the one that works for reflective surfaces and transparent objects) to fluorescent surfaces?[60] In other words, what would be the functional common core of filtering and fluorescence? One way to go would be to say that both filtering (reflectance plus transmittance) and fluorescence together unite under the functional notion of light-transformation, or so: what both ordinary filtering surfaces and fluorescent ones are disposed to do is to transform the SPD of the incident light into that of light leaving the object on being illuminated. In the case of ordinary light filters (without a fluorescent component), this transformation includes no between-wavelengths energy transfer, whereas fluorescence is exactly this sort of transfer. So, object color, in general, is *the disposition to transform the SPD of the incident light into that radiated by the object interacting with the incident light*. A simpler way to put the same point would be to say that ordinary filters and fluorescent materials are all *disposed to produce radiation* in the visible range when illuminated either by visible light or by ultraviolet light. Note that 'radiation' here includes not just light emission and reflection, but also what happens when transparent filters allow certain wavelength bands of the incident light to pass through them. So, object color simply is the objects' disposition to produce radiation in this very general sense (including zero radiation as with totally black objects) as a result of external illumination. This move still does not capture emissive object color, since, as I argued, emissive colors are not dispositions of any sort. But perhaps one can force one's

fantasy a little further and stretch the type physicalist notion to include emitting surfaces under an even vaguer "functional umbrella-notion".

A philosophical worry that this line of reasoning might raise is that what's going on here is simply a Quinian-Goodmanian similarity-finding exercise. As Quine and Goodman famously observed, given almost any pair (perhaps even n-tuple) of entities, one can always find, with a little fantasy, some property – either a physical attribute or an abstract one – that is true of both (all) members of it.[61] These observations are sometimes used to support antirealist approaches to metaphysics. Goodman (and others like Kuhn) argued that what properties, similarities, and types we identify (find as existing) in the world depends essentially on how we conceive of the world, or how our cognitive systems are set up (Goodman, 1965; Kuhn, 1974).

On the other hand, the Kripke-Putnam view of natural kinds (i.e., a "hard-line" realist metaphysics assuming that similarity, difference, properties and kinds are prior to, hence independent of, cognition) suggests that identifying natural kind essences, and deciding whether members of a certain group of entities, particulars, or different samples of material substance belong to the same natural kind, is a matter of scientific discovery. Moreover, when type physicalists about color like Tye or Dretske speak about natural kinds and ontological categories, they assume a Kripkean metaphysics (Tye, 1995, Ch. 7; 2000, pp. 124-125, p. 167n4; Dretske, 1988, p. 58; 1995, p. 89). However, the key discoveries about object color are pretty much made, and they haven't found that redness and other colors are natural kind essences that are remotely similar to being *gold*, *water*, or even *elephant*. Nassau (1997), and Hardin (1988) review such data and demonstrate that in terms of physical bases the dispositions like reflectance are very heterogeneous.

For this reason, type physicalists abandoned the idea of characterizing object color in terms of microphysical properties, and moved to functional properties like reflectance. Alas, object colors in general appear too heterogeneous even in terms of such properties. Even at this level of description it turns out that not all cases of object color are light-dispositions, even though some are. It seems that no convincing generalized notion of object color in terms of functional properties, dispositions, and the like is available. All that is left for type physicalism to try is to further adjust (bend and twist, to be more cynical), the interpretation of the already established empirical phenomena in order to force them into something vaguely reminiscent of a natural kind schema.

But this armchair-based exercise is very far indeed from the original spirit of the theory of natural kinds. *Perhaps*, with a little more fantasy, such an exercise could be done for any arbitrary set of physical entities, resulting in the conclusion that any such set can be characterized by a non-disjunctive, causally effective physical property, or a functional one, that is true of *all and only* the members of that set. This would entail the conclusion that the members of just any set of physical entities constitute a natural kind of some sort (since they have a property that goes as some vague natural kind essence). But surely, such a result would be devastating for the very notion of a natural kind.

Chapter Four: Normal misperception

2.4. Normal misperception: Another escape for defenders of the natural kind view?

Matthen (1988), and Dretske (1995, pp. 91-92) mount another argument for type physicalism about color. The idea is that perhaps we need not capture all types of actual, contemporary color stimuli by a generalized notion of object color. For there are cases of color perception in normal circumstances that are reasonably regarded as illusory, similar to perceptions of other stimulus properties. Think of movie theatres. The picture that appears on the screen does not really move. What actually happens on the screen is a rapid serial presentation of sequential, stationary phases of movement, but not real (i.e., continuous) movement. What we *see* on the screen, however, is smooth, continuous movement. So arguably there is at least a subtle illusion involved here. This illusion comes about because our visual system was not prepared, in evolution, to distinguish this kind of "pseudo-movement" from real movement – for one thing, there are no occurrences of apparent movement (phi phenomena) in our natural environment. An analogous view arguably applies to another aspect of movie performances: color. The screen, when seen in normal illumination, is white – it only looks colored in the dark, illuminated by the projector. Certainly the screen does not have any inherent property that makes it look red, purple, etc., under illumination by white light. So if for an object to be red is for it to have an inherent property that makes it look red under some normal illuminant (i.e., white light), then movie screens are not red. Yet, often enough, they look red. And if something that's not red looks red, there's an illusion going on. Still, seeing colors in the dark counts as quite a normal circumstance: even our evolutionary ancestors

saw stars, fireflies or bioluminescent mushrooms on dark nights. That is, what we have in the movie is a case (or two cases) of *normal misperception*: stable illusory effects in perception that occur in arguably ordinary circumstances of perception.[62] There are other, quite obvious cases of normal misperception: the Müller-Lyer illusion, other shape or magnitude illusions, and figural aftereffects are all good examples.

Having the notion of normal misperception at hand, the key question for type physicalists becomes this. Are there theory-independent (i.e., non-question-begging) grounds for regarding certain types of color stimuli – especially those ones that do not fit into the type physicalist schema – as giving rise to normal misperceptions, rather than veridical color perceptions? (I.e., the claim would be that such stimuli are not genuinely colored, only apparently colored.) If there are cogent reasons to make this move, then the generalization schema need not take those stimuli into consideration. Perhaps in this way, via a reasonable combination of the generalization and normal misperception strategy, type physicalism can be maintained as a plausible view of object color. I will consider the prospects of such a move in this section.

2.4.1. The idea in more detail

First of all, note that the notion of normal misperception itself is a plausible one. The more general idea behind it is that properly functioning complex systems do make mistakes sometimes (Haugeland, 1981, p. 18). Any system that follows some heuristics, or applies hypotheses routinely, is bound to get things wrong once in a while. Heuristic procedures don't always guarantee correct solution, and that is one source of what we call normal misperception – or, more generally, normal error. Artificial intelligence has

taught us that formally correct deductive algorithms are often too slow to be feasible – think of generating the full decision tree of a chess game. For this reason, "quick and dirty" heuristics are used instead, resulting in some speed-reliability tradeoff (Cherniak, 1984).

There are at least two other sources of normal error. The second source is the use of suboptimal algorithms even when the optimal one would be feasible. Interestingly enough, this seems true of human color vision (Wandell, 1995, pp. 314-315). As Wandell observes, in most experiments done so far on color constancy, subjects do not compensate fully for changes in the illumination. When the illumination changes, color appearance changes less than one would expect from the changes in cone absorptions alone (i.e., without assuming some mechanism for color constancy). Cone absorptions at particular areas of the retina directly correlate with changes in illumination, but color appearance does not directly covary with these changes – exactly because we have *approximately* color-constant visual perception. Still, color appearance changes more with changes in illumination than it would if our nervous system used the best possible computational algorithms (Wandell, 1995, p. 315).

The third source of normal error lies in the poor proximal signal of perception. This obviously applies in the case of color vision: the proximal signal for color vision is the *color signal*: the product of surface reflectance and illuminant relative energy distribution. Our vision has no direct and separate access to the two contributing factors of the color signal.[63] The color signal is therefore an imperfect indicator of inherent surface properties like reflectance – but it is the best available one (Matthen, 1988, pp. 12-13).

To summarize and organize these factors, it is useful to look at the work of Matthen and Levy (1984). Sensory input underdetermines the system's output. Interpretation by the perceptual system is added to the sensory input to achieve the output. This step (i.e., the interpretation of poor, imperfect signals) requires that perceptual processing use some sort of heuristics. Perceptual systems interpret proximal signals in terms of distal target stimuli. Interpreting proximal signals that underinform the system about the distal stimuli (e.g., incident light on the retina) is the normal way of functioning of perception. Making errors in interpretation is part of this normal functioning. The source of such errors is (1) interpretation-heuristics that (2) work on poor proximal signals, and (3) the use of suboptimal algorithms.

2.4.2. Why the notion of normal misperception cannot save the natural kind view of object color

2.4.2.1 Comparison with clear cases of illusion

As we have seen, to escape the problem of generalization of type physicalism to nonreflective color stimuli, one could suggest that light sources, fluorescent surfaces, and possibly a number of other stimuli that look to us colored in normal circumstances of perception are not really colored – they give rise to color illusion, or misrepresentation of true color. It seems that color realists with representationalist allegiances are inclined to make this move, though sometimes with reservations (Matthen, 1988 pp. 24-25; Dretske, 1995, pp. 91-92; Tye, 2000, p154).[64]

As I argued in the previous section, the strategy of generalization fails for emitting surfaces. I think it also fails with fluorescence as the common functional core

that applies to both fluorescent and reflective surfaces is quite vague – too vague to be seriously counted as a natural kind essence of any sort. There might possibly be other color stimuli that support the same conclusion (see Hardin, 1988, and Matthen, 1999, pp. 63-64 for such examples), but in the present discussion I will concentrate on fluorescence and light emission. Let us ask the question: is there any good reason to regard emitting and fluorescent surfaces as not genuinely colored, only apparently colored? A good reason here should not, of course, assume the correctness of the reflectance theory. To say that since color is reflectance, emitting surfaces are not genuinely colored would be outright question-begging. What we need in order to save type physicalism is some theory-independent reason for ruling these stimuli out of the realm of "real" colors. Let us see if there is any such reason available.

Here is the strategy I propose to follow. I will review a few cases that are unquestionably those of perceptual illusion, considering what kind of feature makes them illusions. I start with a list of Tye's own examples (in Tye, 2000, Ch. 7), adding one more case to it. Then I check whether any of these "illusion-creating" features applies to the case of emissive or fluorescent color. In addition to this, I raise another problem against the proposal that emitting and fluorescent surfaces are not truly colored. First, here is the list of straightforward illusion cases.

[1] In the case of shape illusions the required theory-independent reasons are obviously given. For instance, in the Muller-Lyer illusion, the two segments look to us as unequal in length when, as a matter of fact, they are equally long. So it is obvious why and how in this case we are perceptually misled, or misinformed. The same or very similar reasons apply to essentially all other shape illusions including Tye's examples (2000, p. 154).

[2] With alleged cases of normal misperception of color (or color illusions), we have to be more careful. Tye mentions two or three such cases in his book (Tye, 2000).

[2a] The first example is the purplish-looking of faraway mountain slopes (p. 159). In this case, the object itself (the mountain slope) does not have a surface reflectance that is in the purplish range (not at any spatial resolution, i.e., averaging the reflectances of smaller local areas over a larger field – see Tye, 2000, p. 158). The reason why we see the faraway mountains as purplish is that the local distal signal (the event of the mountain slope's reflecting light) is distorted during transmission to the perceiver's eye – by the intervening large mass of air, water vapor or droplets, and so on. So in perceiving the slope as purplish we are misinformed about its actual reflectance property.

[2b] Tye's second example: certain simultaneous contrast effects, for instance, when, in a red and black pattern, the black areas appear to have a greenish cast (pp. 153-155). In such cases it is not the transmission of the distal signal to the eye that results in the altered color perception, but rather the processing of certain special color context (contrast) effects by the brain. When, in a pattern of red and black patches, the black areas look slightly greenish, this effect is not due to alterations of the light that travels from the object to the retina, but rather, to some lateral inhibition effects in the retina (or at some more central level of processing). These effects are caused by the red surround of the black areas. Perhaps it is reasonable to say that there is some sort of normal misperception involved here.

[2c] A third case that might be understood along the same line is experiencing blackness in a totally dark room (Tye, 2000, p. 157 and note 11 on p. 168). In such a case one does not see anything black, in the perceptual or success sense of 'see' (Tye, 2000, p. 55 and

note 11 on p. 67) as one does not see anything at all. Therefore in this case there might be theory-independent reason for saying that a normal misperception of color happens here. The reason is simply that there is no perceived object (surface or volume) that is causally responsible for the Ganzfeld-like black sensation. So we experience blackness due to the darkness (i.e., see blackness in merely phenomenal terms) without perceiving anything. Hence we misperceive the darkness (the absence of visual stimulation) as something (an object or surface) black, or so someone might argue. Pretty much the same applies to the case of afterimages (p. 84): as Tye puts the point, when one sees an afterimage, there is nothing one sees – hence there's always an illusion involved in seeing afterimages.

[3] An additional straightforward case is the Phi-phenomenon, or illusory movement (recall the movie example in 2.4 above). Perhaps even in the movie theatre one is subject to a grand illusion, with respect to movement: all one sees on the screen is a rapid serial presentation of subsequent, stationary phases of movement, but not real (i.e., continuous) movement.

Now the question becomes, is there any theory-independent reason to render emitting and fluorescent surfaces as giving rise to normal misperceptions of color? Let us compare the case of emitting and fluorescent surfaces to those of the just mentioned illusions.

(1) First, when we see traffic lights as green, or active TV screens as yellow, we are not misinformed in the way we are in the case of shape illusions. As I have mentioned, it is immediately obvious why and how shape illusions mislead us. Think of the Muller-Lyer

illusion: in this case equal-in-length is misperceived as longer/shorter-than. When we misperceive a rectangle as a parallelogram, the way the edges are connected to each other (the relation between certain constituents of the shape-pattern) is misperceived. If certain simultaneous color contrast effects are analyzed as normal misperceptions (e.g., when the black area looks greenish on a red background) then it is again understandable why and how we are misled in these cases. What we say in such cases is that our color vision system scales, or evaluates, the black area incorrectly, and it is the red surrounding that is causally responsible for this mistake. Similarly for the case of certain shape illusions: it is typically certain special context effects that are causally responsible for the misperception. Remove the inward and outward pointing arrowheads from the Muller-Lyer figure, and the two segments will immediately look equal in length.

To the contrary, the idea that we *mis*perceive stoplights, hot iron, or long-wave-light-emitting active TV screens as red, is anything but pretheoretically obvious. In what respect are we misled when we see stoplights as red or green, and hence can distinguish one from the other by color? The answer seems straightforward: such perceptions are not misleading in any way. They inform us about properties of stimuli, and help us to discriminate those stimuli from each other. When we see hot iron as red there is no special (or arguably misleading) context effect in play, removing which would remove our (allegedly illusory) red sensation, replacing it with a veridical one (what would that be?). The point is that there is no analogy between shape illusions on the one hand, and seeing emitting surfaces as colored on the other that would support the rendering of emitting surfaces cases of illusory color. The same applies to fluorescent surfaces.

(2) Second, is there an analogy between the cases of color illusions mentioned by Tye on the one hand, and the case of emitting and fluorescent surfaces on the other? When we perceive emitting or fluorescent surfaces, there need not be any *distorting* context (simultaneous contrast) effect involved, contrary to the case [2b] above. Simultaneous contrast effects normally obtain between emitting or fluorescent surfaces just as between reflecting ones, but these effects are not necessarily distorting (most often they aren't). When we look at an active color TV screen, simultaneous contrast effects do obtain between differently colored areas, still, we most often see objects depicted on the screen in quite normal color. Nor is the local distal signal (SPD of the emitted light) necessarily substantially altered during its transmission to the perceiver's eye, as in the case of purplish mountain slopes [2a]. Finally, the case of dark rooms and afterimages [2c] does not generalize to that of emitting or fluorescent surfaces either: when we perceive emitting surfaces, there is an object that causally affects our vision, there's optimal circumstances,[65] incident light impinging on the retina, and so on.

(3) How about illusory movement? Once again, there is no parallel here with emitting or fluorescent surfaces. It isn't at all obvious, either pretheoretically, or to theoretically sophisticated minds, that there is no real color on active computer monitors, hot iron, or fluorescent plastics – i.e., that they aren't really colored, only look that way. The problem here is that the definition of movement in terms of ordinary physics, that is, not relying on our perceptual experience as of movement, is available and it is much less controversial than the same kind of definition for object color. We are in the process of looking for a non-controversial (or the least controversial) definition of object color. As part of this process – i.e., before the job is done – we cannot rule out emitting surfaces

from the realm of genuine colors in the same way as we could if we already had, prior to the ruling, a non-controversial notion of object color and that definition had the consequence that emitting surfaces are not genuinely colored. To the contrary, there is a whole host of reasons to regard emitting surfaces as genuinely colored.

2.4.2.2 Reflective surfaces and color space

Here is another argument against rendering emitting and fluorescent surfaces as cases of illusory color stimuli. First step: we know from colorimetry that reflective surfaces cannot match all color perceptions that can arise from perceiving emitting surfaces. There is a wide range of color experiences that can never arise on looking at reflective surfaces – they can arise only when we look at emitting (or fluorescent) surfaces. Emitting and fluorescent surfaces can look much brighter and also much more saturated than ordinary reflecting ones even though, in terms of chromatic hue alone, reflective surfaces can take any value that emitting and fluorescent ones can. In other words, the so-called object color solid (Wyszecki and Stiles, 1967, p. 335) is part of the color space, but not vice versa, i.e., there are parts of color space that are not parts of the object color solid. Next step: If we bite the bullet and say that only reflecting surfaces and transmitting bodies are truly colored, then the consequence is that there is a wide range of color experiences (i.e., those that can arise from perceiving light sources or fluorescent surfaces but not from perceiving ordinary reflecting/transmitting objects) that are *necessarily illusory*. Color experiences with highest brightness and saturation values cannot arise as a result of veridical color perception. The experiences that arise when we look at surfaces emitting pure monochromatic light are perfect examples. We cannot have

those color experiences no matter what reflecting or transmitting surfaces we happen to

look at. To have an idea of the relation of reflective colors to the whole of color space,

look at Figure 2.

Figure 2. Distribution of the color samples of the Optical Society of America Uniform Color Scales (OSA-UCS) in the x,y chromaticity diagram. The OSA-UCS is a collection of 553 color samples (ordinary reflecting surfaces). It samples color space at points that are perceptually equally spaced. The horseshoe-shaped curve with its ends connected by the straight line is the x, y chromaticity diagram – a cross-section of the Yxy (three dimensional) color space wherein lightness is kept constant thus only hue and saturation vary. The cloud of dots in the middle represents the OSA-UCS samples by their x, y chromaticity coordinates. The horseshoe-shaped perimeter corresponds to the chromaticities (perceptual color coordinates) of monochromatic lights from 400 nm (left side, bottom, where the curve meets the horizontal axis) to 700 nm (corner on the right side). The neighborhood of the right corner is the area of red colors; the bottom left corner area corresponds to blue colors and violets. Between the two, along the straight line are found the purples that correspond to no pure wavelength of visible light. The top-plus-upper-left portion corresponds to greens, whereas the mid-right region is that of yellows. Toward the center, roughly where the cloud of dots is densest, saturation approaches zero and the achromatic colors are found. Notice that the OSA-UCS samples occupy only the center region of the color diagram, not extending into regions where saturation is highest – especially so for the purples and the greenish colors. (Measurements of the OSA-UCS samples are my own.)
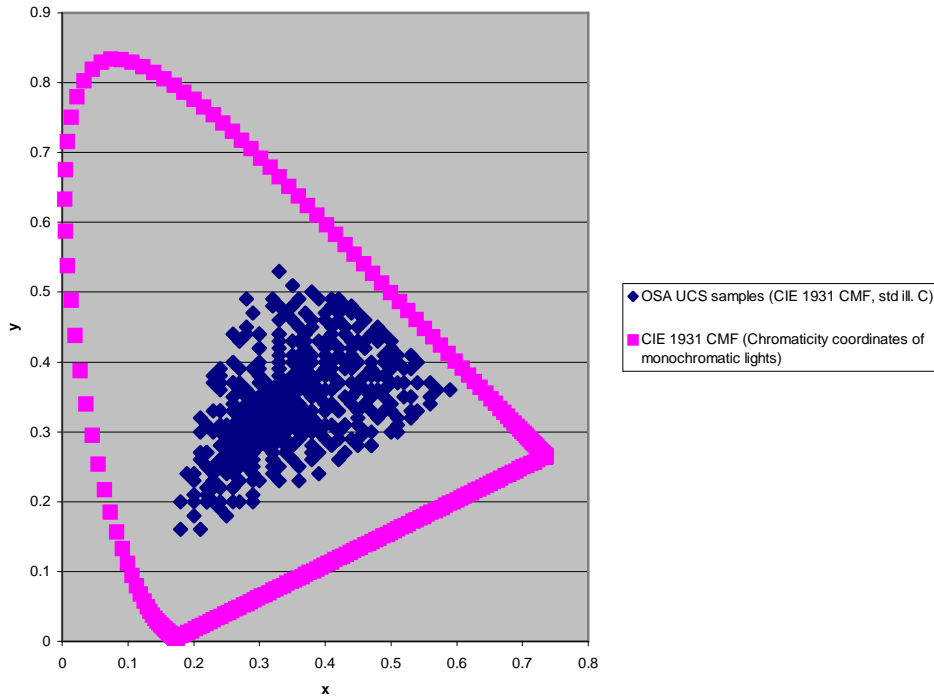
In this figure we can see the distribution of the colored samples of the Optical Society of America Uniform Color Scales (OSA-UCS) in the x,y chromaticity diagram based on the CIE 1931 color matching function (Nickerson, 1977). The OSA-UCS is a collection of 553 color samples (ordinary reflecting surfaces) that sample color space at points that are perceptually equally spaced. In the figure, the horseshoe-shaped curve with its ends connected by the straight line is the x, y chromaticity diagram: it is a cross-section of the Yxy (three dimensional) color space wherein lightness (Y) is kept constant and so only hue and saturation vary. The cloud of dots in the center area represents the OSA-UCS samples by their x, y chromaticity coordinates. The horseshoe-shaped perimeter corresponds to the chromaticities (the x,y perceptual color coordinates) of monochromatic lights from 400 nm (left side, bottom, where the curve meets the horizontal axis) to 700 nm (corner on the right side). The neighborhood of the right corner is the area of red colors. The bottom left corner area corresponds to blue colors and violets. Between the two, along the straight line are found the purples that correspond to no pure wavelength of visible light. The top-plus-upper-left portion corresponds to greens, whereas the mid-right region is that of yellows. Toward the center, roughly where the cloud of dots is densest, saturation approaches zero and the achromatic colors are found. Despite the fact that the OSA-UCS samples are perceptually equidistant, the dots representing them by their x,y color coordinates are not equidistant on the diagram. The reason for this is that equal distances in the x,y chromaticity diagram do not correspond to equal perceived color differences. There are other color spaces that are nonlinear transformations of the Yxy color space and in which Euclidean distances between two points are inversely proportional to the perceived similarities of the hues corresponding to

those points. Notice that the OSA-UCS samples occupy only the center region of the color diagram not extending into regions where saturation is highest. This is especially so for the areas for yellow, green, bluish green, and purple. Highly saturated versions of the colors can only be seen when we look at emitting or fluorescent surfaces. But if the latter are not genuinely colored, only apparently colored, then experiences as of highly saturated (and bright) colors are necessarily illusory. As we can estimate from the diagram, this rendering would affect the larger part of color space – about 70 per cent of its area. That seems too high a price to pay for accepting the normal misperception argument, merely in order to save type physicalism about color.

2.4.2.3 Summary: why the normal misperception escape is blocked

It seems now that none of the ideas that support the rendering of certain classes of stimuli normal *mis*perceptions applies to the case of emitting or fluorescent surfaces. In other words, there does not seem to be any theory-independent reason to view emitting and fluorescent surfaces as cases of normal *mis*perception. So I conclude that seeing stoplights, hot iron and rubies[66] as red are perfectly veridical color perceptions, just like seeing ripe tomatoes as red. As I said above, claiming that stoplights are misperceived as red *because genuine colors are (limited to) reflecting surfaces* would be outright question-begging. If we assume, with the reflectance theorists, that genuine colors are reflectances and nothing else, then we can automatically infer that emitting surfaces aren't really colored. But we lack any cogent reason for such a rendering.

Perhaps, in a Dretskean spirit, one could try to argue thus: most color stimuli are reflecting surfaces in our natural environment – hence distinguishing reflecting surfaces

is the function color vision was selected for (see Dretske, 1995, Ch. 1, Ch. 3). Emitting surfaces are rare in the natural environment, so are fluorescent ones, hence the reactions to them have no (or not much) evolutionary significance. Furthermore, (genuine) colors are the properties the detection of which color vision developed for, in the course of evolution (i.e., the properties whose detection is the function of color vision). Conclusion: genuine colors are reflecting surfaces. I do not find this argument too convincing. It is true that the overwhelming majority of color stimuli in the natural environment consists of reflecting surfaces. However, from this it does not follow that perceiving emitting surfaces as colored cannot carry evolutionary advantage. Perceiving as colored, hence easily identifying, flames, bioluminescent organisms (fireflies, or the glowing green mushroom *panellus stypticus*, etc.), hot metals or live coal, fluorescent and phosphorescent minerals, even the setting sun and the stars can be evolutionarily advantageous for different organisms. Redness, perceived in one perceptual context (red berries among green leaves) can signal food for some organism; in another perceptual context (red-hot pieces of live coal in the soil, accompanied by sensations of heat) it can signal danger to the same organism. Perceiving the redness is perfectly adaptive in both cases, despite the actual fact that in the terrestrial natural environment, most color stimuli happen to be reflecting surfaces.

Chapter Five: No content suitable for phenomenal character

2.5. Defending the first premise: why phenomenal externalism cannot live without the natural kind view of object color

I take it that by now type physicalism about color is effectively refuted. Even if a limited version of type physicalism can be saved in the form of the reflectance theory, it cannot be generalized in any plausible way into a full-blown type physicalist notion of object color that captures the problem cases of non-reflective and non-transmissive (i.e., non-filtering) colors. That is, **P2** of the first, "upper-level" argument in Section 2 above is established:

[**P2**] Type physicalism about object color is wrong.

In this section I turn to **P1** of the same argument:

[**P1**] If type physicalism about object color is wrong, then phenomenal externalism about color experience is wrong too.

As I said in Section 2 above, I endorse a disjunctive physicalist view of object color – roughly one in which stimulus properties that are instances of a particular color (e.g., red) can be characterized by a family resemblance structure (I gave a sketch of the stimulus features determining this structure in 2.3.4). There are red reflecting surfaces, emitting, fluorescent, phosphorescent ones, diffraction gratings, and so on. All that ties these

stimuli together is some crude correspondence in the wavelength range of light they send to the perceivers' eyes – *but that wavelength range itself cannot be identified with the color red*, not even if it can be non-disjunctively characterized.

Let me recapitulate the argument structure that I presented in Section 2 and that will be used to support **P1**.

[P3] Any representational externalist view about color experience has to be able to maintain that object colors play a key role in determining the phenomenal character of color experience.

[P4'] Only type physicalism about color has the resources to support the claim that object colors play an important role in determining the phenomenal character of color experience.

[C2'] Phenomenal externalism about color experience is not compatible with any view of object color other than type physicalism.

C2' is a paraphrase of **P1** (see Section 2 above):

[**P1**] If type physicalism about object color is wrong, then phenomenal externalism about color experience is wrong too.

This premise, together with **P2** that I already have at hand, entails **C1**:

[**C1**] Phenomenal externalism about color experience is wrong.

This conclusion is my target in this dissertation. In the rest of this section I shall defend **P3** and **P4'**.

2.5.1 The third premise (P3)

Once again, **P3** states that any representational externalist view of color experience has to be able to maintain that object colors play a key role in determining the phenomenal character of color experience. I think it is easy to see why this premise is true. **P3** is at the very essence of externalism about representational content in general, and phenomenal character in particular. According to representational externalism, content and phenomenal character do not supervene on internal constitution. In general, this implies that the intrinsic (syntactic-computational, biochemical, etc.) properties do not determine, or cannot in any way fix, the representational content of the internal states of brains/minds. Representational content in general is a relational property, namely some causal relation between brain states and the objects they represent (the causal relation is either lawlike covariation or causal history). Hence the environmental properties that enter, via such brain state–environment relations, into the representational contents of mental states obviously play a key role in shaping those contents – making them what they are. For instance, if the lawlike covariation of perceptual state Q is with skunks and nothing else, then Q has the representational content *that* there is a skunk present. If, in turn, phenomenal character is identified with representational content of some sort, then stimulus properties represented by sensory experiences play a key role in determining the phenomenal characters of those experiences – simply by determining their

representational content. The leading externalist theories of phenomenal character (Dretske, 1995; Tye, 1995, 2000) all endorse such a relation of determination.

2.5.2 The fourth premise (P4')

In this section I start by assessing Tye's theory of phenomenal character, then I show how it generalizes to other versions, including Dretske's. In the two subsections that follow (2.5.2.1 and 2.5.2.2) I assume that Dretske and Tye are strongly committed to physicalist views of object color, and so I will assess their views of the phenomenal in light of the physicalist theory of object color that remains available after type physicalism is ruled out. This view is disjunctive physicalism.

2.5.2.1 Tye's theory and disjunctive physicalism

According to Tye's notion of content, it is lawlike, counterfactual-supporting correlations between stimulus properties and brain states that give rise to content (Tye, 1995, pp. 100-101; 2000, pp. 64-66, 118-122). On this view, if a particular sensory (physiological) state S covaries with the occurrence of stimulus property P in a way that is not merely accidental, but rather, we have reason to assume that, at any occasion where significant anomalies in the circumstances of perception do not obtain, *were P to occur in the organism's sensory field, it would activate S*, then we have reason to say that S (perceptually) represents P. Tye's notion of content is not teleological (2000, Ch. 6); however, he sometimes seems to appeal to causal/evolutionary history (in addition to lawful covariation) as a mediator of content (1995, p. 153; 2000, p. 56). The idea of teleological (sensory) content is roughly that a sensory state S represents stimulus

property P if it has been designed to detect, or indicate, P. What it is that S has been designed to detect is a matter of S's history. Think of the needle positions of a fuel gauge, for instance. The needle position of pointing at 'F' on the plate is designed to indicate full tank. 'Design' can be done either by humans or by Mother Nature: sensory states of the color vision system have been designed (i.e., they have evolved) to detect and discriminate certain surface properties – they were selected for because they did the job of surface discrimination and that enhanced the fitness of the organisms they were the states of.[67]

As we have seen, on Tye's account, perceptual content arises from correlations between stimulus properties and sensory state activations. For object color and color experience the representationalist story goes like this. Redness, the stimulus property, is reliably indicated by certain physiological states in the visual system. By means of this fact about indication, these states count as perceptually representing redness – they acquire the perceptual content that something red is present. This content in turn is the phenomenal character of the experience of red. The experience of red is a physiological state that is contentful, and hence has phenomenal character.

The key problem for this view arises in the following way. As Tye notes explicitly (1995, pp. 194-195; 2000, pp. 124-125), if a sensory state covaries only with some disjunctive property, then that state does not represent anything – it does not have representational content, hence, on representationalism, it cannot have phenomenal character. As Tye adds, disjunctive properties do not themselves enter into causal laws; they do not play causal roles, even though their disjuncts do. Only natural kind essences of some sort (i.e., non-disjunctive, causally effective physical properties) can subserve

causal laws. Moreover, since Tye's notion of content is based on lawful correlations, only properties that can subserve causal laws can give rise to content. So if object colors are truly disjunctive properties, then they cannot give rise to perceptual content that could in turn become the phenomenal character of color experiences. Hence the conclusion: if type physicalism about color is wrong, then Tye's representational externalism about color experience is wrong too.

One could immediately ask, why is it that disjunctive properties are no good for content on the non-teleological account? Why cannot they enter into causal laws? Here is the explanation.[68] It is quite obvious that when it comes to causation, disjunctive properties have a status different from their disjuncts *and conjunctions of properties*. Let **A** and **B** be non-disjunctive, causally effective properties, and **E** some event. '**A** and **B** together cause **E**' is a well-formed causal law: neither **A**, nor **B** alone can cause **E**, but **A** and **B** together can. So, a distinct causal law can take the form

(1) [**A**&**B**] $\rightarrow$ **E**

(where '$\rightarrow$' stands for 'causes').

However,

(2) [**A** or **B**] $\rightarrow$ **E**

is *not* a well-formed causal law. What we should say instead is that there are two causal laws here: (i) **A** $\rightarrow$ **E** and (ii) **B** $\rightarrow$ **E**. So the disjunction here is properly understood as a disjunction between distinct laws, but not as one within the causal antecedent of a single law. In the case of (1), the conjuncts **A** and **B** *cooperate* in evoking **E**; neither one could succeed alone. In the case of (2), however, the disjuncts are *independent* in evoking **E**. At

any particular occasion, if one of the disjuncts is present and is in a position to token **E**, it is totally irrelevant what happens with the other disjuncts – whether they are also instantiated in the local environment, or whether they ceased to exist long ago in the whole world under consideration. Hence (1) cannot be replaced by two distinct laws in the way (2) can. This observation leads us to the conclusion that disjunctive properties do not cause – only their disjuncts do. In any particular case (event of causation) it is one of the disjuncts that is alone responsible for bringing about the effect. Since, in (2) above, **A** and **B** are independent in evoking **E**, that is, they have nothing to do with each other in this respect, why should they be regarded as entering in one and the same causal law? Such a rendering sounds arbitrary, and this is so despite the fact that the effect **E** is the same in the two cases. We know from physics that charged particles in an electric conductor can be made to move in two different ways: either by placing the conductor in an electromagnetic field or by bringing about a voltage difference between two points of it. These two phenomena are closely related, still, the rule that voltage difference produces electric current in conductors, and the rule that electromagnetic fields produce electric current in conductors, count as two different laws of physics, not one and the same law. Similarly, compression into smaller volumes increases the temperature of gases; so do infrared waves. Despite the same effect (temperature increase in gases) we do not say that the two different causes enter *the same law* of physics. Rather, we say that there are two distinct laws here, one being that compression increases the temperature of gases, the other being that exposition to infrared waves increases their temperature.

Here is a slightly different way to formulate the same point. Natural kind essences are properties that endow their bearers with some mechanism that in turn results in

relevantly uniform, homogeneous behavior in various circumstances. H_2O is a dipole molecule, that has a slightly positive and a slightly negative part. As a result of this, water can readily dissolve salts, because its dipole molecules can easily break ionic bonds. Water, however, cannot dissolve glycerids like vegetable oils and animal fat. Alcohol behaves in the opposite way because its molecules are not dipoles. Now, given a set H of entities that can only be characterized by some disjunctive property (i.e., there is no non-disjunctive and causally effective property that characterizes all and only H's members), none of H's members will exhibit relevantly homogeneous behavior due to some physical, chemical, or biological mechanism. The heterogeneity in H's members will not help at all to causally explain why, say, all of H's members can reliably elicit the same effect **E**. Any interesting causal explanation of how an effect **E** arises as a result of heterogeneous causes lies in the functioning of the system that produces **E** as a result of the heterogeneous causes.

Here is an example. Let us call the disjunctive property of *being either my fingerprints, or a particular surface reflectance S_{623}, or the outlines of Australia* property W. W can be specifically causally responsible for carrying out some genetic modification on E. Coli bacteria. All that's needed for this is an effect-instantiator system that consists of (i) a pattern recognition device that is trained or programmed so that when it is input either S_{623} or my fingerprints or the shape of Austalia, it gives the same output signal O, and (ii) an effector device that starts the genetic surgery on receiving O as a signal. It is obvious that all the disjuncts of W have a disposition to genetically modify E. Coli bacteria, since (i) it is counterfactual-supporting that were such an effector system instantiated and given one of W's disjuncts as an input, the genetic modification would

occur; moreover, (ii) we even have an explanatory story of how this counterfactual-supporting link arises. However, the interesting part of this latter explanatory story is about the effect instantiator system, not about property W. With relatively small modifications on the effect instantiator (i.e., re-training the pattern recognition device), W could easily be replaced by a totally different property Q with, say, 273 disjuncts.

Even though set H as a whole (or all of its members) can be characterized by a disposition to produce **E**, such dispositions do not figure in causal explanations (their bases do). But the causally effective bases of this disposition are secondary in causal explanations – their heterogeneity does not play any interesting role in explaining the homogeneity of the effect **E**. For compare the following two explanations. (1) All the disjuncts (all members of H) can reliably elicit **E**. How come? Well, because the system that produces **E** *responds the same way* to all disjuncts. (2) All samples of water can dissolve salts. How come? Well, because all samples of water contain dipole molecules that break ionic bonds easily. It seems, having the non-disjunctive property of *consisting of dipole molecules* is an explanatorily interesting property whereas being a member of H is not. Indeed, disjunctive properties themselves do not play any role in causation and causal explanation over and above what their disjuncts do. Saying that it is *the disjunctive property **A** or **B*** that produces **E** is a vacuous attribution. Either it is property **A** that produces **E** or it is property **B** that produces **E** – there are two separate causal links in play here.

The observation that disjunctive properties do not play interesting roles in causation and causal explanation supports the above idea that disjunctive properties should be regarded as not figuring in causal laws. So comes the bottom line: if there is a

many-to-one mapping between stimulus properties $P_{S1},\ldots, P_{Sn}$ and a sensory/perceptual state E_p, then there is no *single* causal law that connects the different stimulus properties to E_p, thereby subserving E_p's perceptual content. So, E_p has no content (none of the right kind at least: see Tye, 1995, pp. 194-195; 2000, pp. 125-125; see also his 1998), hence, on phenomenal externalism, it can have no phenomenal character either. Since color sensations do have phenomenal character, but they do not have content of the kind required by Tye's theory, phenomenal color character cannot be content as Tye analyzes it.

2.5.2.2 How about disjunctive content?

The above critique of Tye's view leaves open a possible way of defending phenomenal externalism. The idea is that perhaps disjunctive properties can also give rise to some sort of unified perceptual content that can then be identified with phenomenal character. Neither Tye nor Dretske takes this route (Tye, 1995, pp. 194-195; Dretske, 1995, pp. 89-92). Fodor, though not engaged in the problem of phenomenal character, has taken more and more dismissive stands toward disjunctive content (Fodor, 1990a, 1990b, 1994). In Fodor, 1994 he treats disjunctive content as a problem case, under the name 'Twin cases'. 'Water' before 1750 meant *H₂O-or-XYZ*; the systematic, or frequent, occurrence of such cases (together with other problem cases) would mean that the syntactic mapping of semantic properties is not faithful. Still, some people have repeatedly told me that there does not seem to be any immediate obstacle to embracing disjunctive content in explaining phenomenal character.[69] For this reason, I shall elaborate on this idea a bit, trying to make it as clear as possible – though I am not going

to offer a detailed theory along these lines. Then I raise an objection against it that seems to me a serious one. Throughout this subsection, unless I say otherwise, I will assume, like Tye, that content is non-teleological.

Given my case against Tye, someone might argue thus. In the case of a many-to-one mapping between stimulus properties $P_{S1}, \ldots P_{Sn}$, and a perceptual event E_p, there are a number of different causal laws that connect the disjuncts ($P_{S1}, \ldots P_{Sn}$) of the stimulus property to its perceptual effect E_p. Question: do these different laws give rise to the same number of different contents of E_p? If we take this route, it becomes difficult to explain the sameness of E_p's phenomenal character in different circumstances where different disjuncts activate it. Alternatively, we could embrace the idea that perceptual content can arise from a whole bunch of different nomologically stable kinds of causes that have the same effect. That is, the common effect is the activation of a perceptual state, characterized in physiological terms. There is then a "bouquet of causal laws" that underlies the representational content, hence the phenomenal character, of the perceptual state, characterized in physiological terms. This might raise the question, what unifies these separate causal laws into a single, determinate content? And the reply might be: the notion of information. Think of the mathematical notion of information. On it, information is understood as a reduction of uncertainty (or elimination of possibilities). For instance, if we learn that, at time T, a run of experiment E (which has eight possible outcomes: O1-O8) yielded an outcome that is *either O1 or O2 or O3*, we gain information in terms of possibility elimination. That is, we gain *a single piece of information* in this case, just as when we learn that it is exactly O2 that resulted. Now, even if object colors are disjunctive properties, one might point out that in exactly this

sense there is information transfer in color sensation. What a particular color experience does on occurrence is some sort of uncertainty reduction, or possibility elimination. If a tomato looks red to me, I am informed (and can infer) that it has the underlying stimulus property (whatever it is – a determinate property or a disjunct of a disjunctive property), and that it does not have stimulus properties that underlie perceived greenness or blueness.

Given this observation, we could analyze representational content in terms of information in such a way that we do not make a principled divide between partial uncertainty reduction (e.g., that either O1 or O2 or O3 obtains) and more complete uncertainty reduction (e.g., that it is exactly O2 that obtains). All that is different between these two cases is that we gain a smaller piece of information in the former than we do in the latter. Still, in both these cases we gain information, and the delivery of any of these pieces of information can be the function of a sensory state – or it can be the job a sensory state does reliably in non-anomalous circumstances.[70] Having analyzed perceptual content this way, it seems that we can reiterate the representationalist claim that phenomenal character is one and the same thing (by metaphysical necessity) as representational content. So perhaps this way the representationalist evades the problem that arises from the failure of type physicalism about color.

Now let me address the objection. I present the abstract case first, then I give a relevant example. Again, P_{S1}, …P_{Sn} are disjuncts of a disjunctive stimulus property all of which reliably elicit a sensory state E_p. It can easily be the case that some of the stimulus disjuncts are not actually instantiated in the environment in which organisms capable of undergoing E_p live. Say, P_{S1}, … P_{S10} are not instantiated, but P_{S11}, … P_{Sn} are. Or, the first

ten disjuncts occur only extremely rarely in that environment. In this case it is arguable that *the information delivered by E_p tokenings is about the disjuncts that are actually instantiated in the environment, or do repeatedly occur with a considerable frequency.* This is Dretske's opinion (Dretske, 1988, pp. 57-58). His examples are: the job of the doorbell is to indicate people at the door. So if the doorbell rings, it transmits the information that there is someone at the door. The fact that very rarely – say, once in 30 years on average – a squirrel presses the button and makes the bell ring does not make the bell carry the information that either someone is at the door or that a squirrel pressed the button. Rather, in such a case, the inhabitants are misinformed that there is someone at the door – most likely they take the signal to indicate someone at the door. On the other hand, if, for some reason, squirrels make the bell ring frequently, then the doorbell gets to transmit the information that either there is someone at the door or a squirrel passed by. (It is another question whether, in such a case, the inhabitants might want to change the doorbell system so that squirrels cannot make it sound so easily.) Similarly, flies tend to prefer sorbitol to sugar in the experimental lab, and they starve to death (because sorbitol is, though sweet, not nutritious). Despite this, the function of the fly's receptor system remains to indicate sugar (or other nutritional substance); the fly is misinformed by its receptors when it chooses sorbitol and ignores sugar.

To summarize: **if**, out of P_{S1}, …P_{Sn}, only P_{S11}, …P_{Sn}, are present in the environment, P_{S1}, …P_{S10} are not, but were any of P_{S1}, …P_{S10}, to occur, it would elicit E_p just as any of P_{S11}, …P_{Sn} does, **then**, on occurrence, E_p carries the information that either P_{S11} or…P_{Sn} is present. This information constitutes the content of E_p (so we have assumed).

Next step: assume that the normal environment changes, and new, so far counterfactual, disjuncts of stimulus property P_S (say, P_{S7}-P_{S10}) start to occur on a regular basis. The organism of which E_p is a state can adapt to the new circumstances. Given the non-teleological notion of content that we are relying on, E_p's content *is* affected by this environmental change (Tye, 2000, Ch. 6; see also Tye, 1998 for an especially clear presentation of the idea). After the change, E_p will reliably covary with the newly occurring stimulus disjuncts, just as it does with the old ones. That is, after the change, the tokening of E_p will transmit the information that either P_{S7} or … P_{Sn} is present. Change in the information transmitted means change in content. But – and this is the key step – if E_p has phenomenal character and its phenomenal character is its content, then E_p's phenomenal character has to change too – again, *as a result of the occurrence of new stimulus disjuncts that can independently activate E_p*. As I will show in a moment, this is a profoundly implausible consequence.

Before going on, note that if we shift to the teleological, function-based notion of content, then it is *perhaps* arguable that such environmental changes do not affect content. The sensory system of which E_p is a state was selected for the detection of certain domain of properties, among them P_{S11}-P_{Sn}, but P_{S1}-P_{S10} played no part in its evolution. The fact that, at some later point in time, P_{S7}-P_{S10} start to occur regularly in the environment and elicit E_p need not make it the case that E_p *acquires the function to detect* any of P_{S7}-P_{S10}. For function in this context is analyzed in terms of causal (evolutionary) history (Wright, 1973; Matthen, 1988; Dretske, 1995, Ch. 1), and, by assumption, E_p developed in evolution for the detection of P_{S11}-P_{Sn}, but not for the picking up of P_{S7}-P_{S10}. That is, using the teleological notion of content, the plausible reply to such cases of

environmental change is that P_{S7}-P_{S10} give rise to normal misperceptions when they token E_p (see Dretske, 1995, pp. 88-93).

Now let us see why the consequence of change in phenomenal character is so implausible. It is important to understand what this consequence implies. For instance, let us call the *perceived* shade that a typical ripe lemon gives rise to in trichromat humans yellow$_{39}$. Yellow$_{39}$ is a perceptual state identified (well, in principle identifiable) by its physiological characteristics, and it has a phenomenal character. To it there corresponds some physical stimulus property – either disjunctive or not. In the natural environment in which humans evolved, yellow$_{39}$ was evoked by the surface reflectance of ripe lemons in daylight, and perhaps a few other reflecting surfaces. In our contemporary environment, however, a number of other stimuli can probably elicit yellow$_{39}$. Emitting surfaces like color TV screens, computer monitors, fluorescent ones, movie screens illuminated by a projector are a few examples. Even if, in our original, natural environment, yellow$_{39}$ happened to stand for a non-disjunctive property, in our contemporary environment it is reliably elicited by a disjunctive property. On the non-teleological, disjunctive content view we are assuming, this change in the stimulus conditions changes the content of yellow$_{39}$. And if its phenomenal character is its content, then yellow$_{39}$ changes its phenomenal character as well. This means that ripe lemons, the same kind of thing that looked one way to our ancestors, looks to us in some other way, colorwise. Both in our ancestors and in us, ripe lemons elicit perceptual state yellow$_{39}$, physiologically identified, but in them and in us, this state acquires different (externalist) phenomenal characters. The reason for this change is the occurrence of the different sorts of non-reflective yellows that perceptually match the color of ripe lemons.[71]

But this consequence is unacceptable, since it raises a number of very strange questions that we have no idea how to answer. What sort of phenomenal change could this be? How did ripe lemons look to our great-grandparents, as compared to how they look to us? There seems to be no way, even in principle, to answer this question. So should we *really* believe that by the large-scale introduction of new color stimuli, as has happened in the last 150 years, the phenomenal look of the good old color stimuli has changed – simply because the content of the perceptual states that these old stimuli token changes with the introduction of the new color stimuli?

Intuitively at least, it seems quite plausible that ripe lemons do *not* look any different in color to us than they did to our great-grandparents – simply because our visual systems work the same way as theirs did (there is no reason to assume otherwise), and ripe lemons have the same color pigment in their peel as they did 100-150 years ago. Moreover, as I think both common sense and scientific intuition would agree, in determining how a ripe lemon looks to us colorwise, only our nervous systems, the lemon, and the perceptual circumstances obtaining at the spot play a role. Color stimuli that are not present at the same place and time (but if they were, they would look the same color as the lemon) do not play any role in determining the look of the particular lemon that one is looking at, at a given time – exactly because such remote, uninstantiated stimuli do not causally affect the perceptual process. This is, I think, a reasonable intuition. Notice that phenomenal externalism built on some notion of disjunctive content does *not* honor this intuition. Stimulus properties that are not physically present (hence not causally affecting our perception) still become part of the information delivered to the perceiver by the corresponding sensory state tokening. If

each stimulus $P_{S1},\ldots, P_{Sn}$ can reliably token sensory state E_p (and does so regularly) then, on a particular occasion when P_{Sk} does the tokening, the information that is delivered to the organism by the activation of E_p is that *either P_{S1} or ... or P_{Sn} is present*. If it is this information that ultimately determines E_p's phenomenal character, then this phenomenal character is partly determined by something (i.e., stimulus properties $P_{S1},\ldots P_{Sk-1}$ and $P_{Sk+1},\ldots, P_{Sn}$) that are not actually present and hence do not causally influence the particular process of perception. Phenomenal externalism that rejects the notion of disjunctive content (like Tye's and Dretske's theories) can honor the above intuition in the case of veridical perception, but not in the case of non-veridical perception (including hallucination). If the object color lemon yellow is natural kind essence K_{Y39}, then, in each and every case where we veridically perceive something lemon yellow, there is an instantiation of K_{Y39} present (i.e., some object or surface having the property K_{Y39}) which causally influences our color perception. Not so for misperception or hallucination.

As we have seen above, by shifting to a teleological notion of content we can avoid the disjunctive content assumption, together with its strange consequences. Teleological content does not change with the introduction of new color stimuli. If, in the course of evolution, the object color lemon yellow happened to be a non-disjunctive stimulus property (even some natural kind essence), but by now it has become disjunctive, then we still can construe the content of lemon-yellow sensations as built on the historical non-disjunctive property. However, in making this move, we are obliged to rule that all new kinds of color stimuli that perceptually match the color of ripe lemons, but are themselves non-reflective, give rise to normal misperceptions of color – color illusions, that is. As I argued in Section 2.4.2 above, we have no good reason to rule this

way. Finally, there certainly is one way to avoid these problems, and this is by assuming that phenomenal color character is essentially determined by the constitution and internal mechanisms of our visual system. That is, phenomenal character is not representational content of any sort. Rather, it is part of the representational vehicle for object colors. This is phenomenal internalism.

2.5.2.3 Dretske's theory and disjunctive physicalism

As we saw in 1.3.2 above, on Dretske's version of representationalism, it is the stimulus properties themselves that are the phenomenal character of experiences. On this approach, phenomenal characters are identified not with particular physical instantiations of stimulus properties, but rather, with universals. Universals are properties (and kinds) in the metaphysical realist view: properties that exist in the world independently of how they are conceived. On Dretske's view, for each sensory state type there exists exactly one stimulus property (i.e., universal) that that state indicates, hence represents. Such stimulus property universals partially constitute the representational content of sensory states.[72]

I wish to mention just one problem for this view that stems from the failure of type physicalism about color. Object colors are not universals, so there is nothing that could serve to constitute, on Dretske's view, the phenomenal characters of color experiences. There might be room for some argument here, so here is a more careful way to put this point. In terms of stimulus properties, object colors are either not universals at all, or they are universals only in the weakest possible sense of the term. Certainly, if we admit the existence of universals, then there is a sense in which predicates with very

loose structure correspond to universals as well (e.g., 'game', as Wittgenstein sees it; or 'women, fire, and dangerous things', to use G. Lakoff's famous example [Lakoff, 1987]). So object colors are universals in a sense in which almost any arbitrary bunch of entities or properties constitutes a *single* universal. However, on any conception of universals that derives from scientific realism,[73] object colors are not universals at all, because they are not fundamental physical types, nor are they high-level physical types. At the end of the day, the plot remains this: In the case of color perceptions, there is no physical substrate to perceived similarity. All and only red things look to us to have a salient, homogeneous, inherent surface property, but as a matter of physical fact, they have none (at least no causally effective, non-disjunctive surface property). This makes it very hard to maintain any externalist view of color phenomenology.

2.5.2.4 Other theories of object color and phenomenal externalism

In this section I examine three other theories of object color: subjectivism, dispositionalism, and John Campbell's Simple View, in order to see whether any of them would be a suitable account upon which phenomenal externalism about color experience could be built. I make the following general assumption in this subsection. All current externalist theories of representation have it that the relation of representing or aboutness (i.e., that an inner state represents an external state of affairs, or it is about that state of affairs) arises from some sort of causal relation. Representational content is acquired through causal relations that hold between the organism and its environment. Next step: according to representationalism, color experience represents object color. Therefore, for any representationalist account of phenomenal color experience to get off the ground,

there has to be a systematic causal relation between object colors and states of the color vision system that are the color experiences. In other words, assuming that theories of representation are correct in their most general assumption that representation is a causal relation, quite plausibly, color experience can represent object color only if object colors can systematically cause color experiences. Conclusion: any account of object color according to which colors are causally inert, or they are not the causes of color experience, cannot be used to establish phenomenal externalism.

Note that in general it is not true that we can represent only those properties that can cause relevant changes in our cognitive systems. Our concept UNICORN, for instance, has the representational content *unicorn*, even though no unicorn has ever tokened anyone's UNICORN concept. One plausible solution to such a case is that whereas, say, COW *can* be a primitive, or atomic concept, UNICORN cannot (e.g., Fodor, 1990b). The concept UNICORN is an abbreviated description – it is constructed out of basic concepts (e.g., HORSE and HORN) that acquire their representational content from direct causal links to objects in their extension. However, color experiences are unlike conceptual representations in this respect: it is not possible to construct mentally, out of color experiences, other, more complex color experiences that do not correspond to any actual stimulus property in the environment. That's just not the job of color vision.[74] The system of perceptual color representations does not exhibit productivity and systematicity because this system has no powerful generative syntax. Rules of combination within the realm of color experience are very simple compared to linguistic representations, and the complexity of the end product is also strictly limited. We can have color experiences as of unique hues and of binary ones, plus saturated and

unsaturated ones. It is arguable that the experience of, say, purple is syntactically more complex than that of unique red – the former has two decipherable constituents in it (i.e., felt reddishness and felt bluishness) whereas the latter does not (see Jakab, 2000a, 2000b for more detail on this topic). Perhaps Hume was right, and there are shades that we can imagine even though we have never experienced them before. It is perhaps arguable that, say, Mary the colorblind scientist (Jackson, 1982, 1986) could imagine orange if she was shown red and yellow but not other colors. But, on a representationalist account of color experience, this capacity of Mary's to imagine orange is due to the fact that she has successfully perceptually represented red and yellow, that is, red and yellow object colors did cause her experiences as of red and yellow. This story could not be told if it was supposed that colors in general cannot cause color experiences.

Now let us evaluate the above three views of object color with regard to representational externalism. The story is the shortest in the case of subjectivism. On subjectivism, objects are not colored. On different versions of subjectivism, colors are either in the head (i.e., they are mental events, phenomenal color experiences) or they are systematic illusions – properties that we represent objects as having, ubiquitously and erroneously. Even though there are stimuli that are systematically causally responsible for eliciting our color experiences, these properties are so dissimilar to what is suggested by color perception that they do not deserve to be called colors. Clearly, on this view, colors are not causally responsible for eliciting color experiences.

The situation is a little more complicated in the case of dispositionalism about color. One could argue in two related ways. On dispositionalism colors are either not the causes of color experience or if they are, they are so only in a secondary, derivative

sense. Dispositionalism holds that colors are the dispositions of objects to elicit color experience in us. As I mentioned above (Section 2.3.2), dispositions are (or so I take it) functional states, or multiply realizable causal roles. Such causal roles are not themselves causally effective – only their bases, the physical properties that endow their bearers with the dispositions are. What is distinctive about dispositionalism is that it identifies colors with *the dispositions* to elicit color experience, and *not* with the bases of such dispositions. It is physicalism that identifies colors with the bases of dispositions to elicit color experience. These bases are typically identified as high-level physical types like reflectance, emission profile, and so on. So, on dispositionalism, colors are abstract causal roles not themselves causally effective. To say that colors – the dispositions to elicit color experience – are causally effective is either false or it is true only in the secondary, derivative sense that it is the bases of these dispositions (i.e., the accidental color-role fillers) that are literally causally effective.[75] This secondary sense of causation already constitutes sufficient reason to cite colors, as dispositionalism analyzes them, in causal explanations. It makes sense to say that the glass fell and broke because it was brittle; similarly, it makes sense to say that ripe strawberries look red (and are red) because they have a disposition to elicit red experience in us. But from the fact that dispositions figure in causal explanations we cannot conclude that dispositions themselves are literally causally effective.[76] Dispositions can figure in causal explanations simply because they correspond to (or simply, *are*) lawlike causal links, i.e., ones between the bases of the dispositions and their manifestation. But to say that it is the lawlike causal link that causes something is a confusion: the causal link is constituted by the basis property, the event of manifestation, and a specific, predictable interaction (i.e.,

the event of causation) between the two in specific circumstances. All this leaves us with the conclusion that, *strictly speaking*, dispositions do not cause, hence on dispositionalism colors themselves do not cause our color experiences.[77]

Finally, let us turn to the Simple View of color, introduced in Section 1.1.3 above. It should be obvious that were this view correct, it would be the perfect partner for phenomenal externalism. On it, colors are the causes of color experience; they are high-level physical types; they directly determine the phenomenal character of color experience via a strong link of transparency – representationalists could not dream of a better account of object color. *If* it were correct, I stress. Note that Campbell's view seems quite close to type physicalism, as far as we can assess from his concise exposition. For most likely, redness is a *non-disjunctive physical type* if (i) it looks non-disjunctive (i.e., all and only red things *look* to share exactly one salient inherent surface property, that is, the same one with slight variations in each and every case of red objects), and (ii) color experience is transparent with respect to color properties. The key difference between the Simple View and type physicalism is that the former assumes that object colors are not accessible to conceptual grasp – as Tye puts it (2000, p. 149), they are epistemically basic, but metaphysically derivative. Type physicalism denies this and attempts to specify explicitly the high-level physical types that are the colors.

Here is the key problem with the Simple view as I see it. We know a lot about which high-level physical types are specifically causally responsible for eliciting our color experiences. As we have seen, these are surface reflectances, transmittances, emission profiles, and so on. Each object color is a disjunctive property, but probably one with not very many disjuncts. (This is so because, as I argued above, plausibly there is a

solution to the problem of metamerism, at least in the case of broad color categories, and it is metamerism that has led some theorists to conclude that there are infinitely many surface properties – reflectances – that look to us the same color in normal circumstances). Furthermore, what we know, conceptually, about the high-level physical types that are the colors, we know from empirical science, and not from color experience. Color experience does reveal to us something that seems not to be available by any other means (i.e., what it is like to see the colors). But certainly, it does not reveal to us the nature of the high-level physical types that are the systematic causes of color experience (i.e., the types *as identified by science*). At this point Campbell might object that I am begging the question against him, since he made a distinction between transparency to color experience and transparency to conceptual grasp. I reply that this very distinction is dubious, as shown by the case of shape perception. Shape perception does reveal to us the nature of shapes, in a way that is accessible to conceptual reconstruction. In this sense, shape perception is transparent with regard to shapes. But what kind of transparency is it that allows us no conceptual access? At a minimum there seems to be no relevant analogy for this in other realms (see Smith, 1993, p. 272).[78]

In other words, one just cannot be content to reply to the question "But please, what kind of high-level physical types are the colors?" by saying "Well, just look. Don't reflect.", as Campbell wants us to do. Empirical science has taught us the high-level physical types that are the candidates for being identified with the colors. These types are not transparent to perception in the non-controversial sense of transparency (e.g., in the way shapes are perceptually transparent). In the face of this challenge, Campbell wants us to think that colors are some other high-level physical types, ones that are (i) inaccessible

to conceptual grasp, (ii) still transparent, but are so merely to color vision. So either we say that nothing in the natural order responds to Campbell's characterization of object color, or we buy into his dubious notion of transparency and agnosticism about certain high-level physical types. I prefer *not* to do the latter, and so I conclude that the Simple View is mistaken. Therefore it is no way out for phenomenal externalism.

Chapter Six: Individual differences in phenomenology

3. The problem and the second argument

There exists another argument against phenomenal externalism that is independent from the one I presented so far, the argument from the heterogeneous nature of object color. In its general form, the argument from individual differences goes as follows. Even assuming, for the sake of argument, that object colors – at least broad color categories – are sufficiently well characterized perception-independently, in terms of stimulus properties, it is an empirical fact that there are differences between color-normal individuals with respect to their color vision.[79] Even more interestingly, differences of the same kind can be observed between not just trichromat individuals, but groups like women and men, different race and age groups. Color normals are those people who pass the standard tests for trichromacy – for instance the Pseudo-Isochromatic Plates, or the Farnsworth-Munsell 100-hue Test. Such differences between color-normals show up at three different levels (for a review of the relevant data see Block, 1999, pp. 41-44). First, the color-matching functions of different color-normal individuals often differ substantially (Lutze et al., 1990; Neitz and Neitz, 1998; Hardin, 1988, 76-82, and Fig. II-5A-B on p. 77; Block, 1999, pp. 41-42). Second, different color-normals respond differently in color-matching experiments: for instance, females and males typically disagree on the location of their matches. In tests with the Rayleigh anomaloscope, subjects are asked to make the two halves of a screen match in color where one half is illuminated by a mixture of red and green light, and the other half by yellow or orange light. The task is to adjust the intensities of the red and green lights so that the mixture is

indistinguishable in color from the yellow or orange light (Neitz and Jacobs, 1986; Block, 1999, p. 42). In this test, female subjects most often find a match at a point where no male matches (Neitz and Jacobs, 1986, p. 625). Third, there are individual and group differences in the location of unique hues (Block, 1999, p. 43; Hardin, 1988 pp. 79-80; Byrne and Hilbert, 1997, p.272; Kuehni, 2001). These differences are especially large when the stimuli are monochromatic lights (Ayama et al., 1987; Kuehni, 2001, p. 61). The location of unique blue varies between 459-489 nm; locations of unique green vary between 488-535 nm; unique yellow ranges between 544-594 nm; finally, for unique red sometimes a spectral wavelength is chosen, sometimes some extraspectral light (i.e., a mixture of different monochromatic lights). For reflective surfaces, the variation of unique hue locations is smaller, but still significant (Kuehni, 2001, pp. 63, 65).

These differences are related to one another. Between-subject differences in the M- and L-cone sensitivity curves cause a difference in color-matching responses (Neitz et al., 1993, p. 117; Block, 1999, p. 42). Subjects' unique hue choices are also linked to the variations in the crossover wavelengths of their color-matching functions (Kuehni, 2001, p. 65).

These results together indicate that the same color stimulus (characterized in perception-independent terms like surface reflectance) in the same circumstances of perception often appears, to different color-normal perceivers, slightly different in color (Kuehni, 2001, pp. 63, 65; Block, 1999 p. 43). Correspondingly, for two different color-normal subjects in the same circumstances it is often slightly different color stimuli that look, phenomenally, the same in color (e.g., unique green; balanced orange, etc.). (Note that this phenomenon is not the same as metamerism, since, as I just said, for two such

subjects the same reflectance would not look the same in color.) However, if the relevant color-property (i.e., the stimulus property that color experience represents) is one and the same, and circumstances are also identical, then the two subjects' representations of the same color stimulus has to be assigned the same representational content.[80] If, in such a case, the same color property gives rise to different color appearances (perceptions as of different shades) in different subjects, then the problem becomes that color phenomenology differs without a corresponding difference in color content. This is an unacceptable consequence for representationalists like Tye or Dretske who hold that phenomenal character is exhausted by representational content. If color content and color phenomenology vary independently of each other, then the two cannot be identical, nor can phenomenal character supervene on perceptual content (see Block, 1999, p. 40 and note 2 on p. 67; Byrne and Hilbert, 1997, pp. 267-272). The claim then is that data on individual differences in phenomenology falsifies phenomenal externalism.

There are a number of responses given by representationalists to this argument. One such response is by Byrne and Hilbert (1997, pp. 272-274). Block (1999) evaluates a number of possible representationalist responses, and refines the argument. Tye (2000, pp. 89-93) answers Block by trying to fend off the contention that representationalism has no resources to accommodate the phenomenon of individual differences. In the rest of this section I first offer an overview of the debate. Then I continue by challenging Tye's and Byrne and Hilbert's defenses, contending that their responses are correct, but incomplete, and so leave the door open for a slightly improved form of the argument. I shall work out the details of this improvement. The new version of the argument requires, as one of its premises, some empirical phenomena to actually obtain. I shall end this

dissertation by reporting and discussing the results of an experiment that was designed to test whether the required phenomena indeed obtain. If they do, then, I shall claim, the argument from individual differences gains more force and becomes an independent reason to question the correctness of phenomenal externalism.

3.1. Byrne and Hilbert's response

Byrne and Hilbert, after making the claim that color content and color phenomenology are necessarily correlated (Byrne and Hilbert, 1997, p. 267), offer the following response to the problem of individual differences (pp. 272-274). On the evidence, it seems possible that a particular color stimulus Z looks pure green to one subject, and slightly bluish green to another subject. We need *not*, however, hold that green and bluish green are, in general, contraries – that is, we need not believe that if an object at a time is unique green overall, then is it *ipso facto* not bluish green overall. For there may be certain greens that are at the same time bluish greens and vice versa. There is at least an analogous possibility in the realm of shapes. At first glance, being a rectangle and being a diamond seem mutually exclusive attributes. But they aren't in general mutually exclusive. A square, standing on one of its corners, is both a rectangle and a diamond, even though it may not look like a rectangle at first glance. Similarly, a square, standing on one of its edges, does not look like a diamond, but if we change the viewing conditions slightly, (i.e., turn it around by 45 degrees) it will look like a diamond. On the other hand, there obviously are rectangles that are not diamonds, and vice versa. Perhaps the same is true of colors: *certain* colors are both unique green and bluish green. Even though such a color may not look bluish at all under one normal

illuminant, by changing the illumination slightly, it may start to look green with a tinge of blue in it. By admitting that, in general, unique green and bluish green aren't contraries, we have a reply to the problem of individual differences. Moreover, even if there emerges some reason to believe that unique green and bluish green *are* contraries, there still is no disaster, just a bit of nuisance for the type physicalist. For in this case a subject who perceives a unique green surface as slightly bluish, but still predominantly green is subject to a small degree of illusion or misperception, while her color perception is, for the most part, correct. That might happen as living creatures often have sub-optimal mechanisms to implement biological functions – sub-optimal mechanisms that are still good enough to assure fitness and survival.

3.2. Block's anti-representationalist arguments

After reviewing relevant empirical data and formulating the argument, Block (1999) looks at a number of possible representationalist objections and replies to them. In this section I go through the relevant parts of Block's paper, sometimes inserting paragraphs with my own remarks. First we have to make a distinction between *perceptual color categories* that are not themselves color concepts, and *experiences as of narrow shades*. Narrow shades are shades that have no discriminable sub-shades (Block, 1999, p. 46). Perceptual color categories correspond to the perceptual similarity groupings of the colors: for instance, when we look at a rainbow, we see distinct bands of different colors where the stimulus itself has no corresponding bands of different type, just the gradually changing wavelengths of light (Hardin, 1997, p. 293; Shepard, 1997, p. 344; Matthen, 1999, pp. 68, 76). Block observes that if the argument from individual differences is put

in terms of perceptual color categories, then it is relatively easy, for the representationalist, to reply that difference in phenomenal character is accompanied by a difference in representational content. For instance, if Max and Samantha both experience the same mailbox as red, yet their experiences differ slightly in phenomenal character, then the representationalist can reply that this scenario entails a difference in color content. For, say, the shade of the mailbox may be included in Max's perceptual category of pure red, whereas for Samantha, the same shade is included in her perceptual category orangeish red. That is, the two represent the shade of the mailbox in different ways (both being acceptable, in terms of veridicality – perhaps Max's perception fails to detect the tinge of yellow that Samantha's color vision catches).

The next question this scenario raises is how there can be objective (i.e., intersubjectively available) colors if different subjects differ with respect to their perceptual color categories. Block's reply is, there are objects which would be categorized in ideal circumstances as, say, yellow by most or all color-normals, and these objects are objectively yellow. This definition leaves some objects indeterminate in objective color – think of a fabric that looks, to a number of people, green (or bluish green), whereas to many other people it looks blue (or greenish blue) (Block, 1999, p. 45).

At this point I wish to remark that such a perceptual-disposition-based definition of object color is not favored by most representationalists who are targeted in Block's paper. I am sure Tye and Dretske would reject such definitions, insisting on some perceiver-independent terminology instead, such as surface reflectance. For a number of reasons, some of which I tried to provide in the previous sections, Block's approach to

objective color seems inadequate for representationalist views of phenomenal character. For instance, if *being green* is the same as *being the stimulus property that most or all trichromat humans would perceive as green in normal circumstances*,[81] then is this relational property (i.e., being perceived by most other trichromats as green) the one that enters the perceptual content of our experiences as of green? How could such a property causally affect our color perception, and thus acquire a strong position in a causal theory of perceptual content? There seems to be no easy answer to these questions.

Returning to Block's line of reasoning, if the argument from individual differences is formulated in terms of narrow shades and experiences of them, it will not be susceptible to the counterargument from color categorization. When Max and Samantha both look at the same OSA-UCS sample in the same circumstances, they may still experience a slightly different color. That is, they undergo slightly different *perceptually determinate color experiences* (the vivid color experiences that occur on looking at colored objects in good lighting). Since stimulus and circumstances are the same, the perceptual contents of Max's and Samantha's color experiences must be the same as well.[82] Yet there is phenomenal difference, so representationalism is in trouble. Since empirical data suggest that there are such phenomenal differences not underlain by representational ones between females and males, different races, and different age groups, representationalism can only escape by claiming that some groups do not perceive colors veridically – for instance, either men, or black people, or the elderly, undergo systematic color illusions. But this is patently unacceptable – representationalism has led us to an absurd conclusion. If we assume, as we should, that

all of the above groups perceive colors veridically, then what we have is a reductio of representationalism about color experience (Block, 1999, pp. 46-47).

We can put this result in the following way (Block, 1999, p. 48): evidence supports the distinction between two different senses of 'looks': the intentional sense in which two objects look the same in color just in case they are represented as having the same color property, and the qualitative or phenomenal sense that recognizes phenomenal differences that go beyond perceptual content (Block, 1999, p. 48).

The next representationalist objection Block considers goes as follows. A given narrow shade looks different to different groups, therefore different groups represent it as having different colors. Moreover, the representationalist might go on conceding that it does in fact have slightly different colors relative to these groups (Block, 1999, pp. 48-49). If color was defined earlier by means of a relation to trichromat perceivers, why can't we redefine it now relative to different groups of trichromats? The story continues thus: different groups represent one and the same color sample as having different, group-relative color, so their phenomenal difference is underlain by a representational difference, so representationalism escapes. Block's reply to this objection is that there are, as a matter of fact, objective colors – it is enough to go to a paint store to see a lot of them. Hence colors – at any rate, a lot of the colors – are not relative to different groups of trichromats, and so this representationalist objection does not get going.

My remark here is that if narrow shades in the paint shop like Dove White, Benjamin Moore Linen White, Burnt Sienna and the like are all objective, what makes them objective? Certainly not the fact that most color-normals perceive them the same way, because that is false. So Block's definition of objective color that he gave for broad

color categories does not work for narrow shades. What can indeed make colors in the paint store objective is their surface reflectance, and perhaps the mixing procedure by which particular shades (custom colors) are produced from the premixed colors. These are all perceiver-independent characteristics, so it's time to realize that if we want a strong version of the argument from individual differences that proceeds from narrow shades, then we need a definition of object color in perception-independent terms. (Remark ends.)

In a similar vein, the representationalist could object the following way. Assume that Max and Samantha, on looking at the same OSA-UCS sample, undergo phenomenally slightly different color experiences. Following Sternheim and Boynton's descriptive terminology (Sternheim and Boynton, 1966; see also Miller, 1997), assume that the sample looks 60 per cent reddish, 40 per cent yellowish to Samantha, and 70 per cent reddish, 30 per cent yellowish to Max. Since stimulus and circumstances are the same, the phenomenalist would conclude that there is no corresponding representational difference. But the representationalist could just as well insist that there is a representational difference here too – since if the sample looks to Max more reddish, then he just represents it that way – that is, differently from Samantha (Block, 1999, p. 49). Block's reply is that there can be many cases in which there is verbal agreement between Max and Samantha regarding the color of the sample (e.g., both would judge it reddish orange), yet there is a slight difference in their phenomenal characters – a difference that is too subtle to be expressed verbally. As Block suggests, there is no reason to assume that the representational contents differ in such cases (Block, 1999, p. 51).

Inserted remark: Here Block's reply seems to me totally inadequate. Of course there can be differences in perceptual content even if it is virtually impossible to express such differences verbally. To capture such differences was one main reason for which the notion of perceptual (non-conceptual) content was introduced in the first place. Tye would naturally agree that differences in perceptual content can be too minor to be accessible to the conceptual level. Conceptual representation is digital, perceptual representation is analog, so no matter how good the interface is between the two, small pieces of information probably get lost in the process of code transformation. This is one factor that reasonably contributes to the so-called ineffability of sensory experiences (Raffman, 1995).

A better reply to this representationalist objection, formulated in a terminology that Tye and Dretske would prefer, is this. Since perceptually determinate color experiences have the function of representing narrow ranges of surface reflectance, if the same reflectance R activates two different, perceptually determinate color experiences in Max and Samantha, then necessarily, Max represents R as belonging to one narrow range of reflectances whereas Samantha represents it as belonging to another narrow range of reflectances. If, in addition, it is assumed that the two narrow ranges (well-specified types) of reflectance that are the contents of Max's and Samantha's color experiences respectively are non-overlapping (an assumption that does not come for free!), then the conclusion is that one of them misrepresents the sample colorwise (see below for a more detailed explanation). But since both Samantha and Max are normal color perceivers, this conclusion is unacceptable, so representationalism gets in trouble again. Back to Block's line of thought.

As Block admits, in light of the individual differences, there might be a problem with the notion of objective color introduced above. *Prima facie*, the following two principles about color sound plausible. (1) If something looks red in optimal circumstances, then it is red. (2) If something looks red to *me* (or any particular normal color perceiver), then it looks red simpliciter. That is, based on the first person perspective of experience, one can tell authoritatively whether something looks red. However, given the data on individual differences, what looks red to me, might not look red to other people. What color an object looks to me is no infallible guide to what color it looks to others. So trouble emerges as to the compatibility of (1) and (2). The definition of objective color Block offers retains (1) and drops (2): If something looks red to most or all color-normals, then it looks red simpliciter, and if something looks red simpliciter, it is red. Thus it seems that the natural intuition according to which colors are both observational and objective cannot be maintained. If we assume that colors are observational (i.e., I alone can establish the color of objects by looking) then they are not objective: by their own perception others would establish otherwise. If, on the other hand, colors are taken to be objective, then they are not observational (since all or most color perceivers' impressions must be taken into account in establishing the color of objects).

Subjective or person-relative colors are understood by intersubjective differences in color categorization: different subjects' perceptual color categories comprise somewhat different minimal shades. So if Max and Samantha look at the same OSA-UCS sample, they both see the same narrow shade, but they perceptually categorize it in slightly different ways. This is a difference in representation, and it explains their differing phenomenal experiences. Block (1999, p. 52) does not seem to have a strong

reply to this scenario. Instead, he again tries to avoid invoking color categories, and rely only on narrow shades. His key example is this: when Samantha looks at Purple 67 and Purple 68 (say, two samples that are one or two just noticeable differences away), the two chips will look different to her, but it is difficult to assign any representational difference here. She would categorize both samples blue-red, or purple; no conceptual representation could capture their minute difference. Perhaps a representational system of ultimately fine-grained chips could capture the difference in some way, but that is not available to her on the fly, argues Block.

Remark: this is again an inadequate reply. For no good reason, Block insists on capturing the representational difference by some digital representational system – distinct perceptual categories or concepts. But the representational difference can be entirely in the perceptual content that is reasonably thought of as analog, and so it can represent very fine differences. Perhaps as fine as ones corresponding to the minimal phenomenal difference between two color experiences, the representationalist may reasonably argue. Moreover, the digital representation of such minimal differences might be a plausible idea, too. Samantha may be a color expert in a paint factory, having acquired, over the years, a very refined perceptual category system for colors so that she can perceptually recognize quite narrow, close-to-minimal shades.

It begins to seem now that there are problems with Block's anti-representationalist argument. This is enough reason for me to leave off the detailed discussion of Block's paper at this point. He takes it that the last and best chance of representationalism to cope with individual differences is to take a view he calls subjective extensionalism (i.e., subjective extensionalist representationalism: Block,

1999, pp. 55-56). This is the view that the content of color experiences is determined by their extensions: the set of stimuli that evokes, under optimal circumstances, a particular color experience. Subtle individual differences in phenomenology then are coupled with small individual differences in the extensions of color experiences. Such extensions are person-relative to an extent, and so are the contents of color experiences. Block argues that this view is inconsistent as the notion of optimal circumstances presupposes that of objective color – the two are interdefined. (Very roughly, optimal circumstances are the circumstances in which objects look to be the colors they objectively are, and the objectivity of color derives from the sameness of color perceptions of color-normals in optimal circumstances: see Block, 1999, p. 45 and above). This makes a subjectivist version of representationalism, that nevertheless appeals to optimal conditions, a shaky ground (see Block, 1999, 56-61 for detail).

I do not need to tackle this part of Block's argument, as I already pointed out gaps in the earlier part of his reasoning. Subjective extensionalism is *not* the last chance of representationalism, as earlier representationalist objections in his paper received inadequate replies, and hence seem to survive. In sum, I think Block is not entirely successful in refuting representationalism (even though he had all the necessary premises in). So there is room for a reply, and Tye (2000, pp, 89-93) correctly notices this.

3.3. Tye's reply to Block

Tye's reply assumes the details of his representationalist account: his notion of perceptual representations, and his type physicalist definition of object color in terms of surface reflectance. In his reply, Tye does not consider Block's alternative definition of

objective color, or his subjective representationalist moves at all. He begins with the more difficult case, that of narrow shades. The point of his reply is the following. We know that human color vision cannot nearly discriminate every distinct surface reflectance from all others. We can discriminate, in color vision, only metameric classes of SSRs. If Max and Samantha look at the same color sample, the sample has a determinate surface reflectance, but neither Samantha, nor Max represent it as having *that particular surface reflectance*. Their perceptually determinate color experiences both represent the surface reflectance of the sample by some of its generalizable properties: *as belonging to a metameric class*. Correspondingly, many other SSRs are possible that would elicit exactly the same color experience in Max; the same is true of Samantha. Now, if Samantha's and Max's metameric classes in question do not comprise the same determinate reflectances, that is, these classes do not correspond to exactly the same reflectance type, then the properties (reflectance types) that Max and Samantha represent the sample as having will be slightly different as well. But that means a difference in representational content, and hence the way is open to admit a small interpersonal difference in phenomenal character between the two. Next problem: does one of them necessarily misrepresent the sample? Answer: no, if the two metameric classes, or narrow ranges of reflectances, that Samantha and Max represent the sample as belonging to, are overlapping. This is because in such a case the two reflectance properties that uniquely characterize the metamer sets of Samantha and Max respectively[83] are compatible, not mutually exclusive – there are particular reflectances that belong to both types at the same time (one such reflectance is that of the sample they both look at in the example). For instance, Samantha may be a better color perceiver, and so the range of reflectances

that she represents the sample as belonging to may be part of the range of reflectances that Max represents the sample as belonging to. Or the two ranges may overlap neither including the other. That is, *there might be differences between different color-normals in how their perceptually determinate color experiences carve up the set of natural reflectances into metameric classes or (sub)sets*. Tye's key example is that some subjects have better color-discrimination ability than others, and it is reasonable to assume that they correspondingly have a richer realm of phenomenal color experiences (i.e., a larger number of distinct color experiences). This treatment of individual differences in color perception solves another apparent paradox. Samantha and Max see the sample as having slightly different shades, and both their perceptions are veridical, so does the sample have two different shades at the same time? Yes, it can, if at least one of the shades is nonminimal – if it comprises shades that at least some trichromats would be able to discriminate (Tye, 2000, p. 91).

Notice a subtlety. It might appear to someone that Tye's representationalist account of individual differences also has a subjectivist (representationalist) flavor to it. So if Block indeed has an argument to the inconsistency of subjectivist representationalist views, then perhaps that argument applies to Tye's solution as well. I think this is wrong because in Tye's approach objective color and optimal circumstances of perception are defined independently of each other[84]. Redness for Tye is not *defined* as the property that looks red to most or all trichromat humans in normal circumstances (remember, this is Block's definition), rather, for him, redness is a type of surface reflectance. For Block, the explanation goes like this: ripe tomatoes are objectively red *because* they look red to most or all trichromat humans. For Tye, the explanation goes in the reverse direction: ripe

tomatoes look the way they do colorwise *because* they have the surface reflectance (color-) property they do. The way Block defines objective color seems to me to miss the externalist point in representationalism.

As to broad color categories, Tye simply repeats Block's case of individual differences in perceptual color categorization (Tye, 2000, pp. 92-93; Block, 1999, pp. 44-45 and below in his paper): different color-normals may differ in how they perceptually categorize narrow shades. He does not discuss the objections and the representationalist replies that Block conjures up. This turns out to be not a serious problem since, as we saw, Block's anti-representationalist objections in the broad color categories case are, to put it mildly, not one hundred per cent convincing (see the previous section). In addition, Tye has a forceful reply in the narrow shades case, one that deserves careful treatment. This is what I shall provide next; the anti-representationalist counterattack I'm going to offer is intended to apply both against Byrne and Hilbert's and Tye's defenses of the view.

3.4. Reply to Tye, Byrne and Hilbert

Tye (2000, p. 69) formulates the following weak representationalist thesis for the visual modality:

(R) Necessarily, visual experiences that are alike with respect to their representational contents are alike phenomenally.

It is obvious that this is the key thesis that is at stake in the representationalist-phenomenalist debate. Since actual cases are possible in any sense of possibility, if empirical data uncovers scenarios in which phenomenal color experience differs, but there is no room for color-content-difference assignment, then we can conclude that phenomenal variation is possible without variation in perceptual content, and hence phenomenal character cannot be identical with perceptual content. As we have seen from the foregoing review, there are a few theoretical complications that do not make it a very simple task to find uncontroversial cases of content-phenomenal character freeplay. Still, I think individual differences in color vision provide a clear example of such freeplay, and this is what I wish to demonstrate. In what follows I continue to assume Tye's approach to object color, namely that broad color categories, narrower nonminimal shades and minimal shades are all types of reflectance.

What I shall show in what follows is that there can exist an inter-subjective shift between narrow ranges of reflectance and perceptually maximally determinate color experiences. In other words, I shall demonstrate that there exist normal, trichromat subjects (I continue to use the imaginary example of Max and Samantha here) such that the narrow range of reflectances (say, SR_1) that Max sees as unique green is different from the one (say, SR_2) that Samantha sees as unique green. Moreover, SR_1 and SR_2, are non-overlapping, that is, there is no particular surface reflectance that belongs to both ranges. That is, there is no unique narrow range of reflectances that all trichromat humans see as unique green (*mutatis mutandis* for other narrow shades). This indeterminacy results in representational-content-independent changes in the phenomenal

characters of color perceptions between different color perceivers, thereby falsifying Tye's thesis **R**.

In achieving this goal, first I point out in more detail how the philosophical issue of the correctness of **R** comes down to an empirical issue. This requires some theoretical, or philosophical groundwork. Second, I describe an experiment that is designed to demonstrate that the inter-subjective shift that is sufficient to falsify **R** indeed obtains.

3.4.1. A model of perceptual color categorization

Tye (2000, p. 92) writes:

> … there are visual representations both of colors and of (more or less narrow) shades. Since colors comprise or include many shades, in representing X as having a certain color, *my experience effectively classifies it* along many other things whose color shades I can discriminate from X. Such classifications will certainly vary from person to person, and these classifications will be reflected in differences in verbal and nonverbal behavior in certain situations. (My italics.)

According to this passage, there exist two distinct levels of visual representation: that of colors, and that of shades. On p. 91 Tye says that shades can be minimal or nonminimal, where a minimal shade is one for which there exists no other shade that is a shade of it. For instance, when I look at a particular tree leaf in the summer noon daylight, I see a perceptually (maximally) determinate shade of green, i.e., have a perceptual experience of a particular shade of green, say G_{23}. (The experience of G_{23} I will call green$_{23}$.) There are no instances of *the shade G_{23}* (a narrow range of reflectances, or a set of metameric surfaces) that I can discriminate from each other; accordingly, there are no phenomenally distinct color experiences all of which are green$_{23}$ experiences. In other words, minimal shades pretty much correspond to *perceptually determinate color*

experiences (hereafter PDCEs). PDCEs are color experiences that we undergo when looking at particular color stimuli in normal illumination – green$_{23}$ is an example.

Nonminimal shades are categories of minimal shades: wider ranges of reflectances, or sets of different metamer sets. Any *perceptual* experience of a nonminimal shade (e.g., a perceptual experience of green) is *ipso facto* a perceptual experience of a minimal shade, e.g., G$_{23}$. Perceptual color experiences are maximally determinate, at least in normal circumstances of perception. (Barring, perhaps, faint chromatic perceptions in very low lighting.) This is because what we can directly perceive at a given time and spatial location are particular reflecting surfaces ("maximally determinate reflectances"), not any *categories of reflectances* (and nonminimal shades are categories of reflectances).[85] So, in order to perceive a nonminimal shade (e.g., green, or lime green), or have the perceptual experience of a nonminimal shade (e.g., the experience of green or that of lime green), we have to perceive a particular reflecting surface (a surface that is a member of the metamer set which in turn is the minimal shade in question, e.g., G$_{23}$), and thereby have the perceptual experience of a minimal shade (e.g., green$_{23}$). Since G$_{23}$ is green, in perceiving G$_{23}$ one perceives green; in undergoing green$_{23}$ one undergoes an experience of green.

Memories formed of perceptual color experiences, however, are much less vivid than perceptual color experiences themselves. *Perhaps* in recalling or imagining colors one can have the experience of a nonminimal shade without *ipso facto* having the experience of a minimal shade (i.e., a PDCE). Perhaps the faintness of color memories is just this: having the experience of, say, green, without further specification – without the faint experience necessarily being a PDCE like green$_{23}$. As it has been observed by

different authors (Raffman, 1995, pp. 294-295; Tye, 2000, pp. 11-12), our memory for colors is not very accurate. We certainly cannot memorize PDCEs directly, only wider or narrower ranges of PDCEs. (We cannot recognize, based on memory recall, minimal shades: e.g., we cannot recognize the paint color in the paint store that exactly matches the color of the wall in our room.) That is, minimal shades we perceptually represent by PDCEs; nonminimal shades (in general, perceptual color categories) we represent by color category representations that consist of PDCEs.[86] Our memory, it seems, can store only color category representations, not PDCEs themselves.

From the discussion so far it should be clear that PDCEs do not classify together shades that we can distinguish from each other. But, according to Tye, one's color experience (at some other level) does classify together shades that one can distinguish from each other. It seems that we need a little model of perceptual color categorization to proceed further. Based on a number of authors (Raffman, 1995, pp. 294-295; Byrne and Hilbert, 1997, pp. 265-267; Block, 1999, pp. 44-45; Tye, 2000, pp. 11-12), I propose the following two-level scheme. Particular surface reflectances that are picked up by color vision are first categorized by PDCEs: these color experiences correspond to metamer sets. Simplifying matters somewhat (i.e., ignoring illumination dependence in metamerism), given a metamer set M, each one of its members looks the same color (to trichromat humans). This categorization of surface reflectances is inherent in perceptual color experience: those things that are categorized together (i.e., members of metamer sets) cannot be discriminated by our color vision. PDCEs categorize surface reflectances; in addition, as we saw above, PDCEs are not memorizable.

But we can perceptually memorize, recall, and imagine colors to an extent. That is, our color experiences can be stored in memory – not just conceptual/propositional memory as color names, but also as perceptual experiences. However, as we know well, color experiences, when recalled, are much less vivid than they are when we actually perceive colors. Moreover, our perceptual memories of colors enable us only for a much coarser discrimination than does color perception (recall the paint-store example above). Therefore, a second level of color experiences has to be posited: we might call this level *(perceptually) memorizable color category representations*, or simply color category representations. Color category representations are either ranges of PDCEs or, *perhaps*, single faint color experiences that are less specific than PDCEs, and they represent only what is common, perceptually, to certain ranges of PDCEs. An example: try to recall, from memory, the experience of green. The recalled experience is much fainter than most perceptual experiences of green, but it may still convey what is common to all and only PDCEs of green colors. Correspondingly, each color category representation represents a wider range of reflectances (a set of metamer sets), sub-ranges of which (individual metamer sets) we can discriminate from one another. In Tye's terminology, colors and nonminimal shades are both sets of metamer sets, and they are specifically represented by color category representations (for narrow non-minimal shades only experts have color category representations; for colors like red, green, etc., all trichromats have). PDCEs specifically represent individual minimal shades. PDCEs can also be taken to represent, indirectly, or by implication, nonminimal shades or colors, just like the concept *dog* represents (indirectly, by implication) the property of being a mammal. Color category representations are indeterminate with respect to minimal shades, so they do not represent

specific minimal shades (just like the concept *mammal* does not represent dogs). Contrary to PDCEs, color category representations can be stored in perceptual memory, recalled, and used in imagination or coarse discrimination (e.g., given that my car is silverish gray, on seeing a yellow car that is exactly like mine in other respects, I will discriminate it from my car, based on my color memory – i.e., making a memory-based color discrimination).

Not implausibly, conceptual color categories and ordinary color names are attached to wider or narrower color category representations (e.g., red scarlet, cadmium red, green, lime green, etc.), but not to PDCEs. There simply are too many different PDCEs – and too many minimal shades – to name each. So, based on color category representations coupled with color names, we can make memory-based, or "off-line" reports of our color perceptions. We *remember* that Granny Smith apples are light green, Bosc pears are yellowish brown, and so on. I think it is also plausible that at the time of color perception – while our perceptual color-experience lasts, as we look at a stimulus – we can make finer verbal reports on our experience (a PDCE). This we might call "on-line perceptual report": I think the more specific information that is delivered by PDCEs, that is, in most cases, lost for memory storage, is available for conceptual re-coding if it is caught while perception lasts. This is what happens in experiments of the Sternheim and Boynton type (Sternheim and Boynton, 1966; Quinn et al., 1988; Miller, 1997), where the subjects are asked to characterize perceived colors by the aid of unique hue names and percentages (e.g., a predominantly green surface with a tinge of yellow may be characterized as 90 per cent green, 10 per cent yellow).

A color category representation (e.g., that of green) represents the generalizable properties of the shades that are also represented by all the PDCEs (among them, green$_{23}$) falling into the color category representation in question. In Byrne and Hilbert's terms (Byrne and Hilbert, 1997, p. 267), green is a determinable and lime green is a determinate of it; a particular minimal shade of lime green (e.g., G$_{23}$) is a determinate of the narrower color category lime green. The color category representation for green represents various greens; those with different lightness and saturation; slightly yellowish greens, and slightly bluish greens. Despite these differences, the color category representation for green represents all these shades *as green, full stop*; further details about green object colors are represented by narrower color category representations, or PDCEs like green$_{23}$. What color category representations do is abstraction of a relatively simple sort. Once again, only information that color category representations take over from PDCEs belonging to them can be preserved in memory. Of course, there are color category representations for binary hues as well. In addition, certain color category representations can overlap, or have fuzzy boundaries: those for green and bluish green are an example. There might even be differences between different subjects in these overlaps, and the span of their color category representations in general: what I would judge green (perhaps with a negligible trace of blueness), you might judge a more pronounced bluish green.

3.4.2. Incompatibilities

Given the above model, the important point to start with is that neither PDCEs, nor color category representations can represent any object color as, for instance, *both bluish and yellowish*. There exists a three-link chain of incompatibilities with regard to

color and color experience. First, it is a perceptual fact that no surface area can look, to any trichromat human subject, simultaneously bluish and yellowish; likewise for reddishness and greenishness. From this *intrasubjective perceptual incompatibility*, there follows a consequence with respect to object color: there is no such object color as reddish green or bluish yellow. More generally, no object color can be both bluish and yellowish (or both reddish and greenish). Reddishness and greenishness are contraries, or incompatible color properties, as are bluishness and yellowishness. We might coin the term *objective incompatibility* to refer to this relation between complementary colors in general. Were this parallel assumption about object color not made, the access given by color experience to object colors would seem rather strange. Type physicalists about color who maintain that (1) object colors are universals, or natural kind essences (see Tye, 2000, pp. 103, 124-125, note 4 on p. 167) and that (2) object colors play a key role in the determination of the phenomenal character of color experience should, I think, hold that incompatible color perceptions correspond to, or represent, physically incompatible object color properties – properties that cannot combine, or are mutually exclusive (just like their perceptions). If one denies that objective incompatibility follows from intrasubjective incompatibility (plus type physicalism about color) then the question that immediately arises is: if there is such an object color as bluish yellow, then why can we not perceive it as it objectively is, i.e., as bluish yellow? And exactly how do we perceive bluish yellow? The analogy here is being circular vs. being rectangular. These are objectively incompatible properties, and, as a consequence, it is also true that we cannot perceive any object as being both circular and rectangular. (Whereas we *can* perceive an object as both square and diamond-shaped: see Byrne and Hilbert, 1997, p. 274.) If

objective incompatibility is denied, then opponency becomes an arbitrary aspect of color experience that is not grounded in any relevant, corresponding aspect of objective reality – an aspect of perception that depends entirely on the internal constitution of perceivers. That is, although we perceive bluishness and yellowishness as incompatible, still they are, objectively, compatible properties – this is just meant by the denial of objective incompatibility. Needless to say, such an admission would mess up the type physicalist view of color and color perception quite a bit.

If this is not enough argument for objective incompatibility, here is a little more. As Byrne and Hilbert remark (1997, p. 274), a patch that looks unique green can be made to look bluish green by changing the viewing conditions slightly. By the same coin, we might add, the same patch can be made to look yellowish green by changing the viewing conditions in a different way (still staying in the broadly normal range of illumination). This means that intrasubjective incompatibility obtains only for single events of color perception, not for different color perceptions of one and the same subject. The same subject can perceive the same stimulus as either bluish green (at time t_1) or yellowish green (at t_2), but for this a change in the circumstances of perception is needed. Intrasubjective incompatibility means only that we cannot *simultaneously* perceive a surface as both bluish and yellowish (whereas we can perceive it simultaneously as both bluish and greenish).

Now, a physicalist about color may want to avoid endorsing objective incompatibility by appealing to this phenomenon.[87] This is possible if one abandons the claim that object colors are inherent properties of surfaces (like reflectance). That is, one could say that the color of objects is essentially dependent on illumination – it is a

relation between reflectance and illumination, or simply, color is the color signal (the light actually reflected by surfaces). This still would not give one the result that the same surface is *simultaneously* bluish and yellowish, since if the illumination changes, the color signal changes too, and that amounts to a change in color if color is defined as the color signal. If one denies objective incompatibility on the basis of the above phenomenon (i.e., that changing illumination might turn a bluish-green-looking object into a yellowish-green-looking one), but maintains that color is reflectance (i.e., an inherent property of surfaces), then the strange consequence will be that while the illumination changes (within the normal range), our (veridical) color perception of a particular surface changes, despite the fact that the color of the surface (i.e., its reflectance) does not change. In other words, a number of different PDCEs can veridically represent the very same object color, hence the mapping between reflectances and color perceptions becomes one-to-many, and for this reason it does not satisfy the conditions of being a function. There is no single color perception (PDCE) that is *the* veridical perception of a particular surface reflectance in one and the same subject. In sum, the safer way to go seems to be to accept objective incompatibility.

Tye's schema for defining colors in terms of surface reflectance endorses objective incompatibility (Tye, 2000, pp. 159-161): for instance, he claims that for a surface to be bluish, it has to reflect significantly *more* light in the short wavelength range than in the medium and long wavelength ranges together; for it to be yellowish, it has to reflect significantly *less* light in the short wavelength range than in the medium and long wavelength ranges together. These two conditions cannot be simultaneously satisfied; so there can be no object that is both bluish and yellowish.[88]

From objective incompatibility it follows that color category representations like that of bluish green and yellowish green cannot overlap – *not even in different subjects*. That is, on pain of perceptual mistake, there can be no colored stimulus that looks to Samantha bluish green whereas to Max it looks yellowish green. Since any particular surface is either bluish green or yellowish green (or neither, but not both), it follows that no surface can be veridically represented both as yellowish green and bluish green, at any level of (perceptual – or conceptual) representation. If a surface S is, as a matter of fact, yellowish green, and I represent it by activating my color category representation of bluish green, then I misrepresent S colorwise.[89] This constraint we might call *intersubjective perceptual incompatibility*, and it follows from objective incompatibility. Now if Max represents S by his color category representation YELLOWISH GREEN, this is presumably because the PDCE which S gives rise to in Max is a color perception as of yellowish green. By the same coin, if Samantha represents S by activating her color category representation BLUISH GREEN, this is normally due to the fact that S elicits a bluish green PDCE in her. But since, by assumption, S is yellowish green, Samantha misrepresents it at both levels of perceptual organization. That is, what applies to color category representations, applies to PDCEs too: given another surface K, if, on looking at K, Max's PDCE is a yellowish green perception whereas Samantha's PDCE is a bluish green perception, then it follows that at least one of them misperceives K's color. At this point it should now be obvious how the argument from individual differences proceeds.

3.4.3. The core argument

If we can show, empirically, that there is intersubjective overlap between color category representations that, according to representationalism, are incompatible, i.e., if we can show that there are color stimuli that look, say, (even slightly) bluish green for one trichromat subject whereas they look (even slightly) yellowish green for another subject, then we have the following argument against phenomenal externalism.

P1. As a matter of empirical fact, there exist pairs of normal trichromat subjects such that reflecting surface R_0 looks yellowish green to one member of the pair, whereas it looks bluish green to the other member.

P2. R_0 is either bluish green or yellowish green (or neither, but not both).

C. One member of the pair misperceives R_0 colorwise (i.e., perceptually misrepresents R_0's color).

Now, this is a *reductio* of the representationalist (phenomenal externalist) theory of color experience because, on independent grounds, the following is true: if both subjects are normal, trichromat color perceivers, then, *ipso facto*, both their color perceptions (perceptual color representations) are veridical (see Block, 1999pp. 46-47; Tye, 2000, 89-93). So, the conclusion of the just-presented argument is unacceptable. We have to acknowledge that trichromat humans in normal conditions perceive object colors veridically. Alternatively, assuming that R_0 is *de facto* yellowish green, if both subjects represent R_0's color veridically, then both their corresponding perceptual states carry a content **that** *a yellowish green surface is present*. But this content is accompanied by a

phenomenally yellowish green perception in one subject and by a phenomenally bluish green perception in the other. That is, the phenomenal character of PDCEs is incompatible with, or varies independently of, their perceptual content. That is, Tye's thesis **R** is false.

3.4.4. Objections

First objection. The argument so far goes through if R_0 is *de facto* yellowish green and it is perceived by a trichromat subject as bluish green. But if R_0 is unique green and it is perceived as either slightly bluish green or slightly yellowish green, then the argument is not so obviously right. For unique green and bluish/yellowish green, as object colors, are not so obviously contraries.[90] To exploit this option, one could say the following. In the above example, take the whole "indeterminate" range of stimuli – the whole range such that any sub-range of it is perceived by *some* trichromat perceivers as unique green – to *be* unique green. An example is the pure wavelength range between 490 and 520 nm: perhaps that whole range just *is* unique green, simply because we can find trichromat subjects to whom 490 nm looks unique green, and also subjects to whom 520 nm looks unique green (e.g., Byrne and Hilbert, 1997, p. 272). Likewise for any wavelength in between. As a consequence of this, some subjects will perceive 490 nm as slightly bluish green (e.g., those to whom 520 nm looks unique green). Some other subjects (e.g., those to whom 490 nm looks unique green) will perceive 520 nm as slightly yellowish green. It is also plausible that 505 nm will be perceived by some subjects as slightly yellowish green, whereas by others as slightly bluish green. These are really just borderline differences, the representationalist might contend; the whole range of such stimuli (like

the 490-520 nm range) has to be regarded as unique green. Since neither yellowish green, nor bluish green are incompatible with unique green, perhaps this move saves the day for the phenomenal externalist. If R_0 is such that to Samantha it looks slightly yellowish green and to Max it looks slightly bluish green, then what R_0 really is, is unique green. Correspondingly, R_0 is perceived by both S1 and S2 as *predominantly* green, and that is the key point. Slight differences in phenomenology due to individual biological factors do not really matter.

Unfortunately, this escape is not viable. The first problem with it is that the 30 nm range between 490nm and 520 nm covers 10 per cent of the whole visible range (that between 400 and 700 nm). This is somewhat too broad a stimulus range to assign to just one PDCE – i.e., to identify it with just one minimal shade.[91] Recall also the results of Ayama et al. (1987), who report even larger individual variation. Moreover, if I am a subject who finds 490 nm unique green, then to me, 520 nm will plausibly look like a *substantially* yellowish green, not just *almost unique green with a hardly noticeable tinge of yellow*. Between 490 and 520 nm I will be able to make a whole series of chromatic discriminations, but all these discriminations on my part do not reflect any difference in objective color. That does not sound very plausible. A further problem is that if we assign such a broad range to unique green (which is just one, *chromatically* determinate shade), then all other chromatically determinate shades like unique yellow, blue, red, and all the binary hues (of which there is quite a number) should in principle be regarded as comparably broad stimulus ranges. This could render individual phenomenological differences insignificant. However, it seems that there simply is not enough room in the whole relevant stimulus range to accommodate such wide color assignments. Remaining

with the example of spectral colors (i.e., pure-wavelength-emitting surfaces), out of 30-nm-wide wavelength intervals like that assigned to unique green, only ten would fit into the whole visible range (400-700 nm). But we can perceptually discriminate a little more than ten chromatic shades within the range of monochromatic lights. This leaves us unable to assign unique physical correlates (in this case, pure wavelengths emitted by a surface) to the large number of different PDCEs that are elicited by different pure wavelengths. Even though, as Kuehni (2001) remarks, with reflective colors one finds smaller inter-subjective variation, still, these variations are large enough to support the argument. Kuehni (2001, p. 63) found that individual differences in locating unique hues can be up to 4 Munsell 40 hue steps. Four neighboring Munsell chips are not reasonably regarded as being the same color (e.g., unique green) exactly because their surface reflectances differ and most trichromat subjects can discriminate them, despite individual differences in how different subjects would locate unique hues on them.

A related objection can be reconstructed from the remark of Byrne and Hilbert (1997, p. 274): "But even if bluish green and unique green are in fact contraries, this is not a disaster. That many of us misperceive unique green objects is certainly an unwelcome result; but at least (for all the objection says) we veridically perceive them as green, and perhaps that is enough." Perhaps, indeed, this reply also applies to my case with yellowish green perceived as bluish green. Even if yellowish green and bluish green are contraries as I contend, Byrne and Hilbert could still say that seeing a predominantly green surface with a tinge of blue in it as a predominantly green surface with a tinge of yellow in it is a predominantly veridical color perception with a tinge of non-veridicality in it.

Again, this reply misses the point of the critic. The point is that if such cases do indeed obtain, then there just are no (narrow) stimulus property ranges *across different trichromat perceivers* that uniquely correspond to maximally determinate hue perceptions (PDCEs). The reason for this is not simply that one perceiver makes finer discriminations than the other (Tye, 2000, pp. 91-92). If it is also true that different perceivers' assignments between narrow ranges of reflectance and color perceptions are shifted compared to each other, that would entail that PDCEs do not have unique physical correlates in terms of (narrow ranges of) reflectance. In other words, if (1) two narrow, non-overlapping ranges of reflectance[92] R_1 and R_2 are such that to subject S1, reflectances in R_1 look the same color as reflectances in R_2 look to another subject S2 and vice versa, moreover (2) both subjects can discriminate reflectances in R1 and R2 from each other (for instance, reflectances in R_1 look to S1 as unique a green as there can be, whereas reflectances in R_2 look to S2 as unique a green as there can be; plus, reflectances in R_1 look bluish green to S2, whereas those in R_2 look yellowish green to S1) that means that PDCEs have no unique physical correlates in terms of ranges of reflectance. In this case, unique green **is** *either R_1 or R_2*, depending on which subject we consider.

Byrne and Hilbert (1997, p. 274) also mention that if a patch looks unique green, it can typically be made to look bluish green by changing the viewing conditions slightly. This apparently gives rise to a third (version of the already discussed) objection: if a surface reflectance looks slightly bluish green in one viewing condition, then it can be made to look slightly yellowish green by modifying the viewing conditions a bit. This is plausible. So perhaps even slightly yellowish green and slightly bluish green are not really contraries, one might want to add.

The problem again is to find the unique (cross-subject) physical property correlates of the PDCEs in question: the particular slightly yellowish green and slightly bluish green perceptions. In this version of the argument it turns out that there is no one-to-one, or many-to-one mapping between surface reflectances alone and PDCEs: the same surface reflectance can look, to the same subject, slightly bluish green under one *normal* illuminant, and slightly yellowish green under another *normal* illuminant. Similarly for different, but entirely normal simultaneous contrast effects: a particular reflecting surface can, in one context, look slightly bluish green, but when placed in a different layout, it may look slightly yellowish green, both contexts being entirely normal. So the question 'which particular, narrow range of surface reflectances *is* the shade G_{23}?' seems to lose sense. At this point, the representationalist might want to abandon the claim that color (or at least *minimal shade*) is an inherent property of surfaces (e.g., Tye, 2000, pp. 147, 153), and allow that the minimal shade C of an object O is a relational property: a complex (and currently unknown) relation between O's surface reflectance, the illuminant, and the surface reflectances of objects surrounding O (see Tye, 2000, 152-153). For the moment, let us ignore context effects, and say that object color is the same as the color signal.[93] Given a fixed neutral (e.g., mid-gray) background, and a particular illuminant, variations in surface reflectance are indeed specifically responsible for variations of color perception of subjects. This way perhaps we can identify the unique physical correlates of PDCEs (and broader ranges of *reflectance alone* may still work for nonminimal shades, or color categories).

However, even this hope is misguided. For the whole argument that I made above assumes that subjects like Max and Samantha look at a particular surface with reflectance

R_0, under the same illuminant (and in the same color context). That is, it is the same color signal that, say, Samantha perceives as bluish green and Max perceives as yellowish green – this is the assumption I make. Identify minimal shades with ranges, or types of the color signal, and the whole argument from individual differences in phenomenology can be run on these kinds of stimuli. If Max and Samantha have incompatible color perceptions on looking at R_0 *under the same illuminant and against the same background*, that means that the same color signal evokes incompatible PDCEs like slightly bluish green and slightly yellowish green in these subjects. I.e., not even in terms of color signals can we carve out minimal shades (i.e., assign unique physical correlates to PDCEs).

3.5. The question for experimental assessment

The goal of the experiment described below is to demonstrate individual differences in color perception of the kind outlined in the previous sections. To make clear what sort of result would support my argument, here is a list of tables accompanied by explanation. Table 4 shows the individual differences scenario that can be accommodated by Tye's account – indeed, this is the scenario that he himself offered in reply to Block's objections. The subsequent tables show gradually more serious problem cases that cannot be accommodated by Tye's theory. If data from the present experiments show that one or more of the problem cases actually obtain, then I have the empirical premise for my version of the argument from individual differences. First, here is the theoretical possibility that Tye proposed to accommodate individual differences in his schema:

		Stimuli		
		SR$_1$	SR$_2$	SR$_3$
Subjects	Max	UG	UG	UG
	Samantha	BG	UG	YG

Table 4. Individual differences of the kind that can be accommodated by Tye's theory. See the main text for detail.

In Table 4, Samantha has better color-discrimination than Max. SR$_1$, SR$_2$, and SR$_3$ are stimuli, say, three neighboring samples of some color-order system, characterized by their surface reflectance. Max, who has poorer color discrimination, sees all three as unique green, without a trace of blue or yellow in it. Samantha, however, sees SR$_2$ as unique green, SR$_1$ as bluish green, and SR$_3$ as yellowish green.[94] As we saw above (in Section 3.3), in such a case Max's perception corresponds to a non-minimal shade as his unique green experience spans over a wider range of reflectances that includes the three samples that are discriminable for a substantial proportion of color-normals. Max's single color experience on looking at the samples has content that differs from the content of each one of Samantha's experiences. (I.e., Samantha's unique green experience does not have the same content as Max's unique green experience.) Max's experience represents SR$_2$ as belonging to a wider range of reflectances (that includes the other two samples as well), whereas Samantha's experience represents SR$_2$ as belonging to a narrower range of reflectances that does not include the other two samples. Therefore, even the unique green experience of the two should be phenomenally slightly different, as suggested by

the phenomenal externalist view. It might be quite difficult to spell out such a phenomenal difference, but still, the idea that there is such a difference need not be too implausible. Samantha's other two color experiences differ from Max's unique green experience in ways that are easier to conceptualize: on looking at SR_1, Samantha sees bluish green whereas Max sees unique green; on looking at SR_3, Samantha sees yellowish green whereas Max again sees unique green.

Note one thing: it is an empirical question whether interpersonal differences of this kind actually obtain. In the experiment that I will describe below, I found no indication of such a difference. From this it does not follow that no other experiment could demonstrate difference of this kind.

Table 5 shows the first problem case for Tye's theory:

		Stimuli			
		SR_1	SR_2	SR_3	SR_4
Subjects	Max	B" G	B' G	UG	Y' G
	Samantha	B' G	UG	Y' G	Y" G

Table 5. The first problem case for Tye's theory: one-step shift. See the main text for further explanation.

As in the previous case, SR_1-SR_4 are color stimuli characterized by their surface reflectance, say, four neighboring samples of a color order system. B''G stands for a bluish green perception with a more pronounced bluish component, whereas B'G is a bluish green experience with a less pronounced bluish element (similarly for Y'G and

Y''B). The key feature of this scenario is that SR_2 looks unique green to Samantha but bluish green to Max whereas SR_3 looks unique green to Max and yellowish green to Samantha. Note that both subjects can discriminate these two samples (indeed, all four of them), but they disagree on the location of unique green. So this is not a case in which individual differences in color perception come from individual differences in the ability to discriminate colors. There need not be any difference in Samantha's and Max's ability to discriminate colors in this case. What looks unique green to Samantha looks bluish green to Max, and what looks unique green to Max looks yellowish green to Samantha. This indicates a shift between the two in the assignment of phenomenal characters to object color (reflectance) properties. Still, this is a relatively weak problem case because none of the samples is perceived by the two subjects as having incompatible colors (e.g., there is no sample that is perceived by Samantha as bluish green and by Max as yellowish green). I will call this case *one-step shift* in what follows.

Table 6 shows a stronger problem case:

		Stimuli				
		SR_1	SR_2	SR_3	SR_4	SR_5
Subjects	Max	B''' G	B'' G	B' G	UG	Y' G
	Samantha	B' G	UG	Y' G	Y'' G	Y''' G

Table 6. A stronger problem case for Tye's theory: two-step shift. See the main text for explanation.

This case will be called *two-step shift*, because the unique perceptions of Max and Samantha are two samples away from each other on the hypothetical color-order system. Here, one of the stimuli, namely SR_3, is perceived in incompatible ways by the two subjects: to Samantha, it looks yellowish green whereas for Max, it looks bluish green. As I argued above, in such a case the type physicalist must rule that either Max or Samantha misperceives the sample's color, and this ruling is, on independent grounds, unacceptable. Note also that this case includes the one-step shift as well: here, SR_2 and SR_4 are in the same relation as SR_2 and SR_3 in Table 5. If we delete the SR_3 column from Table 6, it becomes equivalent to Table 5.

Here is a third problem case that we might call *one-plus-two-step shift*:

		Stimuli				
		SR_1	SR_2	SR_3	SR_4	SR_5
Subjects	Max	B''' G	B'' G	B' G	UG	Y' G
	Samantha	B' G	UG	Y' G	Y'' G	Y''' G
	Eva	B'' G	B' G	UG	Y' G	Y'' G

Table 7. A third problem case: one-plus-two-step shift. See the main text for explanation.

This case includes the other two in the same way as the two-step shift includes the one-step shift. (If we delete the third row of this table – Eva's data – then what remains is the same as Table 6.) In addition, this scenario is a stronger hint at the one-to-many mapping between narrow ranges of reflectance and perceptually determinate color experiences: SR_3 looks unique green to Eva, yellowish green to Samantha, and bluish green to Max.

The three problem cases together could demonstrate that perceptually determinate color experiences do not have unique correlates in terms of narrow ranges of reflectance (metamer sets). Instead, reflectance – color perception correlations show a substantial between-subject variation. For instance, unique green is one of a narrow range of reflectances for one trichromat subject, and another, non-overlapping narrow reflectance range for another trichromat subject. The other three unique chromatic hues, and presumably, binary hues as well, would have similar properties. Consequently, a given reflectance (or metamer set) can look different in color to different color-normal subjects, in such a way that differences in the phenomenal character of their color experiences cannot be cashed out in terms of perceptual content. In other words, if this is the case, then phenomenal character varies independently of perceptual content, hence Tye's thesis **R** is false.

3.6. Method

Subjects and location. Fifteen subjects (9 females, 6 males, all between 25 and 40 years of age) participated in a unique hue choice and color naming experiment. Seven of the female subjects and five of the males were associated with the Cognitive Science program at Carleton University. Ten subjects were native English speakers, five were non-native speakers. All subjects were trichromats: their color vision was checked by the Pseudo-Isochromatic Plates.[95] The experiment took place in the National Research Council of Canada, Institute of National Measurement Standards.[96] In what follows, I will refer to the female subjects by symbols F1-F9, to male subjects by M1-M6.

Tasks. The experimental session consisted of three tasks. First, subjects were administered the Pseudo-Isochromatic plates, then they were asked to adjust the D&H Color Rule[97], which is a means of assessing individual differences in color perception. Both these tasks were done in a Macbeth lighting booth, under (artificial) daylight illumination.

The third and principal task was to complete a series of color naming and unique hue choice tasks on a computer monitor. The application for this task was developed by the author using the Delphi code-builder software. Throughout the task subjects proceeded by pointing and clicking with a mouse. Forty experimental displays were shown, one at a time. On each display the subjects saw, in a row, nine differently colored squares. There were four types of color series: one consisted of reddish colors, one of greenish ones; a third one contained yellowish colors, and the fourth, bluish ones. First the subjects had to choose, from each presented series of colors, the member that was the purest example of either red, or green, or yellow, or blue (i.e., the task was to choose the

purest red from a red series, the purest green from a green series, and so on). Selections were made by clicking on a button located under the color sample chosen. The subjects were instructed that if they could not find an absolutely pure example of the required category (red, green, yellow, or blue) in a display, then they should choose the one that was the closest. After the unique hue was chosen, subjects named the remaining eight colors in each display by choosing a name from a list of eight color categories occurring under the color samples. After the unique hue choice in a given display was made, the selection buttons under the nine color samples disappeared, and eight color category names were presented under each one of the eight colors that the subject did *not* choose as the purest. (Changing the unique hue choice, and thus restarting the display, remained possible until the subject completed the naming task and proceeded to the next display.)

Subjects selected the appropriate color categories for the remaining eight samples by pointing and clicking on them. They were allowed to select only one name for each color. The following eight color names were offered: Red, Green, Yellow, Blue, YelRed (Yellow-Red), YelGreen (Yellow-Green), BluRed (Blue-Red), and BluGreen (Blue-Green). Subjects were instructed to use 'YelRed' to name colors that look like orange (either lighter, more yellowish oranges, or darker, more reddish oranges); 'BluRed' to name purplish colors: both lighter, more reddish ones like magentas, and darker, more bluish ones like a purplish blue. The name 'YelGreen' was used to name the range of colors from yellowish green to greenish yellow; 'BluGreen' was used to name colors that are either bluish greens or greenish blues. Subjects were encouraged to use the same category name for adjacent color samples in the display if they found it adequate – that is, if they found that neighboring colors deserve the same name simply because the eight

color names supplied were not enough to precisely name the whole variety of colors presented. Subjects were tested individually. Completion of all three tasks required approximately 45 minutes.

Stimuli. It was a key requirement to precisely display the colors on a computer monitor. For this purpose I used a set of functions developed by Rejean Baribeau for transforming the color coordinates of the stimuli into RGB values (and back). In order to assure precise color display, a 21-inch Sony FD Trinitron monitor was used. The monitor was calibrated prior to data collection, and on each day after the experimental sessions all stimuli were separately displayed and measured by a multi-channel spectroradiometer. (On the first day of data collection stimuli were measured both before and after the sessions.) These measurements showed that, over the 17 days period in which data from subjects were collected, the monitor was highly consistent: for all stimuli used in the experiment, the variation of the x, y chromaticity coordinates due to "drifting" of the monitor was within 0.003; the variation of the Y coordinate was within 0.3. In total, 52 color stimuli were used, divided into four color series with 13 members each. The red series contained colors from violet/purple through red to orange; the green series contained colors from yellowish green through green to bluish green; the yellow series contained colors from orange through yellow to greenish yellow; finally, the blue series contained colors from greenish blue to purple. The color coordinates of the stimuli in L*C*h color space were as follows.

Red series: L*=55, C*=60 for each member; hue angles in degrees of samples 1 to 13: 336, 344, 352, 0, 8, 16, 24, 32, 40, 48, 56, 64, 72.

Green series: L*=55, C*=60 for each member; hue angles in degrees of samples 1 to 13: 130, 136, 142, 148, 154, 160, 166, 172, 178, 184, 190, 196, 202.

Yellow series: L*=70, C*=70, hue angles: 54, 60, 66, 72, 78, 84, 90, 96, 102, 108, 114, 120, 126.

Blue series: L*=50, C*=50, hue angles: 182, 195, 208, 221, 234, 247, 260, 270, 280, 290, 300, 310, 320.

Figure 3 shows the distribution of the chromaticities of the 52 stimuli in the x,y chromaticity diagram:
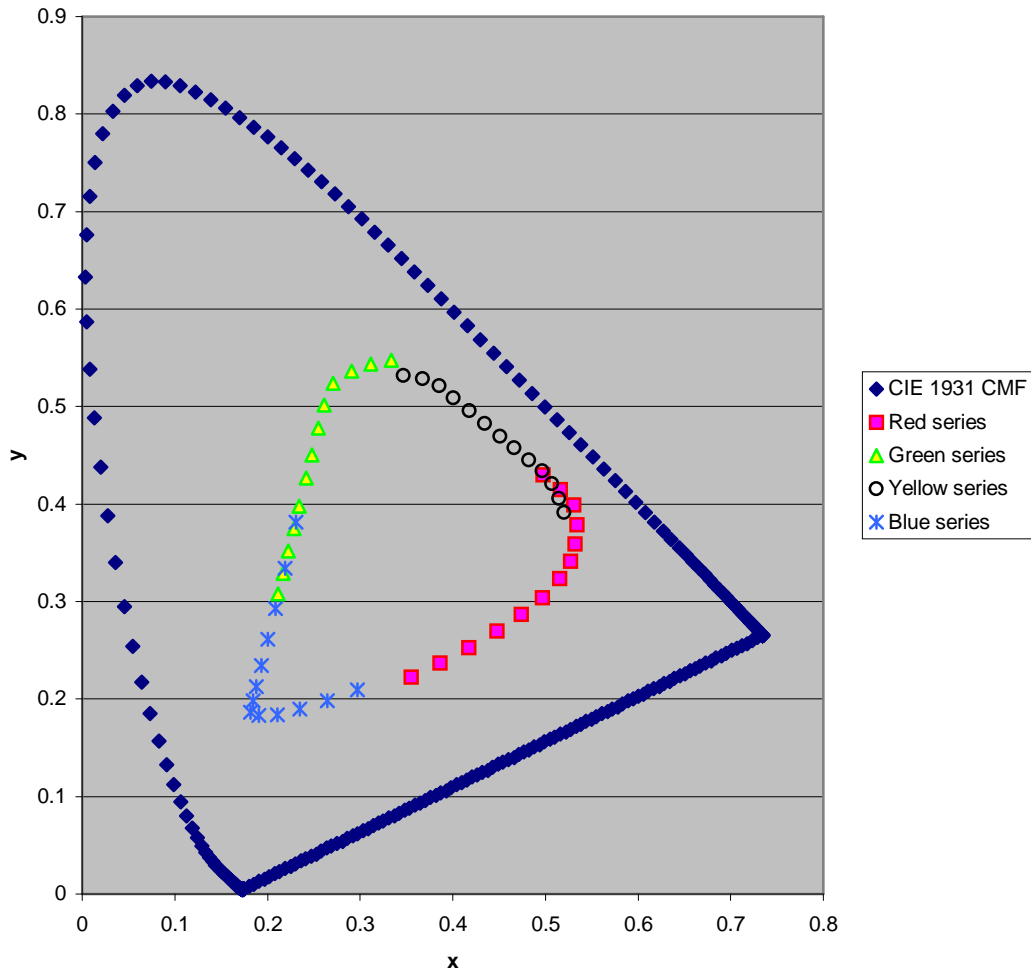
Figure 3. Chromaticities of the 52 stimuli in the CIE 1931 chromaticity diagram, as displayed on the monitor (i.e., displayed are the measurement data provided by the spectroradiometer). Squares: red series, triangles: green series, circles: yellow series, x's: blue series.

In what follows, I will refer to individual samples in the following way: the first sample in the red series will be called RED 1 (i.e., its color coordinates are: L*=55, C*=60, h=336), and so on.

Stimuli were displayed in the absence of external illumination. Each sample was a roughly 15x15 mm square, and the nine colored squares appeared in a joint black frame that was also 15 mm wide on each side (the separation between neighboring squares was

also 15 mm). The rest of the screen was neutral mid-gray – a chromatic match of illuminant D65. Subjects viewed the stimuli from about 40 cm, without using a chinrest.

Experimental design. Forty (40) experimental displays were presented: each color series was repeated ten times. The color samples in the series were never randomized, but the location of the series center (the 7th member of the 13-series) was varied between the 3rd and 7th position in the experimental display, which consisted of nine members. (I.e., if the center of the 13-series appeared in the 3rd position of the display, then the first four members of the 13-series were not shown, only the members 5 through 13. If the center of the 13-series appeared in the 7th position of the experimental display, then the last four members of the 13-series were not shown, only members 1 through 9, and so on.) I will refer to this feature as positioning. The aim of positioning was to prevent the subjects from choosing unique hues in the repeated trials on the basis of location or serial position.

The order of the first 20 displays was: five red series in a row, followed by five green series in a row, then five yellow series and five blue series in the same way. The second 20 displays consisted of five consecutive 'red series – green series – yellow series – blue series' patterns. For each color series and the ten repetitions, the positions of the center (7th) sample of the stimulus series in the experimental display were: 3,7,5,4,6,3,7,5,4,6. This design made it possible to compare the subjects' responses to the same color series when it was repeated in consecutive displays, and when it was mixed with displays of other color series.

3.7. Results

Unique hue choices. For each subject and each color series, the mode, mean and standard deviation of unique choice was calculated. For calculating means and SDs I treated the serial numbers of the samples in each series (1 to 13) as an interval scale. This was reasonable because the distances in L*C*h color space between two adjacent samples within a series were roughly equal. Thus the differences between adjacent samples corresponded to equal perceptual steps, and hues in between any two samples could be obtained by interpolating their chromaticity coordinates. (The only slight deviation from uniformity was in the blue series where in the greenish blue range the steps between adjacent samples were 13 degrees in hue angle whereas in the reddish blue range steps were 10 degree each.) Data were also grouped according to the samples, and the relative frequency of unique choice for each sample was calculated. Table 8 shows the means and standard deviations of females, males, native speakers, non-native speakers, and the entire sample for the four color series.

Table 8. Group means and standard deviations of unique choices. The scale is derived from treating the sample serial numbers as constituting an interval scale (see the main text).

		Females	Males	Native speakers	Non-nat. speakers	Total
RED series	Mean	6.52	6.4	6.49	6.44	6.47
	SD	0.678	0.853	0.784	0.677	0.726
GREEN series	Mean	5.44	5.03	4.9	6.04	5.28
	SD	0.919	0.776	0.471	1.006	0.861
YELLOW Series	Mean	8.03	8.36	8.13	8.24	8.16
	SD	0.568	0.703	0.745	0.321	0.624
BLUE series	Mean	6.85	7.41	7.23	6.78	7.08
	SD	0.783	0.426	0.517	0.983	0.704

A 2-way mixed ANOVA comparing (1) color and (2) native versus non-native speakers was performed. There was a significant main effect of color ($F(3)=34.121$, $p<.001$), and no main effect of the native - non-native difference, nor was any interaction observed. In addition to the data presented in Table 8, I separated, for each subject, the data of the first 20 trials (displays of each color series in a group) from that of the second 20 trials (displays of different color series mixed). A 3-way mixed ANOVA comparing (1) the first and second half of the presentations, (2) color, and (3) gender was performed. Other than for color ($F(3)=44.711$, $p<.001$) there were no significant main effects or interactions obtained. That is, no significant difference in unique hue choices between males and females was observed. Still, at the level of descriptive statistics we find interesting differences between females and males that are worth looking at in future studies. Figure 4 below shows, for the four color series, the relative frequency distributions of unique hue choices along the color samples.

Note that the main effect of color is irrelevant for my purposes: it is caused by the fact that in different color series, the subjects' unique choices centered around samples with different serial numbers. The fact that the whole sample's grand mean was 6.473 for unique red choices, it was 5.28 for greens, 8.16 for yellows, and 7.08 for blues has to do with how I set up the four color series. (Again, the unit of measure here is *sample number* because, as I argued, sample numbers correspond to equal perceptual steps and thus define an interval scale.) It is in principle possible to construct all four color series in such a way that the grand mean of unique choices for all four of them is, say, 7.00. In practice, however, this is a little cumbersome to attain, and it would make no difference as individual differences are assessed independently for the four unique hues.[98]

The absence of a difference between the first half and the second half of the presentations is a good result: it shows that presenting repetitions of the same color series in one group versus mixing up the presentations of different color series has no effect upon the unique hue choices. This signals some sort of stability (resistance to changes in irrelevant factors) of the intersubjective differences that will be examined below. The absence of an effect of the native - non-native distinction can perhaps be taken as an indication that individual differences in the unique hue choices are presumably not due to differences in the ways different people use color names, but rather to prelinguistic, perceptual processes, and that's exactly what I want to demonstrate. (However, I agree that this issue needs further attention.)

Figure 4. Relative frequency distributions of unique choices along the samples. A: red series, B: green series, C: yellow series, D: blue series.

A
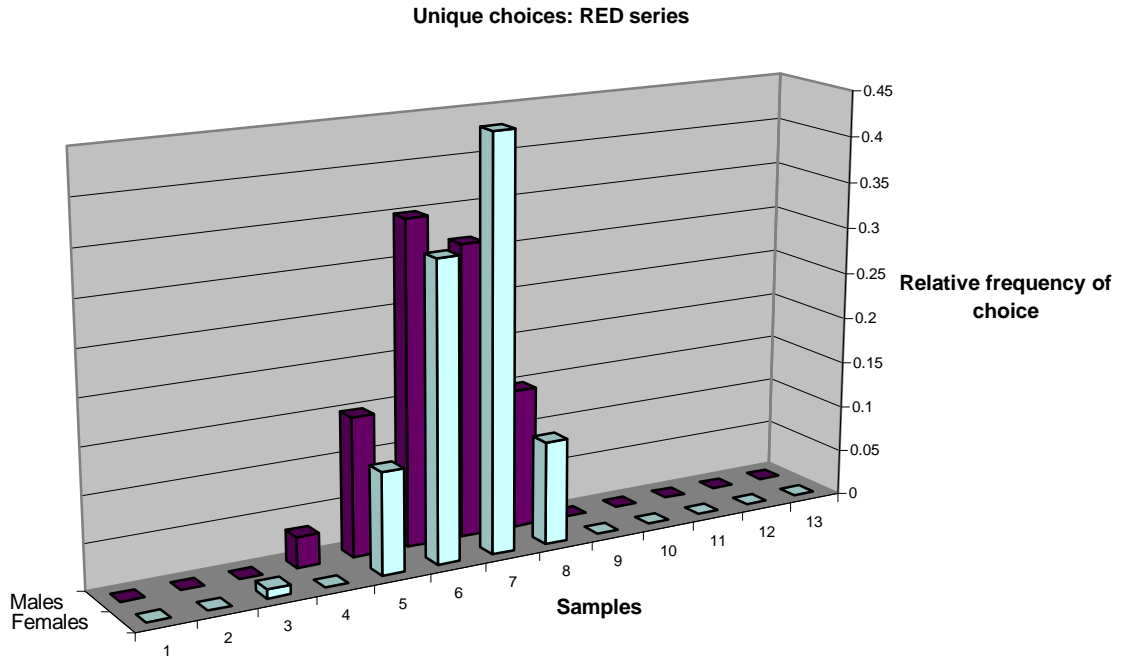
B
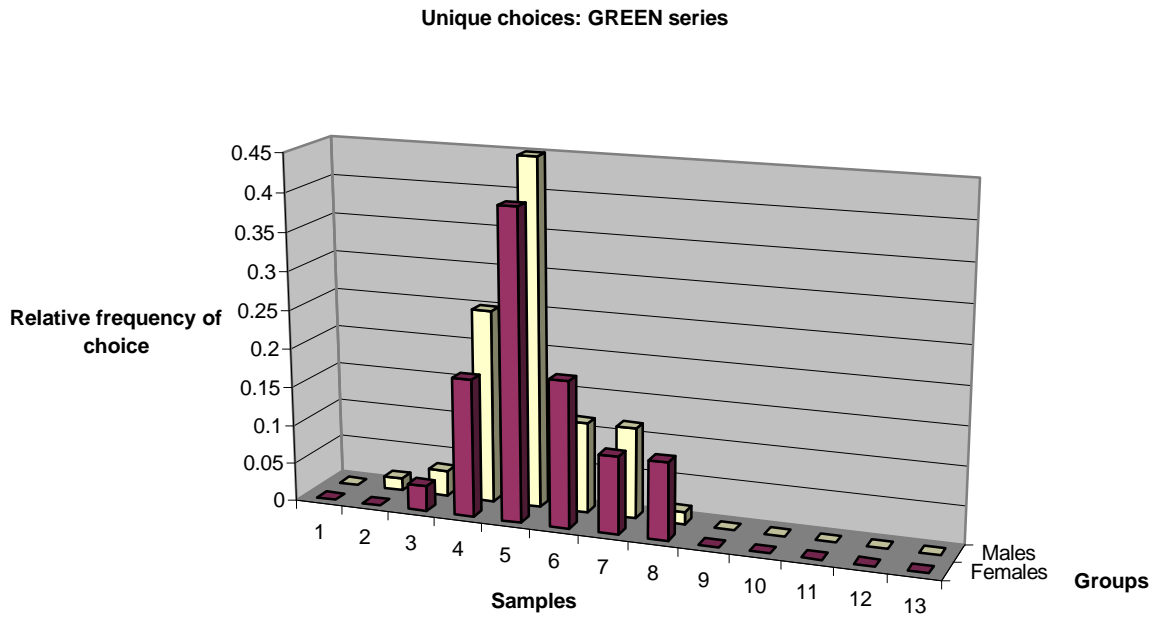
C
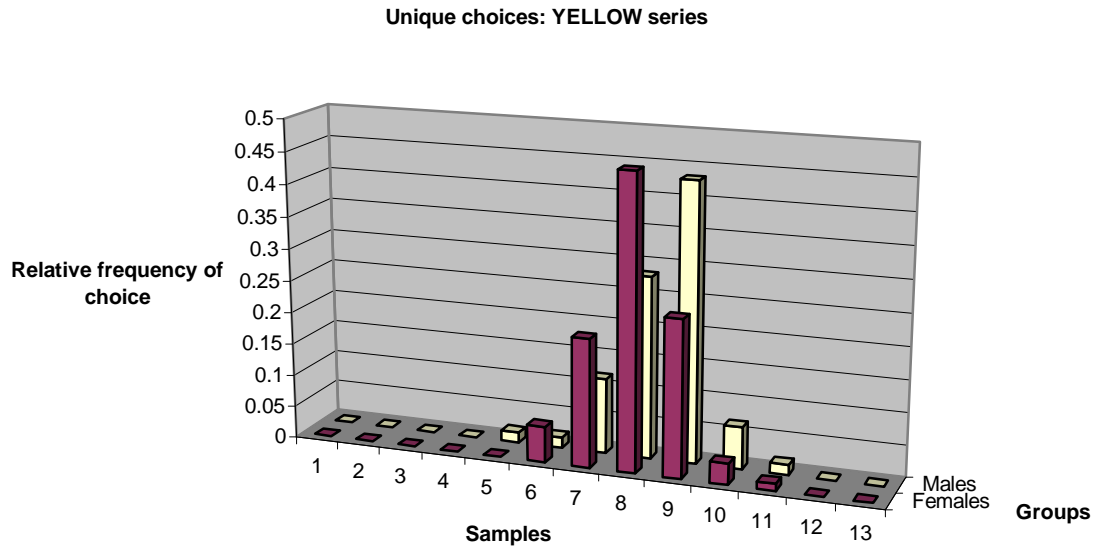
D
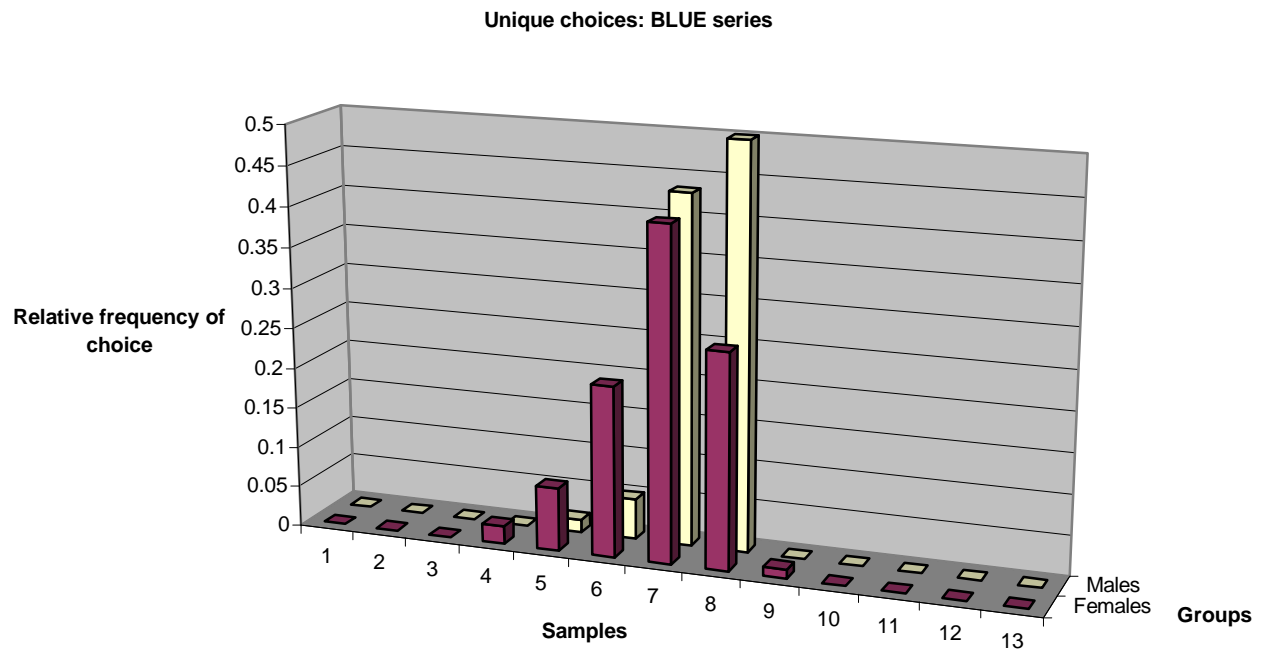
Kuehni (2001) reports slight discrepancies between the sexes in unique hue choices from series of Munsell chips. He examined 40 subjects (22 females, 18 males), but he did not determine whether the differences he found were statistically significant. In the present study, part of the reason for the absence of significant group differences is the small number of subjects examined, resulting in low statistical power for some comparisons.

Despite the absence of group differences, individual differences within the entire group of subjects are substantial, and so are in accordance with what we can expect on the basis of previous studies. Table 9 shows the unique hue choice distributions for the four color series of those subjects who are most different in this respect, and whose data will be further analyzed below.

		Samples												
		1	2	3	4	5	6	7	8	9	10	11	12	13
RED Series	M3							3	7					
	M4				1	6	2	1						
GREEN series	F6			1	5	4								
	F9						1	3	6					
	M1			1	4	5								
	M5				4	4	1	1						
	M6					2	3	4	1					
YELLOW series	F3						4	2	3	1				
	F9							1	1	7	1			
	M1							1	2	5	1	1		
	M4					1		7	1	1				
	M6						1	1	4	4				
BLUE series	F3							4	5	1				
	F5							3	7					
	F9				2	4	4							
	M1							9	1					
	M3							2	8					
	M5							1	9					

Table 9. Frequency distributions of unique hue choices made by subjects who are most characteristically different from each other. The data of these subjects, including their naming responses is examined in more detail below. Empty cells in the table correspond to zeros; numbers in the cells show how many times a given color sample was chosen by a given subject. The sum of each row in the table is 10, as each subject made ten unique hue choices in each color series.

I did not observe any interesting relation between color rule settings and unique hue choices (or namings) worthy of being analyzed further.

Combining unique hue choice and naming data. In order to compare the results to the hypothetical cases described in Section 3.5 above it is necessary to combine unique hue choice and naming data. Such a combination should give us a clearer picture of how the 52 color samples were perceived by the individuals who participated in the experiment. Let me start by some theoretical considerations to motivate my method of combining the two kinds of data.

Since both the unique hue choices and the namings have proven probabilistic, that is, there was an obvious within-subject variation in both tasks, in combining the two kinds of data, some arbitrary decision thresholds are needed in deciding how a particular sample was perceived by a particular subject. For instance, subject F9 named sample BLUE 9 blue nine times and blue-red once. The question arises, what can we say about her perceptions of BLUE 9 in general? She perceived this sample ten times and in essentially the same circumstances in the experiment. Should the fact that she named it blue-red once and blue nine times prevent us from any generalized conclusion about her perception of BLUE 9's color? Of course not. We could, for instance, argue that one blue-red naming against nine blue namings might be due to some irrelevant random event in the subject's color vision system hence, from the data we have at hand, the best idea is to conclude that she perceived Q as blue. However, if she named the sample bluish green four times and blue six times, then we might argue that in her color vision system, there may have been some systematic trace of blue-response on looking at the sample, still, this

blue-response was relatively weak, and so it did not affect her response in all cases. A faint trace of green on a predominantly blue surface is sometimes overlooked. Finally, if the subject named the sample bluish green nine times and blue only once, then it is the unique blue response that can perhaps be interpreted as missing, by accident, a stronger cast of green.

Unique hues correspond to narrow ranges of color space whereas binary hues correspond to wider ranges. In terms of opponent processes, unique red is represented by a positive value in the red-green channel and zero (baseline activity) in the blue-yellow channel. As I have just outlined, we could endorse a strict criterion for unique hue perceptions by saying that if, for instance, a sample is perceived blue in 75 per cent of the cases and bluish green in 25 per cent of the cases, then there already is a small negative response in the red-green channel that tends to affect the subject's response. Therefore the conclusion could be that the subject's perception of the sample was (slightly) greenish blue, not unique blue.

However, we could also endorse a relaxed criterion for unique hue perceptions. We could, for instance, say that if a subject named a sample greenish blue in less than 50 per cent of the cases and blue in more than 50 per cent of the cases, then her perception of the sample was blue. (If a subject named a sample greenish blue and blue both in exactly 50 per cent of the cases, then we can still resort to the above principle that there is a tendency in the subject's color vision to give a negative response in the red-green channel to the sample, and classify her perception of the sample as greenish blue.)

Notice that there is an asymmetry in this system: it does not make as much sense to use a strict criterion for binary hues as it does for unique ones. If a subject named a

sample greenish blue in 75 per cent of the cases and blue in 25 per cent of the cases, then it is reasonable to conclude that her color opponent system responded to the sample by a small negative value in the red-green channel that affected her verbal response in three quarter of the cases; *on average, her response along the red-green dimension was on the green side*.[99] That is, in this case, it is less reasonable to infer that her perception of the sample was unique blue simply because in 25 per cent of the cases she named the sample blue. There can be two explanations of such a pattern of responses: (1) the activity in the red-green channel was present in all cases, but it was too small to affect the verbal response in every case (random factors affect the process of verbal encoding). (2) The activity in the red-green channel itself exhibited some random oscillation, so negative value occurred in only 75 per cent of the cases. If the responses of the color-opponent system indeed show some random oscillation, then a true zero response to a color sample in the red-green channel would be reflected by a few greenish responses and also by a few reddish responses, indicating that the responses of the red-green channel *oscillated around zero*. Such cases do indeed occur in my data set. Here is an actual example: subject M6 named sample YELLOW 8 yellow-red once, yellow-green twice, yellow three times, and chose it as unique yellow four times.

Given these theoretical considerations, here is the method that I used to combine unique hue choice and naming data. I used two criteria: a strict one and a relaxed one for unique hue perceptions. As the first step of the procedure, for each subject and each color sample I added up the number of unique choices and that of unique namings. I assumed that a unique choice is associated with a unique naming: if, on a particular occasion, a subject chose a sample as the purest green, then this response implies that she had a

tendency to name it green (and not yellowish green or bluish green). After this, the sum of unique namings and unique choices was compared to other namings. The perception by a subject of a sample was established in two different ways using the two criteria.

Strict criterion for unique perceptions: If the minority binary response occurred with a relative frequency greater than or equal to 0.25 (25 per cent), then the corresponding binary perception was assigned, otherwise (i.e., less than 25 per cent binary response) the subject's perception of the sample was classified unique. For example, subject F9 named sample GREEN 6 green eight times, chose it as unique green once, and named it yellowish green once. That is, there were 8+1=9 green responses, and one (in ten) – ten per cent – yellowish green response. The assigned perceptual value on the strict criterion was (unique) green. Subject M5 named the same sample (GREEN 6) green six times, chose it as unique green once, and named it bluish green three times. On the strict criterion for unique hues, this response pattern already counted as indicating a bluish green perception. Obviously, the 25% limit is arbitrary, but we need to introduce *some* arbitrary probability threshold.[100]

Relaxed criterion for unique perceptions: the majority response wins after adding up unique choices and unique namings. Example: on this criterion, subject M5's response pattern to GREEN 6 is classified green. Six green namings plus a unique green choice add up to seven. There are three bluish green responses: the total is 7+3=10, and the majority (green responses) wins.

Supplementary principle: contradictory namings cancel each other. Example: subject M4 chose RED 5 as unique red six times, named it red once, yellow-red once, and blue-red twice. In this and similar cases I proceeded thus: I disregarded one yellow-red response

and one blue-red response, leaving a total of eight responses. Out of these, 6+1=7 were red responses, and one was a blue-red response. On both criteria, this counted as a red response (as 1/8<0.25).

Special case for the relaxed criterion: for example, subject F9 chose YELLOW 10 as unique yellow once, named it yellow three times, and yellowish green four times. That is, the yellow responses: 3+1=4 equal the yellowish green responses. In such cases and under the relaxed criterion I assigned binary perception: I decided that F9's perception of YELLOW 10 was yellowish green, not yellow on the relaxed criterion. (I.e., in such borderline cases the relaxed criterion too was slightly biased toward binary hue assignments.) The strict criterion dictates the assignment of binary perception in such cases.

Patterns of individual difference. For each of the four color series, I selected pairs of subjects whose unique hue choices were most different. Table 9 above shows the unique hue choice distributions of the subjects selected. I used the mode of unique choices in selecting such pairs (i.e., the mode here is *the sample*, in a color series, that was chosen as unique most frequently, by a particular subject). All pairs examined were same-sex pairs. I will discuss one pair for the red series, three pairs for the green series, two pairs for the yellows, and four pairs for the blues. Table 10 shows these pairs and their responses to the corresponding stimuli.

Table 10. Different patterns of individual difference obtained with the four color series. M1-M6 and F1-F9 refer to subjects. Symbols under the sample numbers indicate color perceptions of the samples by the subjects. 'R': red, 'G': green, 'Y': yellow, 'B': blue, 'YR': yellow-red, 'YG': yellow-green, 'BR': blue-red, 'BG': blue-green. Dashed lines between the arrowheads indicate the interesting part of the tables. 'XX' in the dashed line marks incompatible color perceptions in two-step shifts. See the main text for further explanation.

_____

RED series

```
Strict criterion:
Sample#   1  2  3  4  5  6  7  8  9 10 11 12 13
M3       BR BR BR BR BR  R  R  R YR YR YR YR YR
M4       BR BR BR BR  R  R YR YR YR YR YR YR  Y
                  >-----------------<
Relaxed criterion:
Sample#   1  2  3  4  5  6  7  8  9 10 11 12 13
M3       BR BR BR BR BR  R  R  R YR YR YR YR YR
M4       BR BR BR  R  R  R YR YR YR YR YR YR  Y
               >--------------------<
```
_____

GREEN series:

<u>PAIR 1:</u>

```
Strict criterion:
Sample#   1  2  3  4  5  6  7  8  9 10 11 12 13
F6       YG YG  G  G  G  G  G BG BG BG BG BG BG
F9       YG YG YG YG YG  G  G  G BG BG BG BG BG
               >------------------------<
Relaxed criterion:
Sample#   1  2  3  4  5  6  7  8  9 10 11 12 13
F6       YG YG  G  G  G  G  G  G BG BG BG BG BG
F9       YG YG YG YG  G  G  G  G  G BG BG BG BG
               >------------------------<
```

PAIR 2:

Strict criterion:
```
Sample#    1  2  3  4  5  6  7  8  9 10 11 12 13
M5        YG YG YG  G  G BG BG BG BG BG BG BG BG
M6        YG YG YG YG YG YG  G BG BG BG BG BG  B
             >---------XX------<
```
Relaxed criterion:
```
Sample#    1  2  3  4  5  6  7  8  9 10 11 12 13
M5        YG YG YG  G  G  G BG BG BG BG BG BG BG
M6        YG YG YG YG YG  G  G BG BG BG BG  B  B
             >-----------------<
```

PAIR 3:

Strict criterion:
```
Sample#    1  2  3  4  5  6  7  8  9 10 11 12 13
M1        YG YG YG  G BG BG BG BG BG BG BG BG BG
M6        YG YG YG YG YG YG  G BG BG BG BG BG  B
             >------XX-XX------<
```
Relaxed criterion:
```
Sample#    1  2  3  4  5  6  7  8  9 10 11 12 13
M1        YG YG YG  G  G BG BG BG BG BG BG BG BG
M6        YG YG YG YG YG  G  G BG BG BG BG  B  B
             >-----------------<
```
_____
YELLOW series:
PAIR 1:

Strict criterion:
```
Sample#    1  2  3  4  5  6  7  8  9 10 11 12 13
F3        YR YR YR YR YR YR  Y YG YG YG YG YG  G
F9        YR YR YR YR YR YR YR  Y  Y YG YG YG YG
                      >--------------<
```
Relaxed criterion:
```
Sample#    1  2  3  4  5  6  7  8  9 10 11 12 13
F3        YR YR YR YR YR  Y  Y YG YG YG YG YG  G
F9        YR YR YR YR YR YR  Y  Y  Y YG YG YG YG
                     >-----------------<
```

PAIR 2:

```
Strict criterion:
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
M1        YR  YR  YR  YR  YR  YR  YR  YR   Y  YG  YG  YG  YG
M4        YR  YR  YR  YR  YR  YR   Y  YG  YG  YG  YG   G   G
                            >------XX------<
Relaxed criterion:
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
M1        YR  YR  YR  YR  YR  YR  YR  YR   Y  YG  YG  YG  YG
M4        YR  YR  YR  YR  YR  YR   Y  YG  YG  YG  YG   G   G
                            >------XX------<
```
_____

BLUE series:

PAIR 1:

```
Strict criterion:
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
F5        BG  BG  BG  BG   B   B   B   B   B  BR  BR  BR  BR
F9        BG  BG  BG  BG   B   B   B   B   B  BR  BR  BR  BR
                        >--------------------<
Relaxed criterion:
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
F5        BG  BG  BG  BG   B   B   B   B   B  BR  BR  BR  BR
F9        BG  BG  BG   B   B   B   B   B   B  BR  BR  BR  BR
                    >------------------------<
```

PAIR 2:

```
Strict criterion:
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
F3        BG  BG  BG  BG   B   B   B   B  BR  BR  BR  BR  BR
F9        BG  BG  BG  BG   B   B   B   B   B  BR  BR  BR  BR
                        >--------------------<
Relaxed criterion:
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
F3        BG  BG  BG  BG   B   B   B   B  BR  BR  BR  BR  BR
F9        BG  BG  BG   B   B   B   B   B   B  BR  BR  BR  BR
                    >------------------------<
```

PAIR 3:

Strict criterion:
```
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
M1        BG  BG  BG  BG  BG   B   B   B  BR  BR  BR  BR  BR
M5        BG  BG  BG  BG  BG   B   B   B  BR  BR  BR  BR  BR
                    >--------------<
```
Relaxed criterion:
```
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
M1        BG  BG  BG  BG  BG   B   B   B  BR  BR  BR  BR  BR
M5        BG  BG  BG  BG   B   B   B   B  BR  BR  BR  BR  BR
                    >----------------<
```

PAIR 4:

Strict criterion:
```
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
M1        BG  BG  BG  BG  BG   B   B   B  BR  BR  BR  BR  BR
M3        BG  BG  BG  BG   B   B   B   B   B  BR  BR  BR  BR
                    >--------------------<
```
Relaxed criterion:
```
Sample#    1   2   3   4   5   6   7   8   9  10  11  12  13
M1        BG  BG  BG  BG  BG   B   B   B  BR  BR  BR  BR  BR
M3        BG  BG  BG  BG   B   B   B   B   B  BR  BR  BR  BR
                    >--------------------<
```

In Table 10, the dashed lines between the arrowheads indicate the portions of the data that I want to focus on: those parts that we can compare to Tables 4 to 7 in Section 3.5. It should be obvious that the strict criterion narrows the ranges that correspond to unique hue perceptions of the subjects. It is worth emphasizing again that I used two different criteria because it may not be very plausible to claim that one and only one of such criteria is the correct one. Based on the behavioral responses I used in this experiment, it is not possible, or at least not easy, to establish with absolute certainty where the boundaries of unique hue categories for different subjects are. Still, I think it is a good result that color-perception assignments do not differ dramatically under the two criteria.

In the table we can see the following results of key importance. First, we can find one-step shifts with the red, green and yellow series. For all the five pairs presented, and

for both criteria, such shifts are obtained consistently. Note that in most of these cases there is more than one way to reconstruct Table 5 from the actual data. For instance, consider the yellow series, Pair 1. Under the strict criterion, column 6 (the column of YELLOW 6) in Table 10 corresponds to the first column of Table 5 (SR_1); column 7 in Table 10 corresponds to the second column of Table 5; columns 8 and 9 together correspond to the third column of Table 5; finally, column 10 corresponds to the fourth column of Table 5. Under the relaxed criterion, columns 5 and 6 correspond to columns 1 and 2 of Table 5 respectively, and column 7, evaluated by the relaxed criterion, suggests an overlap between the two subjects' unique yellow categories that is not represented in Table 5.

As we can now see clearly, one-step shift is compatible, and indeed often co-occurs, with a substantial overlap in the two subjects' unique hue categories. For another example, look at Pair 1 in the green series data. Under both criteria, GREEN 3 and GREEN 4 were perceived as green by subject F6 and as yellowish green by subject F9. Under the strict criterion, GREEN 8 was perceived bluish green by F6 and green by F9; under the relaxed criterion, GREEN 9 was perceived bluish green by F6 and green by F9. This indicates a shift in the two subjects' unique green categories. However, the overlap between their unique green categories is also fairly wide: it spans samples GREEN 6 and GREEN 7 under the strict criterion, and GREEN 5-8 under the relaxed criterion.

Two-step shifts were also observed in two color series out of four: green and yellow. Two such cases were found in the green series, and one in the yellow series. One two-step shift in the green series was observed between subjects M5 and M6 (Table 10, green series, Pair 2). It is dependent on the criterion for unique hue perception, as it

occurs only under the strict criterion. The critical sample is GREEN 6: subject M5 chose it as unique green once, named it green six times and bluish green three times. Subject M6 chose this sample as unique green three times, named it green four times, and yellowish green three times. In both these cases, the binary categories bluish green and yellowish green occurred in three cases out of ten – just enough to affect the color perception assignment on the strict criterion. Still, if we think of the theoretical considerations outlined above, it is arguable that there were subtle opposite tendencies in the color perception of these two subjects. M5 never named GREEN 6 yellowish green, nor did M6 name it bluish green, so perhaps there was indeed a slight bias in M5 toward bluish green perceptions, and a comparable slight bias in M6 toward yellowish green perceptions.

Another two-step-shift in the green series obtained between subjects M1 and M6 (Table 10, green series, Pair 3). Notice two things. First, this phenomenon is also dependent on the evaluation – it occurs only under the strict criterion for unique perception. Second, there are two adjacent samples that seem to have been perceived in incompatible ways by the two subjects: GREEN 5 and GREEN 6. A closer look at the data reveals that subject M1 chose GREEN 5 as unique green five times (!), named it green once, and bluish green four times. He named GREEN 6 green three times and blue-green seven times. Subject M6 chose GREEN 5 as unique green twice, named it green three times, and yellow-green five times. He chose GREEN 6 as unique green three times, named it green four times, and yellowish green three times. The relaxed criterion indicates a sharp one-step shift in this case, apparently without an overlap between the unique green categories of these two subjects.

The two-step shift obtained with the yellow series is somewhat different (Table 10, yellow series, Pair 2). Here the critical sample was YELLOW 8. Subject M1 chose this sample as unique yellow two times, named it yellow once, yellow-green once, and yellow-red six times. (Calculation: the yellow-green response and one yellow-red response cancel each other, so the total number of responses drops to 8. Five of these are yellow-red namings, and three are yellow responses, so the majority response is yellow-green.) Subject M4 chose YELLOW 8 as unique yellow once, named it yellow-red once, and yellow-green eight times. On both criteria this pattern indicates a tendency of the subject toward yellowish green perceptions. An interesting phenomenon is that in the green two-step shift cases, even though the tendency showed up only under the strict criterion, there were no contradictory namings, whereas in the yellow case, which came through under both criteria, there were contradictory namings in both subjects' responses. This might be a phenomenon that deserves further attention.

Finally, we can find at least one case of one-plus-two-step shift in the dataset. If we look at sample YELLOW 8 again, and subjects M1, M4, and M6, we have the following table:

Table 11. One-plus-two-step shift. As in Table 10, M1, M4 and M6 refer to subjects. Symbols under the sample numbers indicate color perceptions of the samples by the subjects. 'R': red, 'G': green, 'Y': yellow, 'B': blue, 'YR': yellow-red, 'YG': yellow-green, 'BR': blue-red, 'BG': blue-green. Dashed lines between the arrowheads indicate the interesting part of the tables. 'XX' in the dashed line marks the point where color perceptions of the three subjects are most divergent. See the main text for further explanation.

```
YELLOW series:
Strict criterion:
Sample#    1  2  3  4  5  6  7  8   9 10 11 12 13
M1        YR YR YR YR YR YR YR YR   Y YG YG YG YG
M4        YR YR YR YR YR YR  Y YG YG YG YG  G  G
M6        YR YR YR YR YR YR YR  Y YG YG YG YG  G
                          >------XX------<
Relaxed criterion:
Sample#    1  2  3  4  5  6  7  8   9 10 11 12 13
M1        YR YR YR YR YR YR YR YR   Y YG YG YG YG
M4        YR YR YR YR YR YR  Y YG YG YG YG  G  G
M6        YR YR YR YR YR YR YR  Y YG YG YG YG  G
                          >------XX------<
```

As we can see from Table 11, the phenomenon is strong in the sense that there is full agreement in the color perception assignments between the strict and relaxed criteria. The marked part of the tables matches Table 7 – the hypothetical case – exactly. Subject M6 chose sample YELLOW 8 as unique yellow four times, named it yellow three times, yellow-green two times, and yellow-red once.

3.8. Discussion

In conclusion, I suggest that the phenomenon of one-step shift was demonstrated for three unique colors out of four (red, green and yellow). In addition, there is some indication that the phenomena of two-step shift and one-plus-two step shift might exist as well, though to confirm this conjecture we need further empirical support. I consider the present experiment to be a pilot study that has to be followed up by further studies. A critical feature of such experiments is the selection of stimuli: too big a gap between neighboring samples would miss individual differences (since all subjects would exhibit

the same response pattern) whereas too small differences would amplify within-subject variation. The key question in such experiments is whether between-subject differences in color perception override within-subject variations, and to observe this phenomenon we need carefully selected stimuli. There are also alternative experimental methods that could be used. One such method (for unique hue choices) is to display only one stimulus at a time, and have the subject adjust its color. This would eliminate color contrast effects that may have obtained between adjacent samples in the present experiment.

The next question is, what is the reason for finding no effect with the blue colors? Part of the reason may be that the perceptual differences between the blue samples were too small on the greenish blue side (BLUE 1 to BLUE 7). Subjects often commented after the experiment that there were more than one samples in the blue series that looked equally good as unique blues, whereas with the other three series the usual complaint was that no sample looked really unique, but each of them had a tinge of some other color component in it. (A typical case was when a subject saw all samples from, say, YELLOW 1 to YELLOW 7 as reddish to some extent, and all the rest as greenish to some extent. About three or four of the 15 subjects made this complaint.) For such cases the subjects were instructed to choose the sample that is the closest to unique. Kuehni (2001, p. 62) used a different approach: he allowed the subjects to identify, as the location of a unique color, the halfway point between two samples.

Some other methodological differences between Kuehni's experiment and mine are also of interest. His aim was also to determine the unique hues, and look at the individual differences in this respect. However, he used ordinary reflecting surfaces (Munsell chips), not emissive stimuli as I did. Perhaps the most important difference

between the two designs is that Kuehni had his subjects choose the four unique hues only once. He did not ask for naming of the remaining samples. This reduced the completion time of the total task to 2-4 min (Kuehni, 2001, p. 62). I think that allowing unique choice in between two samples is a good idea. On the other hand, despite the ambiguity it creates, asking for repeated unique choices and namings to assess within-subject variation is a good idea, too. Every sort of sensation has in it a stochastic component, and around the limits (absolute thresholds, difference thresholds, and so on) this stochastic character becomes crucial. Therefore, asking for repeated responses to the same stimuli is also crucial in psychophysical experiments, whereas asking for just one response to each stimulus might create the misleading impression that the examined phenomenon is deterministic, that is, the subject would always give the same response to the same stimulus in the same circumstance.

Next problem. Against the individual differences I presented, the following objection could be raised. What is going on here is that I look at pairs of subjects in my sample to point out individual differences in color perception that are of theoretical interest. Now, someone could object that I am looking at the outliers of my sample, and find interesting differences only between them. But then, the actual individual differences in phenomenology that allegedly support my argument are the exception rather than the rule. So I am using the exceptional cases in trying to refute phenomenal externalism, instead of drawing upon the majority of the cases and conclude that, at least in this particular experiment, I failed to find any interesting between-subject differences that could support my anti-representationalist conclusion.

This objection needs a reply. I think there is reason to take individual differences seriously in this case. First of all, even though only about 5.7 per cent of all the *pairs* exhibit one-step shift or two-step shift[101], we can see from Table 9 above that five of the six male subjects, and four of the nine female subjects participate in at least one of the six interesting pairs. Viewed this way, the phenomenon of individual differences is more than a marginal one. Second, unlike the case of shape perception, there is no normative ruling against outliers in color vision. A subject whose best unique yellow is sample YELLOW 7 is just as normal a color perceiver as another subject whose best unique yellow is YELLOW 9, even if one of them perceives YELLOW 8 as greenish yellow, and the other, reddish yellow. If the task is to choose unique green from a series of greenish colors, and most color-normals choose stimulus X, whereas a few of them consistently choose stimulus X+3 in the series, then the question arises, is there any reason to believe that the outliers misrepresent the greenish colors? The only reason for holding that the outliers misrepresent the colors is because they diverge significantly from the group average. They see unique green somewhere else in the series than do most other color-normals. Compare this with an analogous experiment in shape perception. One could give to experimental subjects the following task. Given a series of more or less circular objects, as in Figure 5, subjects have to choose the one that is the best circle.
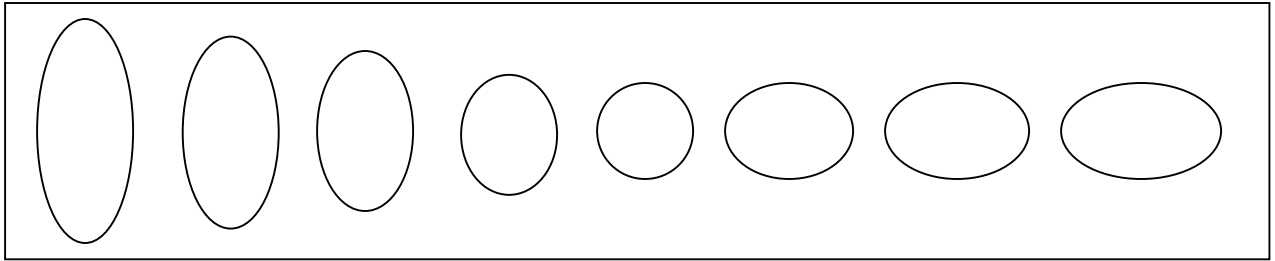
Figure 5. A shape perception task that is analogous with the unique hue choice task presented in this section.

It is quite likely that most subjects, women and men alike, would choose the best circle in such a case. (Differences between adjacent shapes can be minimized making the task more difficult.) In Figure 5, the best circle is the fifth from left. If a few subjects would consistently choose, say, the fourth shape from left, then there would be two quite different reasons to hold that these subjects misperceive the figures. The first (and irrelevant) reason is that they diverge from the majority. The second (relevant) reason is that they consistently fail to correctly assess the relation between the horizontal and vertical dimensions of the shapes. As a matter of fact, the fourth figure from left is an oval, not a circle: its vertical diameter is longer than its horizontal one. Here, in the case of shape perception, there is an objective criterion for misperception, in addition to conformity to the group majority. As a consequence, outliers here are reasonably regarded as misrepresenting the figures. Not so in the color case where the norm for trichromacy (i.e., passing the standard color vision tests) does not include the exact perceptual location of unique hues – nor are there any perceiver-independent stimulus features that make certain object colors "objectively" (i.e., perceiver-independently) unique.

Tye would probably resist this conclusion and claim that he managed to "objectify" the unique-binary distinction (Tye, 2000, pp. 162-165). If his schema for defining object colors in terms of surface reflectance were correct, then a perceiver-independent distinction between unique and binary hues would indeed fall out of it. As Tye suggests, unique green is the reflectance for which $M^* >_s L^*$ and $S^* =_a L^*+M^*$ obtains. Similar criteria are provided for the other three unique hues: see Section 2.2.2. above. Given such an objective distinction between unique and binary hues, Tye can go on to claim that normal trichromat subjects should perceive exactly those colors as unique green that are objectively unique green. Outliers who locate unique green somewhere else misperceive the green colors because they fail to correctly identify the point of balance between S^* and L^*+M^* – in exact analogy with the case of ovals and circles.[102]

Alas, Tye's schema is hopelessly mistaken, and therefore his proposal for the distinction between unique and binary hues is also entirely off the mark. Thus he has no basis to claim that outliers in the unique hue choice task misperceive the colors any more than the majority does. Note also that my crude specification of broad color categories in terms of reflectance in Section 2.2.5 does not suggest any perceiver-independent division corresponding to the perceptual division of unique and binary hues.

My conclusion in this section is that outliers in the present color vision experiment do not in any sense misperceive the colors; this is a disanalogy with relevant cases of shape perception. My outliers are color-normals (as suggested by the color vision tests), and so they can be compared to the majority – that is, to other color normals. Different color normal individuals locate the unique hues at different points of the

stimulus space without committing any perceptual mistake. The only upshot remains that color phenomenology varies independently of the relevant (content-bestowing) stimulus properties.

4. Concluding remarks: arguments against representational externalism

As I noted in the previous sections, there are different arguments against representational externalism about color experience. I made two such arguments: one from the failure of type physicalism about color, and one from individual differences in phenomenology. I think it is possible to set up other arguments as well. I did not attempt to argue from unity and the unique-binary distinction (see section 1.4 above); nor did I use the phenomena of color contrast to argue against type physicalism and representational externalism. Indeed, the terms 'color contrast' and 'simultaneous contrast' hardly occur in the present work, even though they are of central importance to color perception (see for instance Fairchild, 1998; Lotto and Purves, 2000). The relevance of simultaneous contrast phenomena in the present context is this: one and the same colored surface can appear, to the same observer and in the same illumination, radically different in color in different color-contexts (surrounds). None of the relevant color-contexts are reasonably regarded as 'abnormal', since they might occur in our everyday environment. This observation can give rise to arguments according to which there is virtually no correlation between perceiver-independently characterized stimulus properties and phenomenal color experiences. The absence of such correlations can then be used to support either subjectivist views about color or at least internalist views of color experience. As I said above, I think subjectivism about color can be avoided by

some neat philosophical moves like relativizing the colors and endorsing a disjunctive physicalist view of object color. However, as I argued, internalist views of color experience cannot be avoided. That is, what it is like to see colors cannot be fully understood in terms of what properties of the environment are represented by color vision.

Notes

Notes for Chapter One

---

[1] That is, no relation to perception figures as an essential element in the property of *having/being a particular shape S*.

[2] A quick analogy here is with firing a shotgun. No matter what evokes the effect (e.g., pulling the trigger, or heating the barrel), the effect is the same – the firing. More details will come in later sections.

[3] Pigeons are a key example in Thompson et al's paper.

[4] This observation does not go uncontested – this is one of the key issues discussed in the present dissertation.

[5] Not everyone would accept the idea that dispositions are not causally effective. I will discuss this point, and give my reasons for holding that dispositions are causally inert in 2.3.2 and 2.5.2.4 below.

[6] I.e., if there are non-dispositional properties at all, then having a temperature of, say, 400 °C is certain level of kinetic energy of the molecules – a non-dispositional property that is "intrinsic" to the macrolevel objects that have such a temperature.

[7] In Tye's view, 'inherent property' means *local feature* – a feature that does not involve anything away from the object or surface of which it is a property (Tye, 2000, pp. 147, 153). Reflectance in Tye's view counts as an inherent property, but the ratio of the reflectances of a target surface and the surrounding, background surface does not (see his 2000, pp. 152-153). Obviously we cannot identify inherent properties with intrinsic (i.e., nonrelational) properties, since reflectance, an allegedly inherent property, is undeniably a relational property – it is a disposition. Still, as I understand him, Tye strives for retaining at least some weakened notion akin to intrinsic properties that applies to colors, in order to defend the type physicalist view that colors are natural kind essences by means of which surfaces with the same color belong to the same natural kind. Natural kind essences are typically intrinsic properties, or at least the paradigm examples like being H_2O or having the atomic weight 79 are.

[8] Newton's prism and a few other examples (like rainbows) might strike one as a counterexample to the claim that light rays themselves are invisible. However, in these cases light is reflected or scattered by the objects which it passes through (the glass volume of the prism, and the volume constituted by the myriad water droplets in a cloud) and our color perception is due to these phenomena. It is the prism (or a certain part of the clouds) that appears colored in such cases, somewhat similarly to the case of an illuminated movie screen. It is the normally white screen that looks differently colored, not the lights coming from the projector that illuminate it.

[9] Tye, just like Hilbert, recognizes that object colors are derivative, "uninteresting", anthropocentric properties, even if they are physical types of some sort. However, it is evident that despite this insight, Tye strongly holds onto a natural kind view of object color (Tye, 2000, Ch. 7; see also pp. 124-125). Thus the notion of anthropocentric natural kinds is not my invention. How plausible this dilution of the notion of natural kinds is on other counts, I am not sure. On this liberal approach, quite obviously, shapes are natural kinds as well – including shapes that we identify perceptually.

[10] Switching between subjects will not evade this conclusion, Campbell argues. For what we want to make sense of is the possibility that the particulars and their color properties in our environment might have existed in the absence of any human or other perceiver. In making sense of this possibility we still cannot appeal to identification of particulars from no point of view, because there is no such identification (Campbell, 1993, p. 260).

[11] I am not entirely convinced. Campbell here makes two related claims that are nevertheless distinct. The first is that relation to some subject necessarily figures in identifying particulars (Campbell, 1993, p. 259). The second is that we identify particulars via their spatio-temporal location, and that is a contingent feature (p. 261). The second claim can be true without the first, and I think it is also more plausible than the first. For instance, in identifying heavenly bodies, we typically make reference to their spatial location, and this would become necessary if molecule-by-molecule duplication of them occurred. Still, no relation to any particular perceiver, or some ideally situated perceiver, figures in the identification of stars and planets (as opposed to spoons, for instance). Relation to the Solar System is not possible to eliminate completely from the localization of stars and planets, simply because the Solar System is part of the Universe, and even if

localizations in it were explicitly specified relative to Alpha Centauri, the system would still contain information about the location of the Solar System relative to any other astronomical object.

[12] Not just any three wavelengths in the 400-700 nm range would work. The three primary wavelengths have to be visually independent: no additive mixture of two of them should be a visual match of the third. Moreover, even with three visually independent primaries, there will always be some test light SPDs that cannot be matched by an additive mixture of those primaries. To match these test lights, it is necessary to move one or even two of the primaries to the test side of the circular bipartite field. Mathematically this means that though most visual matches are of the form: $\mathbf{t} = e_1\mathbf{p}_1 + e_2\mathbf{p}_2 + e_3\mathbf{p}_3$, some matches are of the form: $\mathbf{t} = -e_1\mathbf{p}_1 + e_2\mathbf{p}_2 + e_3\mathbf{p}_3$. See Wandell, 1995, p. 84.

[13] More exactly, $\mathbf{y}(\lambda)$ is a rough approximation to the brightness of monochromatic lights of equal size and duration (Wandell, 1995, p. 87).

[14] I.e., such intrinsic properties of mental (brain) states are necessary but not sufficient to explain representational capacity.

[15] This might strike some people as a controversial claim. There are theorists (the so-called epiphenomenalists) who argue that the discrimination, recognition, etc. happens on the basis of ordinary (biological, computational) properties of perceptual (brain) states, but not via their corresponding phenomenal characters. For even though phenomenal character arises from brain states, the two are not identical, and only brain states are causally effective, phenomenal characters are not. I am not particularly attracted by this view, so I don't mind tacitly assuming that whatever they are, phenomenal characters are causally efficacious.

[16] "Lawfully" refers to some high-level psychologival laws here – psychological links that we do not frequently call a *law* in ordinary discourse, because they are often too particular. For instance, certain visual perceptual states are perceptions of chipmunks because they reliably covary with the occurrence of chipmunks in one's visual field. This is a result of perceptual learning – the formation of perceptual categories that correspond to chipmunks. In important respects, such links are like laws of nature: they are counterfactual-supporting, and we even have psychological explanatory stories of how such links obtain (e.g., pattern-recognition by neural networks in the brain, etc.).

[17] According to Neale (1990, pp. 124-129), a hyperintensional linguistic context is one which (1) does not allow the substitution of co-referential expressions within its scope, (2) is sensitive to more than the truth value and truth conditions of its operand, and (3) allows quantification into its scope. Psychological/propositional attitude terms like 'believes' are hyperintensional. Even though Mark Twain = Samuel Clemens, from the fact that Mary knows that M.T. wrote *Huckleberry Finn's Adventures* it does not follow that Mary knows S. C. wrote *Huckleberry Finn's Adventures* (1). Moreover, the following makes sense: Mary believes that **there exists exactly one person** who wrote *Huckleberry Finn's Adventures* (i.e., it was not two or more writers who wrote it in cooperation) (3). The feature (2) raises difficult questions and it is not easily captured by a quick example.

[18] However, see Section 2.5.2.4. below where I discuss Campbell's views further. What I say there is relevant to Tye's notion of transparency as well.

[19] As we have seen, Tye endorses a covariation theory of representation: representational content is determined by *covariation in optimal circumstances* (Tye, 1995, pp100-101). Misrepresentation arises when circumstances are non-optimal. An important question (although one I am not going to address here) is whether we can define optimal circumstances without making reference to selection history.

[20] In the case of human artifacts it would be relatively easy to cash out causal history in terms of the synchronic notion of, say, expectations of the designer, or user. E.g., the user of the car takes (expects, or interprets) the speedometer to carry information about speed, and this interpretation fixes the function of the speedometer. In the case of natural systems, such expectations can only be realized by the environment: the environment 'expects' living organisms to exhibit adaptive features, not maladaptive ones, or else they get selected out. However, synchronic 'expectations' of the current environment may or may not be in accordance with the selection history of organisms. A feature that has been selected against for a long time, and hence survived only in the form of an infrequent recessive allele can abruptly become adaptive as a result of a change in the environment. In other words, there is some sort of a lawful connection between designers' intentions and users' expectations (prospective users are typically told what to expect from a machine they consider buying), whereas selection history and current 'expectations' of the environment are contingently related.

[21] I am grateful to Dan Ryder for discussing Dretske's views with me, and letting me know Dretske's answer to my questions about his view. I am also grateful to Andrew Brook for discussions on this point. For the view of universals that Dretske is relying on, see Armstrong, 1989.

[22] By the term 'color objectivism' Thompson means physicalist views in a strong sense – views that I categorize as type physicalist. In his 1995 book Thompson analyzes in detail the reflectance theory of color, with a focus on D. Hilbert's (1987) account.

[23] Linearity is important and it is most often tacitly implied in isomorphism claims. If we allow for nonlinear transformations to occur in the mapping between physical stimulus dimensions and perceptual similarity spaces then again almost any stimulus similarity space becomes isomorphic with almost any perceptual similarity space. For instance, Thompson (1995, pp. 125-128) argues that the color-stimulus space consisting of three axes corresponding to values of integrated reflectance in the short, middle and long wavebands, as assumed by Hilbert (1987), is not isomorphic with perceptual color space with the dimensions of hue, saturation, and lightness. This claim is properly taken to imply that it is *linear* isomorphism that does not obtain between these spaces. For there surely is some complex nonlinear transformation that takes triplets of integrated reflectance, transformed by parameters characterizing perceptual context, illumination, observer sensitivity, and other relevant factors, into those of hue, saturation and lightness. Color science has not completely recovered this complex transformation yet, but details of it are already known (see for instance Fairchild, 1998). Note also that most sensory functions are nonlinear mappings between measurable stimulus properties and perceptual similariy spaces. Weber-Fechner type sensory functions are logarithmic transformations; Stevens type sensory functions are power functions (including, as a special case, linear transformation where the exponent is 1).

[24] Moreover, since no stimulus properties have these properties, no stimulus properties are the colors, or so the subjectivist argument continues (see Hardin, 1988; Matthen 1999).

Notes for Chapter Two

[25] For instance, as Block (1999) argues, the following cases are actual. Given two subjects, S1 and S2, and two color samples C1 and C2, C1 looks unique green to S1 but, say, slightly bluish green to S2, whereas C2 looks slightly yellowish green to S1 and unique green to S2. That is, the very same stimulus looks different in color to two trichromat perceivers in normal circumstances. Therefore, it is not true in full generality that, say, C1 is unique green (unique green is either a narrow shade or a narrow category of shades: unique greens differ from each other only in saturation and lightness, but not along the chromatic dimension of color space). The very same reflectance (or type of reflectances) can look different in color to different trichromat humans. This case, if carefully established, goes beyond the principle of perceiver-relativity (1.1.3). That is, unique green (a color category characterized by its perceptual look) is one narrow range of reflectances for me, and another, non-overlapping narrow range of reflectances for you, hence, unique green is at best a perceiver-relative shade (see Block, 1999, p. 63). I will give more details about this case in section 3.

[26] Even if we take the variability of metamer sets with individual perceivers into account this generalization is claimed to be true of the metamer sets of any particular perceiver. There exists some variation across normal trichromat subjects with regard to their metamer sets (Hilbert, 1987, pp96-97).

[27] A restriction has to be added here: two reflectances that have the same TIR are perceived as the same color if they appear in the same visual context. Simultaneous contrast effects are not explcitly considered in Hilbert's proposal. See Hilbert, 1987, p111, footnote 8).

[28] This sounds like a quite arbitrary stipulation – but I use it merely as a shorthand. There are many SSRs that look achromatic gray but whose reflectance is not a constant function of wavelength.

[29] Matthen seems not to distinguish as carefully between opponent process signals and integrated reflectances as does Tye (2000, pp. 160-161). But his point (in Matthen, 2001) is essentially the same as Tye's: metamers can be united under non-disjunctive types of reflectance, resulting in a characterization of broad color categories in terms of surface reflectance.

[30] In response to light increments, vertebrate photoreceptors show a hyperpolarization response in a graded fashion: the greater the light incerment, the stronger the response. In response to decrements in light, the photoreceptors show depolarization and secretion of synaptic transmitters. Under normal circumstances, the light falling on any given receptor is constantly fluctuating, and the receptor responds with a fluctuating

polarization of its membrane (DeValois and DeValois, 1997, pp. 99-100). It is shown by in vitro single cell recording studies that in the absence of light cone membranes exhibit a steady inward flow of sodium ions called the dark current. Dark current is regarded as the baseline level of activity, or "zero signal". When the photopigment absorbs light from a flash, it responds by a hyperpolarization of its membrane due to a slowdown of the inward flow of sodium. Then, as a result of overcompensation in restoring the resting potential, a depolarization follows. The amplitude of the photocurrent response increases with the stimulus intensity (Wandell, 1995, pp. 89-90). That is, zero photon absorption results in baseline activity, or zero signal; if there is photon absorption, then there is hyperpolarozation. If there is a decrease in photon absorptions per time unit, then there is a graded depolarozation – a cone signal that corresponds to a *decrement in positive value*. But just as there is no absorption rate that is less than zero photons per time unit, there is no cone signal that is interpreted as a negative value by the processing stages into which it is input. Cone signals indicate either no photon absorption or the presence of absorption.

[31] ASSC 5 Conference, Durham, North Carolina, May 27, 2001.

[32] Interface reflection (otherwise called the specular component) is the mirror-like component of reflection, responsible for shininess, that is typically not wavelength-selective [except for cases like the color *gold*]. In the overwhelming majority of the cases, only body reflection is wavelength-selective, and so this component determines the perceived color of reflecting surfaces. Body reflection emerges with the same likelihood in almost any direction [see Wandell, 1995, pp. 292-293]; ideal matte surfaces have only body reflection, no interface reflection.

[33] Measurements were taken at the National Research Council of Canada, Institute of National Measurement Standards. My measurements are not guaranteed to exhibit the high degree of accuracy that characterizes other measurements taken in that institute by more up-to-date equipment; however, as I said, for purposes of the present research, they are perfectly fine. I am grateful to Réjean Baribeau for providing me with equipment and assistance to collect my measurements.

[34] Surface reflectance is a function of wavelength; so is the SPD of illuminants. Mathematically, the color signal is the product of these two functions. Multiplying an SSR function with an SPD function which is constant over wavelength, with all its values being 1 does not change the values of the SSR function.

[35] Tye (2000, pp. 159-165) is more explicit on this point.

[36] This happens in the case of Stevens-type sensory functions, where the exponent of the power function characterizes the sensitivity of the perceiver to the particular aspect of the stimulus. The input of the Stevens transformation is some physical (perceiver-independent) measure of the stimulus; the output characterizes the sensory/perceptual response of the organism. See below in this section.

[37] In addition, it almost certainly has to include other parameters characterizing human observers. McCann et al. (1976, pp. 449-450) say the following. In order to achieve a good correlation between (1) triplets of reflectances characterizing the displays in the Mondrian experiments and (2) color perceptions, light reflection, weighted by cone spectral sensitivities, has to be further transformed using a power function with exponent 1.3 (apparently a Stevens-type sensory function whose exponent characterizes the sensitivity of observers). As McCann et al. note, they used this transformation to compensate for the fact that equal increments in reflectance do not represent equal increments in sensation. See below for further discussion.

[38] Surfaces that have the same tristimulus values under a specific illuminant are metamers under that illuminant – or identical in reflectance.

[39] Byrne and Hilbert (1997, p265) offer only one color-definition in Tye's style (that of green), without mentioning the need for any correction. I showed in the previous section why that definition is wrong. Since Byrne and Hilbert do not speak about corrections of the schema at all, the considerations in this section do not apply to their formulation. Tye remarks that in developing his generalized schema he was influenced by Byrne and Hilbert's definition (Tye, 2000, note 20 on p. 168).

[40] McCann et al. measured integrated triplets of reflectances, weighted by cone spectral sensitivities using three broadband telescopic photometers whose sensitivity extended the whole visible range. One of these photometers was equipped with a set of color filters whose transmittance approximated the sensitivity (probability of absorption) of the short wave cones (e.g., highest sensitivity of these cones around 445 nm was modeled as highest transmittance by the filters covering the photometer's sensor; the low sensitivity around 525 nm was modeled as low transmittance, and so on). The other two photometers were equipped with filters modelling the other two cone sensitivity curves. Triplets of weighted integrated refectances were measured thus: the three filter-equipped photometer was pointed at a sample surface (e.g., a Mondrian area, or a Munsell chip), then they were pointed at a standard white surface. Weighted integrated radiances

(i.e., three different weightings of the color signal by the filters) of the sample surface were then divided by those of the standard. See McCann, 1976, p. 449-453.

[41] 'abs' stands for absolute value.

[42] A properly chosen second-order polinomial function would do this favor for us.

[43] See the next section for the notion of perceiver-dependent properties.

[44] The CIE tristimulus values of colored surfaces are also perceiver-dependent properties, since they appeal to human color-matching functions, that is, linear transforms of cone spectral sensitivities (see above in this section). Tristimulus values can be regarded either as systematically corresponding to (describing?) cone activity ratios in standard circumstances of perception, or as values of arbitrary but useful mathematical transformations (functions) that give the same values for surfaces that we percive as the same in color (in standard circumstances). Calculating tristimulus values is an abstract mathematical transformation. The reason why we use it so frequently is that an equivalent mathematical transformation is implemented by our visual system – hence tristimulus values predict perceived color. There are no measurable stimulus properties, instantiated in the absence of human perceivers, that can be identified with (i.e., that are) the tristimulus values.

[45] Assuming that the members of such sets occur against the same background.

[46] Again, integrals of reflectance above the sensitivity ranges of the cones are *not* the same for members of metamer sets.

[47] Well, except for normal misperceptions – see below.

[48] Tye (electronic communication) gave this reply when I raised to him the problem of category mistakes (see the previous section).

[49] All graphs in this figure are displays of my own measurement data.

Notes for Chapter Three

[50] This is a reasonable claim since as we perceive it, the redness of ripe tomatoes, and that of stoplights or hot iron is very similar – phenomenally, or perceptually, they are the same kind of thing, despite small variations in shade.

[51] Note that the transmittance of a film is a function, among other things, of the thickness of the film. A thin layer of red wine is pink, or pale red; a thick layer of red wine is dark red.

[52] Emissive color changes more frequently in time than reflective color (think of color TV screens), but surface reflectance changes in time as well: think of cameleons' skins or tree leaves in the fall.

[53] However, the manifestation of reflectance, namely actual light reflection, is a physical event.

[54] And a similarity to fluorescent surfaces.

[55] Hence, in this particular case, it lacks any interesting theoretical interpretation.

[56] When calculating color-matching between reflective surfaces and, say, color monitor displays, the match is sought between (i) the color signal that arises at the reflecting surface (i.e., reflectance times external illumination) and (ii) simply the SPD of light emitted by the monitor. According to the principle of color matching, these two quantities, when multiplied by the standard trichromat observer's color-matching functions, should be equal – this predicts a match in perceived color for the standard observer.

[57] The reason why it does not look infinitely bright is well explained by the adaptation and limits of sensitivity of our visual system. But of course, these latter factors cannot be built into a type physicalist notion of object color.

[58] As far as I can tell, this idea is my own – I did not find it anywhere in the literature.

[59] Except for refraction – the change in speed and direction of the light ray when it enters the solid or liquid transmitting medium (see above and Nassau, 1997, pp. 24-28).

[60] And fluorescent transparent volumes that also exist.

[61] By the same coin, one can always find qualitative *differences* between any two non-identical particulars, at some level of abstraction.

Notes for Chapter Four

[62] It is not always easy to distinguish between normal and abnormal circumstances of perception, or, in other words, between what counts as the stimulus and what counts as part of the circumstances. In the movie case we might count illumination by the projector as part of the stimulus. The problem is, there is always something slightly "abnormal" in alleged cases of normal misperception. In cases where the illusion is stable, and it resists beliefs and perceptual learning, it is arguable that either the stimulus or the circumstances are not perfectly normal, in the sense that our perception has not been prepared in evolution to pick up and interpret *that* stimulus, in *that* circumstance, in a veridical way. So perhaps there is no such thing as an entirely *normal* misperception. If, on the other hand, one wants to maintain that there is such a thing as normal misperception, then one has to accept that normality (whatever it amounts to), is not immediately destroyed if there is some unusual aspect of the perceptual situation. Matthen (1988, pp. 11-13) nicely introduces the notion of normal misperception.

[63] However, the visual system might have an indirect acess to the overall brighness of the illuminant (Shepard, 1997, p. 324; Maloney and Wandell, 1986). For example, the brightness of the illumination might be estimated, independently of the light scattered by surfaces, from the brightness of the sky, and from the evidence for the presence of shadings.

[64] It is important to note that there is a remarkable change in Matthen's views on color that took place between his 1988 and 1999 pieces. In his earlier paper he defends type physicalism whereas in the latter one he no longer does so – even though he remains a color realist and a physicalist.

[65] One might argue that when we look at a color TV screen in a dark room, the circumstances of perception aren't normal as there's no external illumination present. I'm not totally convinced, but let us accept this objection. Still, when we look at an active computer monitor in an office illuminated by tungsten bulbs or fluorescent tubes, this objection does not apply, and we still perceive the emissive colors of the monitor perfectly well.

[66] I.e., fluorescent ones: see Nassau, 1997, pp. 10-13.

Notes for Chapter Five

[67] For philosophical purposes, there are significant differences between these two notions of representational content. Causal history (evolutionary history) is causally inefficacious: two organisms that are molecule by molecule duplicates (or as similar qualitatively as they can be, in every relevant respect), can nevertheless have radically different causal histories. An ordinary example is that of two cars of the same type made in different countries: two Toyota Tercel '94s can have exactly the same structure, performance, look, that is, exactly the same causally effective physical properties, despite the fact that one was made in Canada and the other in Japan. A fantastic example is that of Swampman, an exact duplicate of, say, me who was brought about by some cosmic coincidence from scattered organic matter in a swamp. Swampman has no human ancestors, and no evolutionary history. On the teleological notion of content, his sensory states have no representational content, as they were not designed, or evolved, to indicate anything. On the non-teleological notion of content, however, Swampman's perceptual states do have content: given that his physical constitution is like mine, the lawlike, counterfactual-supporting correlations between his perceptual states and stimulus properties immediately obtain. On independent grounds, it seems intuitively plausible to many philosophers that such a swamp creature would have conscious experience, and according to the representationalist view, this can happen only if he has perceptual content. Only the nomic correlation (non-teleological) notion of content provides the representationalist with this conclusion.

Another fantastic thought experiment, relevant in this context, is that of brains in a vat. Imagine that John Smith's brain is removed from his skull and is placed in an appropriate solution to maintain its biochemical functioning. In addition, the nerve endings are cut off from Smith's sense organs and hooked up to a supercomputer that supplies them with appropriate input signals while processing their output. By assumption, this simulation is perfect: Smith's nerve endings receive the same kinds of inputs that they did in his skull. Most philosophers' intuition is that in such a case John Smith would continue to have the kinds of conscious experience he had before, i.e., he would not notice any change. Now, if his conscious experience is normal, then, for the representationalist, his perceptual contents must remain normal too. The teleological notion of content provides the representationalist with this conclusion: John Smith does not

lose his evolutionary history by being "envatted". However, it is arguable that by cutting off the sense organs from the nerve endings, one breaks the counterfactual-supporting psychological laws that couple stimulus properties and sensory state activations together. Even if John Smith envatted was placed in ordinary circumstances of perception, he would not perceive anything – his perceptual systems are radically altered if not completely destroyed. So, on the non-teleological notion of content, he cannot have any perceptual content that could support the intuition, for the representationalist, that he has conscious experience. See Tye, 1995, p. 153 for a similar line of argument.

[68] Some philosophers deny that there are such things as disjunctive properties at all. I am liberal in this respect: as far as I am concerned, there can be disjunctive properties (i.e., it might make sense to speak about disjunctive properties). This does not save phenomenal externalism anyway.

[69] There were two people who mentioned this objection to me, independently of each other: John Kulvicki at the Tucson 2000 conference, and Dan Ryder at the "Consciousness and Emergence" conference at the University of Western Ontario, in April, 2001. At the latter event, there emerged a lunch-table discussion of this issue with the participation of Dan Ryder, Jillian McIntosh, William Seager, and Andrew Bailey. I am grateful to all these people for raising very interesting ideas in the discussion. However, it seemed to me that we were unable to come up with anything like a decisive objection against the disjunctive content idea. I shall raise an objection in the main text that, as far as I can tell, is my own, though it may have been facilitated by the just-mentioned discussion.

[70] The function of a system and the job a system does reliably need not be the same. Remember the distinction between the teleological and non-teleological notions of content.

[71] Here is another relevant example. Brown, like black, is regarded as a contrast color because we can only see brown in appropriate color contexts. Looking at a brightly illuminated chocolate bar through a cardboard tube whose inside is painted black, the chocolate bar will look orange – quite a surprising experience. In terms of surface reflectance, the chocolate bar is similar to, say, an orange peel. Both of them reflect relatively few light in the 400-550 nm range; the reflectance of both rises abruptly around 550 nm, and stays high until 700 nm. The difference is that the *average* reflectance of brown surfaces is significantly lower than that of orange ones. Perceptually, brown is blackened orange or blackened yellow. It might occur to someone that emitting surfaces (light sources) never look brown. The common sense intuition is something like that there is no such thing as "brown light". However, the idea that emitting surfaces never look brown is false: we can and do see a whole variety of browns on computer monitors or color TV screens. Such screens provide us with simultaneous displays of various color patches, hence the appropriate color contrasts to perceive certain areas of them as brown. What remains true is that, for instance, a single light source in darkness never looks brown – it can only look orange or yellow. But such a case is analogous with looking at the chocolate bar through a tube with black interior. Returning to the problem in the main text, I conjecture that before the invention of color TV screens and other emitting surfaces that can display complex stimuli, emissive brown was *very* rare in our environment. That is, perceptions of brown were once elicited almost exclusively by reflecting surfaces. In our contemporary, man-made environment, there are a lot of emissive brown stimuli around.

[72] I am grateful to Dan Ryder for discussing Dretske's views with me, and letting me know Dretske's answer to my questions about his view. I am also grateful to Andrew Brook for discussions on this point.

[73] According to scientific realism, it is science, ultimately physics (in its finished form), that represent the criterion of what there is. Ontologically significant predicates are those essential to the formulation of the correct physical theory. An opposing view of universals is sometimes called *apriorism*: it is the idea that we can determine what universals there are by mere armchair reflection on the stock of predicates in our language. To every non-equivalent predicate in a natural language there corresponds a separate and distinct universal. This may occur to someone as a too liberal criterion for the existence of universals; many realists think that restrictions need to be placed on the Platonic schema of universals (Loux, 1998).

[74] Well, not that one could not find *marginal* counterexamples to this rule. Supersaturated colors are one such example. But it remains the case that such cases are marginal and by no means essential to the function of color vision – unlike our capacity to construct new conceptual representations that do not correspond to any actual object or property.

[75] Again, the key difference between dispositionalism and disjunctive physicalism is that on the former, color is the role (the disposition), whereas on the latter, color is the accidental role filler (the basis of the disposition).

[76] See Smith, 1993, pp. 270-271 for what seems to me an example of such a line of reasoning.

[77] The situation is similar in the case of more complex functional states like propositional attitudes. Accepting the functionalist intuition, "beliefs cause behavior" is, strictly speaking, false. Beliefs are, functionalism tells us, functional states or causal roles. It is the belief-role-fillers (in humans, neurological states) that cause behavior. Beliefs in general are not identical with neurological states. Still, it is the relevant neurological states, the ones implementing beliefs and other propositional attitudes that participate in causal interactions. Beliefs are causal roles; causal roles do not cause; what happen to *play* these roles do.

[78] The (real or apparent) revelation of the phenomenal character of experiences is not a relevant analogy as phenomenal characters are not environmental stimulus properties that perception is after, but colors are – see Section 1.1.3.

Notes for Chapter Six

[79] We can assume that type physicalism about color is correct for broad color categories, or we can accept the family resemblance view of object color that I suggested in the previous part. The argument from individual differences can be formulated in both cases. In this sense the two arguments are independent of each other. The conclusion of the argument from individual differences is that there is something like a "freeplay" between object color properties and phenomenal color experiences. This, however, prevents us from finding unique correspondence relations between perceptually determinate color experiences (perceived shades as opposed to perceptual color *categories*) and determinate narrow ranges of object color properties like types of surface reflectance. This means that, for the population of color-normals, the object color unique green becomes disjunctive simply by means of the individual differences – unique green is either reflectance type R1 (for color-normal subject S1) or reflectance type R2 (for subject S2), … and so on for all subjects or groups of subjects that have relevantly different trichromat color vision.

[80] This is not exactly right as it stands – a complication is ignored. I shall describe and discuss this complication later. My first aim is to give a schema of the anti-representationalist argument from individual differences in color phenomenology – a somewhat simplified schema to which I can add the further wrinkles as I proceed.

[81] Block (1999, p. 45) writes: "The objective nature of color … derives from the overlap between persons with normal color perception. There are objects which would be categorized as 'blue' under ideal circumstances by everyone with normal color vision, and that's what makes them objectively blue."

[82] This conclusion is contested by Tye (2000). See below for detail.

[83] I.e., the non-disjunctive replectance properties that are true of all and only Samantha's and Max's metamer set respectively. Remember, we are assuming a hardline type physicalist view of color.

[84] BTW, in his two books about consciousness, Tye does not offer a definition for normal/optimal circumstances of perception. For his reasons, see Tye, 1995, note 16 on p. 226.

[85] I.e., we cannot perceive the type of reflectance that is green, only instantiations of this type, but such instantiations are particular (maximally determinate) reflectances, instances of minimal shades. Greenness as an abstract property (reflectance type) is always realized together with further differentia specifica that are not necessary for being green. Analogy: we cannot directly perceive "the generalized mammal", only instantiations of a certain mammal species like a particular dog, for the same reason as with green.

[86] I.e., the *category* that is represented consists of minimal shades; the *category representation* is either a range of PDCEs or, as in the case of faint color memories, it is a single, less determinate color experience (take the latter as a speculative proposal – I shall not argue for it, nor is this idea necessary for the argument I'm making). See below in the main text.

[87] None of the type physicalists I know of (most importantly A. Byrne, D. Hilbert, and M. Tye) take this route. It seems they are all prepared to endorse objective incompatibility, so what I mention in this paragraph of the main text is just a possible theoretical route.

[88] As I argued earlier, Tye's schema for defining object colors in terms of surface reflectance is wrong. However, in the argument from individual differences I do not need this result – assuming that Tye's schema is right does not affect the latter argument. This shows that the two arguments I offer in this dissertation are quite independent of each other.

[89] Recall Block's notion of person-relative color (Block, 1999 and Section 3.2 above). The idea is, different color-normal subjects may differ in how they categorize minimal shades perceptually. A particular surface

may be categorized by Samantha as orangeish red whereas by Max as plain red, while both their perceptions are veridical. However, it follows from objective incompatibility that if a surface is categorized by one of them as blue-red, and by the other as yellow-red, then there is a mistake involved in one of their color perceptions, since no surface can objectively be both bluish and yellowish: being blue(ish) and being yellow(ish) are incompatible stimulus properties.

[90] On Tye's schema (2000, Ch. 7) we can avoid concluding that unique green and bluish (yellowish) green are contraries, by saying that the meanings of "approximately equal" and "significantly more/less" are related by a fuzzy boundary. I.e., in his schema (for reflecting surfaces only), unique green surfaces reflect approximately the same amount of light in the short wavelength range as in the medium-plus-long wavelength one; bluish greens reflect significantly more light in the short wavelength range than in the medium-plus-long wavelength range. Now what is significantly more as opposed to *in*significantly more (i.e., approximately equal) is a matter of discussion. Tye was careful enough to avoid any specification here. See also Byrne and Hilbert (1997, pp. 272-274 on this point).

[91] The experience of unique green is chromatically as determinate as it can be. There are no *chromatic* variations of unique green (and the experience of it): unique green is the color that is neither bluish nor yellowish. There are variations of *unique green as experienced*, in terms of lightless and saturation, but when speaking about monochromatic lights, brightness and saturation values are also pretty well specified. Monochromatic lights look highly, perhaps *maximally*, saturated. The intensity of monochromatic lights and their corresponding perceived brightness can be regarded as fixed at an arbitrary but reasonable level – this helps to avoid questions arising from the Bezold-Brücke effect. But if all three dimensions of color space are fixed, then we arrive at a PDCE – a fully specified color experience. This is why I say that the experience of unique green, in the context of monochromatic lights as stimuli, is essentially a perceptually determinate color experience.

[92] In actuality such non-overlapping ranges of reflectance are not very different; indeed, they are quite close to one another.

[93] Colorimetry makes this assumption (see e.g., Wyszecki and Stiles, 1967/1982). In colorimetry, object color is (often) identified with the color signal: surface reflectance times illuminant spectral power distribution (Wyszecki and Stiles, 1967, p. 279).

[94] Throughout this section I always imply the sameness of perceptual circumstances for the hypothetical subjects whose color perception is compared.

[95] American Optical Company, 2nd Edition.

[96] The author is grateful to Rejean Baribeau and Jessica Cox for for making available their laboratory and equipment, and providing assistance with the preparation and data collection.

[97] 1967 Edition, Davidson and Hemmendinger, 2857 Nazareth Rd. Easton, Pa. More recent address: Hemmendinger Color Laboratory, 438 Wendover Drive, Princeton, New Jersey, 08540

[98] In general, there is no hue angle in L*C*h color space that corresponds to, say, "standard unique red" (the same holds for the other three unique hues). In colorimetry there are no such things as standard unique hues, let alone color stimuli (e.g., particular surfaces with a determinate reflectance) that look chromatically unique to the overwhelming majority of color normals. However, it would be too early to declare this at this point in the text, since this indeterminacy is the very phenomenon that I want to demonstrate in the present experiment. If one set up the four color series in such a way that the grand mean of unique hue choices is 7.00 for each of them, that alone would still not make it the case that there is a sample in the series (i.e., the seventh) that is the "official unique hue". This is because even if the average unique choice in a series is sample 7, still, the majority (or *all*) of the subjects might choose some sample other than the seventh as unique (e.g., all females consistently prefer sample 6 whereas all males consistently prefer sample 8).

[99] Note that it is possible to argue the same way if a subject named a sample blue in 75 per cent of the cases and greenish blue in 25 per cent of the cases. This motivates the use of the strict criterion for unique hue perceptions.

[100] This is true in general when we want to classify values of a probabilistic variable. For instance, it would be impossible to establish just noticeable differences (jnd's) between stimuli without introducing (essentially arbitrary) probability thresholds. For what counts as noticing a difference between two stimuli? Noticing it in 100 % of the cases? Psychophysicists prefer less strict criteria: the tradition is to use 75 per cent. That is, if a subject signals a difference between two stimuli in at least three quarter of the cases, then

he is credited with noticing the difference – his jnd is the stimulus difference that he notices with a probability of 0.75.

[101] Given 15 subjects, there are 15*14/2=105 different pairs. Out of these, 6 are of interest (or perhaps 10, if we count in the blue series cases). 6/105=0.0571, that is, 5.7% of the pairs shows at least one of the interesting phenomena.

[102] Byrne and Hilbert (1997, pp. 279-281) offer a very similar solution to the problem of unique-binary distinction.

References:
Akins, K. (1996). Of Sensory Systems and the "Aboutness" of Mental States, *Journal of Philosophy*, 337-372.

Armstrong, D. (1989). *Universals: An Opinionated Introduction*. Boulder: Westview Press

Ayama, M., Nakatsue, T., Kaiser, P., K. (1987). Constant hue loci of unique and binary balanced hues at 10, 100, and 100 Td. *Journal of the Optical Society of America A*, 1136-1144.

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

Boghossian, P., A., Velleman, D., J. (1997a). Colour as a Secondary Quality. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 1. (The Philosophy of Color) Cambridge Mass. The MIT Press, 81-103.

Boghossian, P., A., Velleman, D., J. (1997b). Physicalist Theories of Color. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 1. (The Philosophy of Color) Cambridge Mass. The MIT Press, 105-136.

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

Chalmers, D. (1996). *The Conscious Mind; In Search of a Fundamental Theory*. Oxford University Press.

Cherniak, C. (1984). Computational Complexity and the Universal Acceptance of Logic, *Journal of Philosophy*, **81,** (12), 739-758.

Cohen, J. (2000). Color Properties and Color Perception: A Functionalist Account. PhD thesis, Rutgers University, New Brunswick, New Jersey.

Davidson, D. (1970). The Individuation of Events. In *Essays in Honor of Carl G. Hempel.* Ed. N. Rescher, Dordrecht, Holland: Reidel, 216-234.

Davidson, D. (1984). Truth and Meaning. In D. Davidson: *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford

Davies, M. (1997). Externalism and Experience, In: Ned Block, Owen Flanagan and Güven Güzeldere (Eds.): *The Nature of Consciousness*, Cambridge Mass. MIT Press

DeValois, R., L., DeValois K., K. (1997) Neural Coding of Color. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 2. (The Science of Color), Cambridge Mass. The MIT Press, 93-140.

Dretske, F. (1988). *Explaining Behavior; Reasons in a World of Causes*. Cambridge Mass. The MIT Press

Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, Mass. MIT Press

Evans, G. (1982). *Varieties of Reference*. Oxford, Clarendon Press

Fairchild, M., D. (1998). *Color Appearance Models*. Addison-Wesley Publishing Company

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. (1990a). A Theory of Content, I: The Problem. In J. Fodor, *A Theory of Content and Other Essays*, Cambridge Mass. The MIT Press

Fodor, J.A. (1990b). A Theory of Content, II: The Theory. In J. Fodor, *A Theory of Content and Other Essays*, Cambridge Mass. The MIT Press

Fodor, J. A. (1994). *The Elm and the Expert*. Cambridge Mass. The MIT Press

Gibbard, A. (1995). Visible Properties of Human Interest Only. *Philosophical Issues*, **7** (Perception), 199-208.

Goldman, A. (1970). *A Theory of Human Action.* Englewood Cliffs, N.J.: Prentice-Hall.

Goodman, N. (1965). *Fact, Fiction, and Forecast*. Indianapolis: Bobbs-Merrill.

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

Hardin, C., L. (1995). Color Subjectivism. In A. Goldman (Ed.): *Readings in Philosophy and Cognitive Science*, Cambridge Mass. The MIT Press, 493-507.

Hardin, C., L. (1997). Reinverting the Spectrum. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 1. (The Philosophy of Color), Cambridge Mass. The MIT Press, 289-301.

Hatfield, G. (1992). Color perception and neural encoding: does metameric matching entail a loss of information? In M. Forbes and D. Hull (Eds.): *PSA, 1992*, 2 volumes. East Lansing, MI: Philosophy of Science Association.

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

Hilbert, D., R. (1987). *Color and Color Perception. A Study in Anthropocentric Realism.* Stanford: Center for the Study of Language and Information

Hoffman, D. (2001). The Data Problem for Color Objectivism. *Consciousness and Cognition*, **10**, (1), 74-77.

Hunt, R., G., W. (1982). A model of colour vision for predicting colour appearance. *Color Research and Application*, **7**, (2), 95-112.

Hunt, R., G., W. (1991). *Measuring Color*. Second Edition, Ellis Horwood Limited, 1991.

Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Quarterly* **32**, 127-136.

Jackson, F. (1986). What Mary Didn't Know. *Journal of Philosophy* **83**, (5), 291-295.

Jackson, F., Pargetter, R. (1997). An Objectivist's Guide to Subjectivism About Color, In Alex Byrne and David R. Hilbert (Eds.), *Readings on Color*, Vol. 1. (The Philosophy of Color), Cambridge Mass. The MIT Press, 67-79.

Jakab, Z. (2000). Ineffability of qualia: a straightforward naturalistic explanation. *Consciousness and Cognition*, **9**, (3), 329-351.

Jakab, Z. (2000). Reply to Thomas Metzinger and Bettina Walde. *Consciousness and Cognition*, **9**, (3), 363-369.

Johnston, M (1992). How to Speak of the Colors, *Philosophical Studies*, **68**, 221-263.

Kim, J. (1976). Events as Property Exemplifications. In M. Brand and D. Walton (Eds.): *Action Theory*, Dordrecht, Holland: Reidel, 159-177.

Kuehni, R., G. (1999). Towards an Improved Uniform Color Space, *Color Research and Application*, **24**, (4), 253-265.

Kuehni, R., G. (2000a). Uniform Color Space Modeled with Cone Responses, *Color Research and Application*, **25**, (1), 56-63.

Kuehni, R., G. (2000b). A Comparison of Five Color Order Systems, *Color Research and Application*, **25**, (2), 123-131.

Kuehni, R., G. (2001). Determination of Unique Hues Using Munsell Color Chips. *Color Research and Application*, **26**, (1), 61-66.

Kuhn, Th. S. (1974). Second Thoughts on Paradigms. In *The Structure of Scientific Theories*. Frederic Suppe, ed. University of Illinois Press.

Land, E., H. (1977). The Retinex Theory of Color Vision. *Scientific American*, **237**, (6), 108-128.

Land, E., H. (1997). Recent Advances in Retinex Theory. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 2. (The Science of Color) Cambridge Mass. The MIT Press, 143-159.

Lewis, D. (1997) Naming the Colours. *Australasian Journal of Philosophy*, **75**, 325-342.

Lotto, R., B., Purves, D. (2000). An empirical explanation of color contrast, *Proceedings of the National Academy of Sciences*, **27**, (23)

Lutze, M., Cox, J., Smith, V., C., Pokorny, J. (1990). Genetic Studies of Variation in Rayleigh and Photometric Matches in Normal Trichromats, *Vision Research*, **30**, (1) 149-162.

Lycan, W. (1996). *Consciousness and experience*. Cambridge Mass. The MIT Press

Maloney, L., T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *Journal of the Optical Society of America* A, **3**, (10), 1673-1683.

Maloney, L., T., Wandell, B., A. (1986). Color constancy: a method for recovering surface reflectance. *Journal of the Optical Society of America* A, Vol. 3, No. 1, January, 29-33.

Matthen, M. (1988). Biological functions and perceptual content. *Journal of Philosophy*, **85**, (1), 5-27.

Matthen, M. (1999). The Disunity of Color. *Philosophical Review*, **108**, 47-84.

Matthen, M. (2001). What Colors? Whose Colors? *Consciousness and Cognition*, **10**, (1), 117-124.

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

Maund, B. (1995). *Colours: Their nature and representation*. Cambridge, UK: Cambridge University Press.

McAdam, D., L., (1997). The Physical Basis of Color Specification. In Alex Byrne and David R. Hilbert (Eds.), *Readings on Color*, Vol. 2. (The Science of Color), Cambridge Mass. The MIT Press, 33-63.

McCann, J., J., McKee, S., P., Taylor, T., H. (1976). Quantitative studies in Retinex Theory; A comparison between theoretical predictions and observer responses to the "Color Mondrian" experiments. *Vision Research*, **16**, 445-458.

McGilvray, J., A. (1994). Constant colors in the head. *Synthese*, **100**, 197-239.

McLaughlin, B. (2001). The Place of Color in Nature. Manuscript.

Meyer, G., W. (1988). Wavelength selection for synthetic image generation. *Computer vision, Graphics, and Image Processing*, **41**, 57-79.

Miller, D.L. (1997). Beyond the elements: investigations of hue. In C. L. Hardin and Luisa Maffi (Eds.), *Color categories in thought and language*, Cambridge University Press, 151-162

Nassau, K. (1997). The Causes of Color. In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 2. (The Science of Color), Cambridge Mass. The MIT Press, 3-29.

Neale (1990). *Descriptions*. Cambridge Mass. The MIT Press

Neitz, J., Jacobs, G. (1986). Polymorphism of Long-wavelength Cone in Normal Human Color Vision, *Nature*, **323**, 623-625.

Neitz, M., Neitz, J. (1998). Molecular Genetics and the Biological basis of Color Vision, in *Color Vision: Perspectives from different disciplines*, Ed. W. G. Backhaus, R. Kriegl, and J. S. Werner, Berlin: DeGreuter

Neitz, J., Neitz, M., Jacobs, G. (1993). More than Three Different Cone Pigments Among People with Normal Color Vision, *Vision Research*, **33**, (1) 17-122.

Nickerson, D. (1977). History of the OSA Committee on Uniform Color Scales, *Optics News* (Optical Society of America), Winter issue

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.

Quinn, P. C., Rosano J. L., Wooten, B. R. (1988). Evidence that brown is not an elemental color. *Perception and Psychophysics*, **43**, (2), 156-164.

Raffman, D. (1995). On the Persistence of Phenomenology. In Thomas Metzinger (Ed.): *Conscious Experience*. Imprint Academic, Schöningh, 293-308.

Ross, P. (2001a). The Location Problem for Color Subjectivism, *Consciousness and Cognition*, **10**, (1), 42-58.

Ross, P. (2001b). Locating Color: Further Thoughts, *Consciousness and Cognition*, **10**, (1), 146-156.

Shepard, R., N. (1997). The Perceptual Organization of Colors: An Adaptation to Regularities of the Terrestrial World? In Alex Byrne and David R. Hilbert (Eds.): *Readings on Color*, Vol. 2. (The Science of Color), Cambridge Mass. The MIT Press, 311-356.

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

Shoemaker, S. (1996). Qualities and qualia: What's in the mind?, in S. Shoemaker: *The first-person perspective and other essays*, Cambridge University Press

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

Sternheim, C., E., Boynton, R., M. (1966). Uniqueness of perceived hues investigated with a continuous judgmental technique. *Journal of Experimental Psychology* **72**, 770-776.

Thompson, E. (1995). *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception*, London and New York: Routledge.

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.

Tye, M. (1995). *Ten Problems of Consciousness.* Cambridge Mass. The MIT Press

Tye, M. (1998). Externalism and memory. *Proceedings of the Aristotelian Society*, Supp. Vol.

Tye, M. (2000). *Consciousness, Color, and Content.* Cambridge Mass. The MIT Press

Wandell, B., A. (1995). *Foundations of vision.* Sinauer Associates, Sunderland, Massachusetts

Werner, J., S., Wooten, B., R. (1979). Opponent chromatic mechanisms: Relation to photopigments and hue naming. *Journal of the Optical Society of America*, **69**, 422-434.

Wright, L. (1973). Functions. *Philosophical Review*, **82**, (2), 139-168.

Wyszecki, G., Stiles, W., S. (1967). *Color Science – Concepts and Methods, Quantitative Data and Formulas*. John Wiley and Sons, NY.