NOÛS 40:2 (2006) 205–240

# *Sensory Awareness Is not a Wide Physical Relation: An Empirical Argument Against Externalist Intentionalism*[1]

Adam Pautz
The University of Texas at Austin
The Australian National University

## 1. Introduction

**Intentionalism** is very popular. Here I will work with a simple form for convenience. On this form of Intentionalism, to have an experience is to stand in **the sensory representation relation** or, as I shall say, **the property-awareness relation** to a cluster of properties (Dretske 1999). The properties we are aware of are the ostensible properties of external objects or parts of one's body. In veridical experience these properties are instantiated in one's environment or body, while in non-veridical experience they are not. In hallucination, for instance, one is aware of properties but not anything that instantiates the properties; one is aware of a cluster of free-floating properties (Dretske 1995, 1999, 2003; Tye 2000). Finally, the qualitative character of experience is determined by the totality of properties one is aware of (the totality of properties one sensorily represents). Roughly speaking, necessarily, if two people are aware of the same properties, then they have the same experience-type.

Of course, we are not only aware of properties; in veridical experience, we are also aware of objects and facts. But object-awareness and fact-awareness do not determine qualitative character (Dretske 1999). It is what properties things seem to have which determines qualitative character. Here I will simply speak of the "awareness relation". But it is to be understood that I am exclusively interested in the nature of the property-awareness relation.

There is much to recommend Intentionalism. Unlike Disjunctivism, Intentionalism is a common factor theory. The common factor between

veridical and non-veridical experience is the sensory representation of a cluster of properties. At the same time, unlike adverbialism and certain qualia-based theories, Intentionalism is supposed to accommodate the **transparency** or **relationality** of all experience, even hallucinatory experience. By this I mean, roughly, the apparent fact that when we focus on our experience, even if it is hallucinatory, we focus on *something or other* that we experience. On Intentionalism, the *something or other* in question is a property, which may or may not be instantiated where it seems to be. Other theories share these twin virtues—for instance the Sense Datum Theory and the Meinongian Theory—but it may be argued that Intentionalism is superior to such theories because it avoids the postulation of strange particulars as the common objects of veridical and non-veridical experience. For these reasons, while I do not accept the simple form of Intentionalism that I will work with here, I believe that some form of Intentionalism is correct.[2]

On Intentionalism, states of sensory consciousness are relational properties: they divide into an intentional relation and a *relatum* (a complex of properties, a proposition, or other intentional object). On the simple form of Intentionalism I will be working with here, the relation in question is the awareness relation (the sensory representation relation) and the *relatum* is a cluster of properties, which I will call the **Q-properties**. The Q-properties are the properties the awareness of which determines qualitative character. They include colors, sounds, the pain qualities we feel in various parts of our bodies, as well as so-called "primary" qualities such as shapes. Therefore, on Intentionalism, the hard core of the mind-body problem, namely the problem of sensory consciousness, becomes two problems: the problem of finding a place for the awareness relation in the physical world, and the problem of finding a place for the Q-properties.

**Reductive Intentionalism** adds to Intentionalism the claim that the awareness relation is identical with a physical relation and the Q-properties are identical with physical properties. Here I use 'physical' in a broad sense, so that functional and topic-neutral properties and relations count as physical. Most Reductive Intentionalists defend **Externalist Intentionalism** (Dretske 1995; Lycan 1996, 2001; Tye 1995, 2000). By this, I mean the conjunction of Intentionalism and the claim that the awareness relation is identical with a *wide* physical relation between sentient creatures and properties outside the head. By 'wide physical relation', I mean a physical relation which is such that neurobiological duplicates living under the same laws can differ in what they bear it to: for instance, a relation defined in terms of causal covariation under optimal conditions (Tye 1995, 2000) or biological function (Dretske 1995). Externalist Intentionalism entails **Externalism**: the thesis that neurobiological duplicates living under the same laws can differ phenomenally. It is not surprising that most Reductive Intentionalists are Externalists, because our most plausible accounts of intentional relations in general are wide. Most Externalists also claim that the Q-properties are necessarily

identical with the physical properties that the states of our sensory systems track. For instance, colors are identical with reflectance properties, sounds are identical with properties of sound waves, tastes and smells are identical with chemical properties, felt pains are identical with types of bodily disturbance, and so on. Call this **Physicalism about the Q-properties**. Hence, according to Externalists, the state of having an experience with a certain qualitative character is identical with a wide state: the state of standing in a wide physical relation to a certain physical property (Tye 1995, 162–163). By being in such states, we are directly aware of physical properties which may or may not be instantiated in the environment or body. (Of course, we are not aware of them *as* physical properties. See Tye 2000, 55–56.) Further, the qualitative character of experience is fully determined by the physical properties one is thus aware of.

I am opposed to Reductive Intentionalism in any form. Instead I defend **Primitivist Intentionalism**. I accept the Brentano/Meinong thesis that the awareness relation, in other words the sensory representation relation, is a primitive relation between individuals and properties which is not identical with any physical relation. I also hold that the Q-properties, for instance colors and felt pains, are primitive properties which cannot be identified with any physical properties of external objects or bodily regions. In fact, I believe that, while these properties exist (in my view we are related to them in experience), nothing in fact instantiates them, including our own experiences.

However, I do not arrive at Primitivism about the awareness relation and the Q-properties on the basis of the usual *a priori* arguments: the conceivability of certain kinds of spectrum inversions, the conceivability of individuals physically like us in all narrow and wide respects but lacking sensory awareness (Chalmers 1996, McGinn 1997), or the conceivability of worlds physically like our own in the pattern of instantiation of reflectance properties and the like but in which Q-properties are not instantiated (Chalmers 2006). In my view, such arguments are not required. My case for Primitivism relies on **Dependence**: the claim that internal factors play a role in determining the qualitative character of experience (Pautz 2004). Consider a case in which two possible individuals are in different internal neural states and consequently have different behavioral dispositions. Dependence is no stronger than the claim that, at least in some such cases, the correct verdict is Different Experiences: the individuals involved have different experiences. I do not accept this claim on the basis of intuition; rather, as we shall see, I accept it on the basis of our best empirical theories of qualitative character, together with a plausible principle concerning the link between experience and behavior.

In this paper, I will carry out a step in the argument from Dependence to Primitivism about the awareness relation. I will set to the side the issue of the Q-properties. In particular, I will use Dependence, together with a

principle concerning the link between experience and behavior, to develop a counterexample to Externalist Intentionalism: the conjunction of Intentionalism and the claim that the awareness relation is identical with a *wide* physical relation. The counterexample is a case of the noted kind involving color vision. The correct verdict in this case, I will argue, is Different Experiences. But the conjunction of Intentionalism and the claim that the awareness relation is a wide physical relation delivers the mistaken verdict of Same Experiences. Hence the Intentionalist cannot identify the awareness relation with a wide physical relation.

My plan is as follows. In §2, I provide some scientific background on color vision. (Those familiar with the science of color vision may wish to skip this section and proceed directly to the counterexample.) In §3, I turn to the counterexample. In §4, I reply to objections. Finally, in §5, I suggest that the proper response to the counterexample is not to abandon Intentionalism. Rather, it is to combine standard Intentionalism with the claim that internal factors play a role in determining what colors, pains, tastes and other Q-properties we are aware of. I then mention some problems which anyone who accepts this view must confront.

## 2. Background

First, a terminological preliminary. I accept Intentionalism, but it will be well to state the argument in a way that is neutral between competing accounts of the nature of color experience. For this purpose, I will introduce some neutral terminology for characterizing color experience. Say that a color experience is **reddish** iff it resembles the color experiences actually normally produced by objects that we call 'red'; say that a color experience is **greenish** iff it resembles the color experiences actually normally produced by objects that we call 'green'; and so on. Here I mean to fix the reference of these terms, not give their meanings. This terminology may be extended to other types of experiences in the obvious way. As an Intentionalist, I would say that a color experience is reddish iff in having it we sensorily represent (are aware of) a reddish color; but the terminology is compatible with other views.

I will begin by describing some facts about color experience. Then I will show how, according to the so-called opponent process theory of color vision, the explanation of these facts is to be found in the brain.

Some color experiences are **unitary**. For instance, some reddish experiences are neither yellowish nor bluish. We have exactly four types of color experiences that are unitary: unitary reddish, unitary greenish, unitary yellowish, and unitary bluish color experiences. All other color experiences are **binary**. For instance, every orange experience is to various degrees *both* reddish and yellowish. Besides the four unitary color experiences, every color experience we are capable of having is a binary combination having

two unitary components in a certain proportion (Hurvich 1981, chapter 1). Once we have had a little practice we can come to impressive agreement on how to characterize color experiences in terms of the four hue primitives, using percentages to express degree. So, for instance, we might describe a purple experience as 60% reddish and 40% bluish. Some binary combinations are excluded, however. We have no red-green or yellow-blue experiences. (Maybe we do under unusual experimental situations. See Crane and Piantanida 1983 and Billock et al. 2001.) In this sense, *reddish* and *greenish*, and *yellowish* and *bluish*, form **opponent pairs**. We may say, then, that color experiences vary along two chromatic dimensions: red/green and yellow/blue, on which they may take positive or negative values.

The claim that color experience has a unitary-binary opponent structure is to some degree based on direct introspection. It is also supported more indirectly by countless psychological experiments on hue-cancellation and color naming (e. g. Boynton et al. 1964) and is generally supposed to receive some support from cross-cultural linguistic evidence, although the details are controversial (Kay et al. 1991; MacLaury 1997, chapter 2).

The so-called **opponent process theory** of color vision shows us that the best explanation of the character of the quality space for color experience is to be found in the brain (Hurvich 1981, Hardin 1988, Gordon and Abramov 2004, De Valois and De Valois 1993, De Valois 2004). I will work with a simple form of the theory.

There are three **cone-types** on the retina, each most sensitive to a different range of wavelengths within the visible spectrum. Since these ranges of wavelengths overlap considerably, in normal situations cones of all three types are simultaneously active.

Further downstream, according to the opponent process theory, the outputs of the cone systems are summed and differenced to create two neural channels. Each channel is a system of neurons with similar response profiles. Depending on the light received by the eye, the overall firing rate of a channel can be above or below the base rate of firing, or "positive" or "negative" for short. One channel, which I will call the $\phi$ **opponent channel**, corresponds to the red-green direction of color space. When it assumes a positive state of activation we have a reddish experience; when it assumes a negative state of activation we have a greenish experience. The other channel, which I will call the $\psi$ **opponent channel**, corresponds to the yellow-blue direction of color space. When it assumes a positive state of activation we have a yellowish experience; when it assumes a negative state of activation we have a bluish experience. Thus, according to opponent process theory, the values a color experience has on the phenomenal parameters red/green and yellow/blue are *determined by* the overall (relative) levels of activity in the $\phi$ and $\psi$ opponent channels. The opponent channels, so to speak, provide a **neural interpretation** for the quality space for color experience in the brain.

The **chromatic response curves** are shown in Figure 1. They are determined indirectly by a psychophysical method known as hue-cancellation (Hurvich and Jameson 1955). According to the opponent process theory, they reflect the levels of activity assumed by the $\phi$ and $\psi$ opponent channels in response to monochromatic light across the visible spectrum. In particular, the thin line represents the activity of the $\phi$ channel above or below base rate in response to monochromatic light across the visible spectrum, while the other line represents the activity of the $\psi$ channel. (The figure also features an achromatic curve which we may ignore here.)

Given information about the light entering an individual's eyes, one may use the response curves to predict the character of the resulting experience. For instance, if one calculates from the response curves that the light entering an individual's eyes produces relative $\phi$ activity of $+0.6$ and relative $\psi$ activity of $-0.4$, then one may predict that she will have an experience that is 60% reddish and 40% bluish (Hurvich 1981).

Now we are in a position to explain the phenomenal facts noted above. We have exactly four unitary color experiences because we have a unitary color experience just in case either the $\phi$ channel or the $\psi$ channel is put into a positive or negative state of activation while the other is put into neutral balance, and there are exactly four ways in which that can happen. For

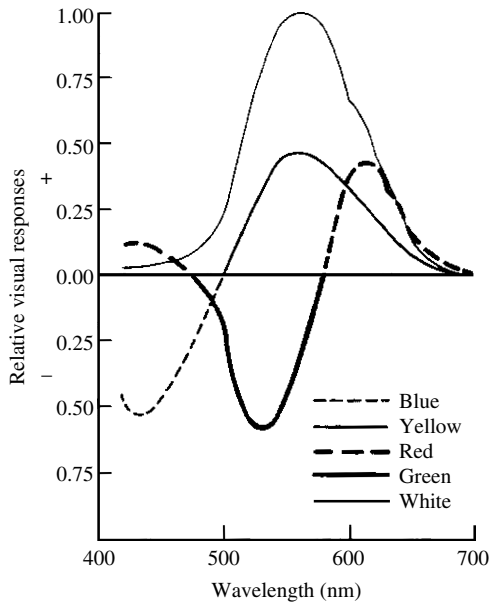



**Figure 1.** Chromatic and achromatic response curves for an equal energy spectrum and a single observer. From Leo M. Hurvich 1981. Reprinted with permission from Leo M. Hurvich.

instance, we have a unitary greenish experience when 500 nm monochromatic light enters our eyes, since, as Figure 1 shows, this puts our $\phi$ channel into a negative state of activation and our $\psi$ channel into neutral balance (in other words, 500 nm is the $\psi$ channel *zero crossing*). The various binary color experiences result from the simultaneous (positive or negative) activation of both the $\phi$ channel and the $\psi$ channel. For instance, when 580–700 nm monochromatic light enters the eyes, so that the $\phi$ channel assumes a positive state of activation (coding reddishness) and the $\psi$ also assumes a positive state of activation (coding yellowishness), we have a reddish-yellowish experience. Resemblances among color experiences are explained by resemblances among opponent channel states. The absence of binary reddish-greenish experiences and binary yellowish-bluish experiences (under normal circumstances) is explained by the fact that it is impossible to put either the $\phi$ channel or the $\psi$ channel into a state of both positive and negative activation at once.

The opponent process theory also explains many psychophysical phenomena. For instance, **complementary after-images** are neatly explained in terms of neural rebound. When we have a reddish experience, for instance, the $\phi$ channel is in a positive state of activation. When we cease to have the experience, the $\phi$ channel tends to rebound and go into a negative state of activation, producing a greenish afterimage. It also provides convincing explanations of countless other psychophysical phenomena which I cannot go into here, including the forms of color deficiency, the Bezold-Brucke phenomenon, certain facts about subjects' patterns of discrimination, hue-cancellation, color naming, temporal resolution for flickering stimuli, mixture thresholds for color stimuli, and second site adaptation effects.

The opponent process theory also receives some direct physiological confirmation. Direct recording of neurons has revealed that the outputs of the cones are in fact summed and differenced to create two systems of neurons in the LGN (a kind of relay station between the eyes and the visual cortex) roughly corresponding to the red-green and yellow-blue cardinal directions of color space. However, there are some discrepancies between their activity and color experience. The activity of these neurons does not match the psychophysically-derived response curves shown in Figure 1. For this reason and others, these systems of neurons cannot constitute the $\phi$ channel or the $\psi$ channel hypothesized in order to explain the psychophysical data (De Valois and De Valois 1993). Chromatically opponent cells are found in the different areas of the visual cortex, but even at the population level there is no evidence that they fall into distinct chromatic classes (Lennie 1999, 240). Color scientists have proposed detailed multistage models which are supposed to solve these problems (De Valois and De Valois 1993; De Valois 2004; Guth 1991). No doubt the model I am working with here is oversimplified; but I work with it for the purposes of illustration.

According to opponent process theory, opponent channel activity (or something like it) plays a direct role in explaining the character of color experience. For this to be so, the following must obtain:

> **C-Dependence**: Opponent channel activity plays a direct role in determining the character of color experience. By virtue of their reflectance profiles, an object reflects certain light and sets up certain opponent channel activity in us. In turn, that activity **directly** determines the character of the resulting color experience.

C-Dependence is compatible with a variety of philosophical views on color experience, for instance the Identity Theory, Functionalism, and Dualism. The modally weakest interpretation of C-Dependence is a Dualist one. On this view, C-Dependence is supported by a brute psychophysical law directly linking opponent activity with color experience. C-Dependence is also compatible with multiple realizability. Granted, it implies that neural differences above a certain level of grain make for phenomenal differences. In particular, it implies that differences at the level of opponent channel states, which are likely to have repercussions on downstream processing and behavior, make for phenomenal differences. But it does not imply that neural differences below this level of grain make for phenomenal differences.

Note that C-Dependence says only that internal factors **play a role** in determining color experience. For this reason, it does not entail **Internalism** about color experience: the strong thesis that internal factors *completely* determine color experience, so that neurobiological duplicates living under the same laws have the same color experiences. Nevertheless I will argue in the next section that C-Dependence is enough to rule out all the versions of Externalism that I know of.

Why accept C-Dependence? First, it is supported by an inference to the best explanation. Nothing in the outside world can explain the unitary-binary character of color experience, the comparative resemblance relations among color experiences, or the psychophysical phenomena listed above. So, the explanation must lie in the brain, even if it does not take the simple form discussed here. Second, systems of neurons whose activity approximates the hypothesized $\phi$ and $\psi$ channels have been discovered in the early visual system.

Dependence is not limited to color vision. In the other sensory modalities, the empirical case for the direct role of internal factors in determining qualitative character is perhaps even stronger. The extent or intensity of a bodily disturbance is a very poor predicator of the intensity of the resulting pain. For instance, doubling the strength of electric shock causes more than a doubling of pain intensity (Stevens 1975). By contrast, pain intensity is **linearly correlated with** neuronal discharge frequencies in many cortical

areas (Coghill et al. 1999, p. 1936; Porro et al. 1998; Coghill et al. 2003). Consequently, it is agreed among neuroscientists that neuronal discharge frequencies in some brain circuits directly determine pain intensity, although which brain circuits these are remains a subject of considerable debate. Likewise, resemblances among taste experiences in qualitative character are matched by resemblances among across fiber patterns in the brain (Smith et al. 2000). This suggests that qualitative resemblance among taste experiences is internally-determined. Thus, while I focus here on the opponent process theory for purposes of illustration, my argument does not depend on that theory. If one rejects the opponent process theory, then one may consider another example. In fact, the argument does not hinge exclusively on Dependence at all; as we shall see, I believe that counterexamples of the kind I will discuss may also be generated on the basis of a plausible principle concerning the link between experience and behavior.

Owing to the explanatory gap, we may not know the why or wherefore, but this much is known: qualitative character is systematically configured and constrained by processing taking place beyond the receptors. Since the modally weakest interpretation of Dependence is a Dualist one according to which it is supported by brute psychophysical laws linking experiences with internal states, and since *ceteris paribus* laws may be presumed to hold in nearby counterfactual situations, we may assume that the same applies to nearby counterfactual situations.

## 3. A Counterexample to Externalism

It may seem that Dependence does not pose a problem for Externalism. For instance, suppose that the Externalist explains sensory representation in terms of biological function. In the actual world, if two people undergo different levels of $\phi$ and $\psi$ activity, then they are in states that have the biological function of indicating different reflectance properties. In general, on Externalism, in the actual world, individuals who are in suitably different neural states are guaranteed to represent different properties, and so are guaranteed to have different experiences.

However, Dependence does create a problem for Externalism. But to see this we must go to across-world cases. Here I will describe one such case. It concerns an individual in the actual world and his counterpart in a nearby counterfactual situation. The structure of the case is as follows. The actual individual and his counterpart have different postreceptoral wiring as a result of differences in their evolutionary histories. In consequence, when they are presented with the same external property, they are put into different neural states and have different behavioral dispositions; yet the different neural states have, in their respective worlds, the biological function of indicating, and are optimally caused by, *the same* external property. I will argue on the basis of Dependence that the individuals involved have

different experiences. For those who are not convinced by the empirical argument, I will argue for the same conclusion on the basis of a plausible *a priori* principle concerning the link between experience and behavior. But, I will argue, all versions of Externalism deliver the mistaken verdict that the individuals involved have qualitatively identical experiences. Therefore, Externalism is false.

*Maxwell and Twin Maxwell*

To begin with, note that the response curves of the $\phi$ and $\psi$ opponent channels shown in Figure 1 are the result of countless small mutations and selection pressures. Selection pressures may determine some features of the response curves. For instance, they may determine that fruit and the background foliage normally produce quite different opponent channel states, so that we may easily pick out the fruit. But they do not determine the exact shapes of the response curves. There are many possible response curves such that if our opponent channels had had those response curves then we would have had or could have had the same fitness that we actually enjoy. So, selection pressures do not discriminate between the alternatives within this class. There is no adaptive story to be told about why the response curves have the exact shapes that they do. This means that it is just a matter of chance that they have the exact shapes that they do. The series of mutations which determined them might have been different; the mutations might not have been selected against and therefore might have stuck; and together these very small differences might have resulted in differences in the response curves.

Now let $W$ be the closest possible world in which the following two conditions hold. (I assume a unique closest world for simplicity.) First, owing to chance differences in our evolutionary history, the response curves of the $\phi$ and $\psi$ opponent channels are different from the actual response curves shown in Figure 1 in some particular way. (It does not matter how.) Consequently:

> **Different Processing and Behavior**: Because of evolved differences in their wiring, if an inhabitant of the actual world and an inhabitant of $W$ view physically identical objects under the same conditions, different levels of activation are set up across their postreceptoral $\phi$ and $\psi$ channels. Consequently, their behavioral dispositions, in particular their sorting dispositions, are slightly different (in ways described below).

Second, in $W$, although the response curves of human beings' *postreceptoral* $\phi$ and $\psi$ opponent channels are different from the actual ones shown in Figure 1, the response curves of the three *cone receptor types* remain unchanged. Consequently:

**Same Causes**: If an inhabitant of the actual world and an inhabitant of *W* view physically identical objects under the same conditions, they are put into different opponent channel states; however, those opponent channel states are caused under normal conditions by the same external reflectance properties.

The reflectance of an object is given by the proportion of incident light the object is disposed to reflect at each wavelength in the visible spectrum. A reflectance property is any property which can be defined in terms of reflectances.

By **normal conditions** I mean merely *statistically* normal conditions. Normal conditions involve normal lighting conditions (daylight or something comparable), the normal operation of the visual system, and so on. Later I will argue that the above also holds under optimal conditions, which is a more teleological notion.

Let us say that reflectance property *R* is **mapped onto** opponent channel state *O*, relative to species *S* and world *W*, iff under normal conditions *R* causes the tokening of *O* in the visual system of species *S* in world *W*. In the actual world, reflectance properties are mapped onto certain opponent channel states in humans; by Different Processing and Behavior, those same reflectance properties are mapped onto different opponent channel states in *W*.

Let us suppose that the two mappings do not preserve unitary-binary structure in the following sense. Reflectance properties which are, in the actual world, mapped onto binary opponent channel states are, in *W*, sometimes mapped onto unitary opponent channel states, and *vice versa*. (A binary opponent channel state is one that involves activation in both channels, while a unitary opponent channel state is one that involves activiation in only one channel.) It is because the two mappings do not preserve unitary-binary structure that we and our counterparts are behaviorally different, as we shall see. But let us suppose that the two mappings approximately preserve resemblance structure in the following sense. By and large, reflectance properties that are mapped onto similar (dissimilar) opponent channel states in the actual world are also mapped onto similar (dissimilar) opponent channel states in *W*.

I take it that Different Processing and Behavior, on the one hand, and Same Causes, on the other, are compatible. I define the $\phi$ and $\psi$ channels and their levels of activation in internal, neural terms. Anything can cause anything. So, there are possible worlds in which we are put into different levels of $\phi$ and $\psi$ activity, and in which our behavioral dispositions are consequently slightly different, but in which those opponent channel states are normally caused by the same external reflectance properties. I am stipulating that *W* is such a world.

It follows that one cannot dispute my claim that in *W* our opponent channel states and behavioral dispositions are different than they are in the actual world. This is a stipulated feature of the case.

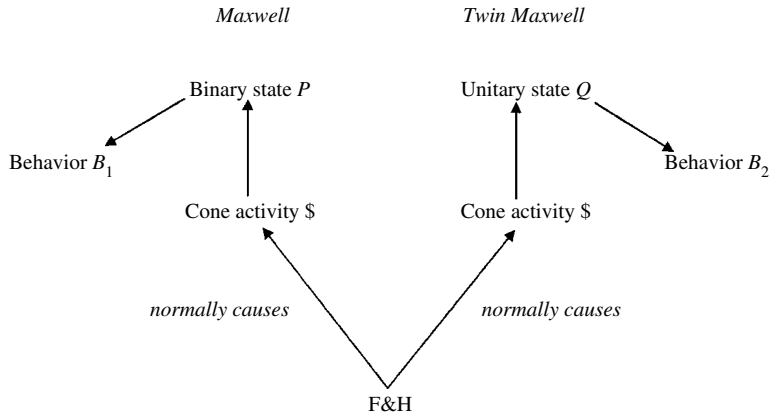
Given the basic physical facts of the two worlds as I have described them, are our color experiences in the actual world and in *W* the same or different? For concreteness, let us focus on a particular case: that of Maxwell and Twin Maxwell. To begin with, some background on what the Externalist will say about colors.

I have been speaking of the unitary-binary distinction as a distinction among *color experiences*. But, on Intentionalism, it is primarily a distinction among *colors*. For instance, a binary color experience is just one in which one is aware of a binary color. Typically, Externalists are Physicalists about the colors: for instance they identify colors with reflectance properties. So, they must explain the unitary-binary distinction among colors in physical terms. To account for the unitary-binary distinction among colors, Tye (Tye 2000; Tye and Bradley 2001) postulates four hue-magnitudes: reddishness, greenishness, yellowishness and bluishness. He identifies these with complicated reflectance properties F, G, H and I. (The details are not relevant here.) He says that binary colors—colors which are "perceptual mixtures"—are conjunctions of two of these properties while unitary colors are not such conjunctions. Call this objectivist, Physicalist account of the unitary-binary distinction among colors **the hue-magnitude account**.[3] I am a Primitivist about colors and the unitary-binary distinction, but let us assume, for purposes of argument, that some such physicalist account is correct. The argument I will be developing here still applies. For Tye combines this account with an Externalist account of the representation of hue-magnitudes. This is the source of the problem I am raising. As we shall presently see, the Externalist component of the view means that it delivers the wrong verdict about *W*.

Now back to Maxwell and Twin Maxwell. Suppose that Maxwell and Twin Maxwell view a square object with the reflectance property F&H (that is, on the hue-magnitude account, the color red-yellow) under normal conditions. (On a vastly oversimplified version of the hue-magnitude account, F&H is the property of normally reflecting light such that $L > M$ and $(L + M) > S$, where L, M and S are the long, medium and short components of the visible spectrum. But, again, the details do not matter here. By the way, I will sometimes write as if Maxwell and Twin Maxwell view the same square although this is not the case as they occupy different worlds.) By Same Causes, the receptor systems of Maxwell and Twin Maxwell are the same. So, in Maxwell and Twin Maxwell, F&H causes the same pattern of activation, \$, across the three cone types. However, by Different Processing, \$ leads to different opponent channel states further downstream. In Maxwell, it leads to equal positive levels of activity across the $\phi$ and $\psi$ channels. Call this opponent channel state *P*. In

Twin Maxwell, by contrast, $ leads to positive activity in the $\phi$ channel and neutral balance in the $\psi$ channel. Call this different opponent channel state $Q$. In short, while in the actual world F&H is mapped onto $P$, in $W$ it is mapped onto $Q$. Consequently, Maxwell and Twin Maxwell's behavioral dispositions are somewhat different.

Diagrammatically:




My argument against Externalism concerning this case now rests on two premises. The first premise is as follows.

**First Premise**: The correct verdict in the case of Maxwell and Twin Maxwell is Different Experiences: Maxwell and Twin Maxwell have different color experiences of the square.

I have two arguments for this premise. The first relies on C-Dependence. By C-Dependence, opponent channel states play a direct role in determining the unitary-binary character of color experience. $W$ is a very nearby world. So we may presume that the same holds in $W$. This is obviously the case on a Physicalist interpretation of C-Dependence: for instance on the view that color experiences are somehow constituted by opponent channel states. But it is also the case on the weakest, Dualistic interpretation of C-Dependence according to which it is supported by a contingent, fundamental law directly linking color experiences with opponent channel states. For as $W$ is a nearby world we may presume that the same fundamental laws hold there. (Everyone will grant that if we had so evolved that the average mass of humans had been different the law of gravitation would still have held. By analogy, if we had so evolved that the response curves were different, the relevant fundamental psychophysical law would still have held.) Now

Maxwell is in binary opponent channel state *P*, involving positive activity in both the $\phi$ and $\psi$ opponent channels, while Twin Maxwell is in unitary opponent channel state *Q*, involving positive activity in the $\phi$ opponent channel and neutral balance in the $\psi$ opponent channel. So, by C-Dependence, Maxwell has a binary experience of the square while Twin Maxwell has a unitary one. In particular, Maxwell has a red-yellow experience, while Twin Maxwell has a unitary red experience.

True, the binary opponent channel state *P* (in the actual world) and the unitary opponent channel state *Q* (in *W*) are normally caused by the same reflectance property, namely F&H, that is, the property of normally reflecting light such that $L > M$ and $(L + M) > S$. On Externalism, this reflectance property is necessarily identical with a binary color. But if opponent channel states play a **direct** role in determining the unitary-binary character of color experience, in accordance with C-Dependence, then this should be irrelevant. The first state should still result in a binary color experience and the second should still result in a unitary one. (For more here, see Pautz 2005.)

Even if a "Two-Factor Theory" is correct, the verdict of Different Experiences follows. In the case of Maxwell and Twin Maxwell, the external factor is the same, for both are in states that are normally caused by F&H. But the internal factor is different. Therefore, even on a Two-Factor Theory, the correct verdict is Different Experiences.

Now for my second, back-up argument for Different Experiences. By Different Processing and Behavior, Maxwell and Twin Maxwell not only undergo different levels of $\phi$ and $\psi$ activity; they also have different behavioral dispositions as a result. In the actual world and in *W*, the same reflectance properties are mapped onto different opponent channel states. By stipulation, the mappings approximately preserve resemblance structure. Reflectance properties that are mapped onto similar (dissimilar) opponent channel states in the actual world are also mapped onto similar (dissimilar) opponent channel states in *W*—only different ones. So, Maxwell and Twin Maxwell's sorting dispositions are more or less the same. But, by stipulation, they do not preserve unitary-binary structure. This means that, under rare circumstances, Maxwell and Twin Maxwell will exhibit different sorting behaviors.

Consider an example. In the actual world, F&H is mapped onto the binary opponent channel state *P*. Let *R* be a similar reflectance property. Let us suppose that in the actual world *R* is mapped onto opponent channel state, *P′*. Further, let us suppose that *P* and *P′* are similar in that both involve positive activation of both the $\phi$ and $\psi$ channels—just slightly different levels of activation. So, F&H and *R* objects normally produce in Maxwell quite similar opponent channel states. Consequently, he is to a certain high degree disposed to sort together an F&H object with an *R* object.

In *W*, F&H is mapped onto the unitary opponent channel state *Q* involving positive $\phi$ activity and no $\psi$ activity. What about *R*? Suppose that, in *W*, *R* is mapped onto binary state *Q′*, involving relatively high positive $\phi$ activity and a very slight degree of positive $\psi$ activity. So, whereas in Maxwell both F&H and *R* are mapped onto binary opponent channel states, in Twin Maxwell F&H is mapped onto a unitary opponent channel state while *R* is mapped onto a binary opponent channel state. In consequence, Twin Maxwell is "less disposed" than Maxwell to sort together F&H and *R* objects.

No doubt there will be other subtle behavioral differences between Maxwell and Twin Maxwell. For instance, the number of just-noticeable-differences between two stimuli may differ between them.

By stipulation, then, Maxwell and Twin Maxwell's behavioral dispositions differ somewhat. Now consider the following principle:

> **The Experience-Behavior Link**: If two actual or possible individuals have qualitatively identical color experiences, then they have the same color-related behavioral dispositions. If two actual or possible individuals have suitably different color experiences (for instance, they are not merely black-white spectrum inverted), then they have different color-related behavioral dispositions. Likewise for other kinds of experiences.

Some comments. First, paraplegics and the like are not counterexamples: for they nevertheless have dispositions to behave. Nor are individuals who have the same color experiences and yet behave differently because of different background mental states. Again, their dispositions to behave are similar, provided that they have the same other mental states. Second, there is a question about the modal status of the link between experience and behavior. Some might say that the Experience-Behavior Link must be restricted to individuals in this world or nearby worlds: for instance, those who hold that experiences are necessarily identical with brain states and that the connection between brain states and behavior is contingent. Others might say that it holds with respect to all possible or actual individuals: for instance defenders of certain forms of Functionalism. The present argument is neutral between these views. Whatever view one takes on the nature of experience, some form of the Experience-Behavior Link has great intuitive appeal. At least in this world and nearby worlds, there is a harmony between experience and behavior.

Twin Maxwell is not a swampman or a brain in a vat. It should be uncontroversial that both Maxwell and Twin Maxwell have color experiences. The only question is: do they have the same color experiences or different color experiences? By stipulation, their color-related dispositions

are different. Therefore, by the Experience-Behavior Link, they cannot have the same color experiences. For then their color-related dispositions would be the same. Can you imagine two normal individuals having the same color experiences, and yet totally different color-related behavioral dispositions? I find this inconceivable, but at the very least it is counterintuitive. Therefore, given their different color-related behavioral dispositions, Maxwell and Twin Maxwell must have different color experiences, if they have color experiences at all.

In short, by Different Processing and Behavior, Maxwell and Twin Maxwell not only undergo different levels of $\phi$ and $\psi$ activity; they also have different behavioral dispositions as a result. This makes it very implausible that the case of Maxwell and Twin Maxwell is a case of alternative neural realizations of the same color experience. On the contrary, the only reasonable verdict in this case is Different Experiences. So even if we forget about the opponent process theory and C-Dependence, we should accept Different Experiences.

This back-up argument for Different Experiences is relatively *a priori*. It does not take sustained empirical investigation to know that there are counterfactual-supporting links between color experience and behavior. So, while the title of the paper suggests that the argument is wholly empirical, the argument also has an *a priori* component.

By C-Dependence and the Experience-Behavior Link, the correct verdict is that, owing to the different response curves of their $\phi$ and $\psi$ opponent channels, Maxwell and Twin Maxwell regularly have somewhat different color experiences of the same objects under the same conditions. They exhibit a case of **Altered Spectrum**. This strikes me as non-negotiable. So far the argument ought to be uncontroversial. The crucial question is whether any of the versions of Externalism can deliver this verdict. I will now argue that they cannot. This is my second premise:

**Second Premise**: All versions of Externalism deliver the verdict of Same Experiences.

Recall that Externalist Intentionalism combines Intentionalism with the claim that the awareness relation is a wide physical relation between sentient creatures and properties outside the head. There are many such relations to choose from. Let the **S-role** be the functional role characteristic of experiences. Michael Tye (2003, 175–176) says that the S-role is being poised to influence the formation of appropriate beliefs and desires, while Fred Dretske (1995, 19) says that it is having the function of supplying information to a cognitive system for use in the control of behavior. Here now are some wide physical relations that the Externalist might identify the awareness relation with:

**The Optimal Cause Relation**: **x** is in a state that plays the S-role and that would be caused by the instantiation of property **y** were optimal conditions to obtain
**The Indication Relation**: **x** is in a state that plays the S-role and that has the implied indicator function of indicating the instantiation of property **y**
**The Asymmetric Dependence Relation**: **x** is in a state that plays the S-role and that asymmetrically depends on the instantiation of property **y**
**The Input-Output Relation**: **x** is in a state that plays the S-role and that under optimal conditions is caused by the instantiation of **y** and that in turn enables **x** to distinguish objects that have **y** from objects that do not and to engage in behavior appropriate to an object having **y**
**The Success Relation**: **x** is in a state that plays the S-role and that leads to successful action when and only when an object is present that has property **y**, where *successful action* is action that satisfies the subject's desires
**The Consumer Relation**: **x** is in a state, S, that plays the S-role and if S is to help the parts of **x** that use (or "consume") S to perform their functions in a historically normal way, then property **y** must be instantiated[4]

Externalism, then, comes in different versions. Each identifies the awareness relation with a different wide physical relation on the list. But all versions deliver the verdict of Same Experiences, for Maxwell and Twin Maxwell's wide relations to the outside world are the same in spite of the neurobiological and behavioral differences between them.

On Tye's Externalism, the awareness relation is in effect identical with the Optimal Cause Relation. (He adds a condition about asymmetric dependence which may be ignored.) This relation is defined in terms of optimal conditions. To show that Tye Externalism delivers the mistaken verdict of Same Experiences, I must first show that optimal conditions obtain in both Maxwell's situation and Twin Maxwell's situation. Call this the **Symmetry Thesis**.

Tye explains optimal conditions as follows:

> For evolved creatures, such conditions for vision involve the various components of the visual system operating as they were designed to do in the sort of external environment in which they were designed to operate. Here, there is no interference—no genetic abnormalities to throw things off, no peculiarities in the outside setting. Everything is as it should be. (Tye 2000, 138)

Consider Maxwell first. His wiring has not been interfered with in any way. There are no inverting lenses involved, or anything of the sort. His wiring is the product of natural selection. As he views the square, his visual system is operating as it was designed to do in the sort of external environment in which it was designed to operate. Moreover, there are no peculiarities in the outside setting. It is impossible to identify anything amiss.

Parallel remarks apply to Twin Maxwell. As he views the square, his wiring has not been interfered with in any way. There are no inverting lenses involved, as in Block's Inverted Earth example (Block 1990), or anything of the sort. Granted, his postreceptoral wiring is different from Maxwell's. Consequently, he is put into opponent channel state *Q* rather than *P*. However, by stipulation, Maxwell's postreceptoral wiring, like Maxwell's, is the product of natural selection. It is just as it should be. So, when he views the square, his visual system, no less than Maxwell's visual system, is operating as it was designed to do in the sort of external environment in which it was designed to operate. Moreover, there are no peculiarities in the outside setting. The lighting and other conditions are normal. As in the case of Maxwell, it is impossible to identify anything amiss.

Therefore, on the input side, everything is as it should be in both situations. The situations are perfectly symmetrical. Given Tye's explanation, it seems to follow that if optimal conditions obtain in the case of Maxwell, then they must obtain in the case of Twin Maxwell. But it might be suggested that optimal conditions is also a partly output-oriented notion (Tye 1995, 154–155). Maxwell and Twin Maxwell must also engage in adaptive and appropriate behavior. But this is also the case.

To see this, first consider a contrasting case, which is a variation on a case due to Colin McGinn (1989). Percy and Twin Percy view a square object. Percy is put into neural state *F*. Twin Percy's wiring is different, so he is put into different neural state *G*. As a result of being in *F*, on viewing a square object, Percy behaves in a way appropriate to squares. Therefore his behavior is adaptive and appropriate. But, as a result of being in *G*, Twin Percy behaves in a way appropriate to circles. In consequence, his behavior is maladaptive. Moreover, suppose that circle things also put Twin Percy into *G*. Then he sorts together squares and circles. He sorts objects into unnatural classes. Therefore his behavior is also inappropriate. Here the Tye Externalist might say with some justification that optimal conditions obtain for Percy but not Twin Percy. Because it reduces considerably his evolutionary fitness, it is difficult to imagine that Twin Percy's different wiring resulted from evolution. It must be the result of some kind of tampering or interference. But suppose that it is the result of evolution. It may still be said that optimal conditions do not obtain in the case of Twin Percy because his behavior is maladaptive and inappropriate. Then it is not the case that *F* and *G* both *optimally* track the property of being square. So, Tye Externalism has no commitment to the claim that both of their experiences

*represent* the property of being square, and hence no commitment to the claim that they have the same shape experience. In other words, in this case, Tye Externalism avoids the intuitively mistaken verdict of Same Experiences. The case of Percy and Twin Percy is not a counterexample to Tye Externalism. It is only a counterexample to simple, purely input-oriented forms of Externalism.

But the case of Maxwell and Twin Maxwell is completely different. The difference in wiring between them is the result of evolution, and it does not result in a difference between them in evolutionary fitness. In normal circumstances, their behavior is more or less the same and equally adaptive. We may suppose that they are equally able to discriminate objects. For instance, we may suppose that, in *W* as in the actual world, fruit and the background foliage produce very different opponent channel states, so that the fruit may easily be picked out. So, Maxwell and Twin Maxwell are able to negotiate the world equally well. Granted, under rare conditions, Maxwell and Twin Maxwell exhibit subtle sorting differences, as we have noted. But these subtle sorting differences do not affect their fitness. Unlike Percy and Twin Percy, they have **equal evolutionary fitness**.

Likewise, Maxwell and Twin Maxwell's behaviors are equally appropriate. By stipulation, in *W* as in the actual world, dissimilar reflectance properties are mapped onto dissimilar opponent channel states. In consequence, neither Maxwell nor Twin Maxwell sorts together objects with dissimilar reflectance properties. For instance, let UF be the reflectance property that the Tye Externalist would identify with the color unitary red. (On a grossly simplified version of the hue-magnitude account, UF is the property of normally reflecting light such that $L > M$ and $(L + M) = S$.) Neither Maxwell nor Twin Maxwell is disposed to group the F&H square with UF objects, for in both *W* and the actual world F&H and UF are mapped onto two quite different opponent channel states—just different ones in the different worlds.

Granted, as we have seen, because unitary-binary structure is not preserved, Maxwell and Twin Maxwell's sorting dispositions differ very slightly. Consider again the example used above: Twin Maxwell is less disposed than Maxwell to sort together the F&H square he is viewing with objects possessing the reflectance property *R* because whereas in Maxwell both F&H and *R* are mapped onto binary opponent channel states, in Twin Maxwell F&H is mapped onto a unitary opponent channel state while *R* is mapped onto a binary opponent channel state. But one cannot say on that account that either sorts object inappropriately. By virtue of his opponent channel states, Maxwell is disposed to sort objects into certain classes. The classes are fairly "unnatural", for the reflectances of the objects in the classes do not objectively resemble one another to a high degree. By virtue of his different opponent channel states, Twin Maxwell is disposed to sort objects into slightly different classes. But we may stipulate

that these classes are not less "natural" than the classes into which Maxwell is disposed to sort objects. Consequently, although Maxwell and Twin Maxwell's sorting dispositions differ, one cannot say that one has "appropriate" sorting dispositions, while the other has "inappropriate" sorting dispositions.

I believe that I have said enough to establish the Symmetry Thesis, but here are two arguments which support it further. First, a kind of "knowledge argument". Suppose that you have the concept of optimal conditions and are given a complete description of both situations in non-phenomenal terms. You are given the full story about the reflectances of objects, Maxwell and Twin Maxwell's behavioral dispositions, and so on. Suppose further that you do not know which world is the actual world, so that actual-world chauvinism cannot intrude. You consider the two worlds impartially. In all respects relevant to optimal conditions, the situations are perfectly symmetrical. Hence the only reasonable conclusion to draw is that optimal conditions obtain in both worlds. Suppose that you said that optimal conditions fail to obtain in one of the worlds. Which one? How could you possibly justify the claim that optimal conditions obtain in one of the worlds but not the other? There is **no symmetry breaker**. So, setting aside the bizarre view that whether optimal conditions obtain with respect to a situation is epistemically opaque, we must say that optimal conditions obtain in both situations.[5]

Here is a second argument for the Symmetry Thesis. Suppose that, in spite of my arguments, the Tye Externalist rejects the Symmetry Thesis and accepts instead the **Asymmetry Thesis**: optimal conditions obtain in the actual world but not in $W$. For instance, suppose that he says that optimal conditions obtain in the actual world but not in $W$ on the faulty grounds that $\boldsymbol{B_1}$ is uniquely appropriate or well-tuned to the objective pattern of similarities and dissimilarities among the reflectance properties of objects, where $\boldsymbol{B_1}$ is our actual set of sorting dispositions. (I use bold to indicate that the defender of the Asymmetry Thesis holds that $\boldsymbol{B_1}$ is uniquely appropriate.) Hence, optimal conditions fail to obtain in $W$ and, more generally, in any world in which our sorting dispositions are other than $\boldsymbol{B_1}$. This claim, it seems to me, has an extraordinary consequence.

Consider a time, $t$, in the history of the actual world before the evolution of the visual system in its current form. At $t$, there were many different possible sets of sorting dispositions we might have evolved which are such that if we had evolved any one of them, then we would have had or could have had the same evolutionary fitness that we actually enjoy. Within this class is $\boldsymbol{B_1}$, the magical set of sorting dispositions required, according to the Asymmetry Thesis, for optimal conditions to obtain. But also within the class is $B_2$, our sorting dispositions in $W$, and many others: $B_3$, $B_4$, $B_5$, and so on. Had we evolved any one of these, optimal conditions would not have obtained in the actual world. Selection pressures do not discriminate between the members of

this class. Now, at $t$, $\boldsymbol{B_1}$ was just one possible set of sorting dispositions among the others, and it was no more likely that we would evolve $\boldsymbol{B_1}$ than it was that we would evolve $B_2$ or $B_3$ or $B_4$. . . and so on. So if the Asymmetry Thesis is true, then it will have to be taken to be a fluke that we actually evolved $\boldsymbol{B_1}$: the magical set of sorting dispositions which, according to the Asymmetry Hypothesis, are perfectly tuned to the objective pattern of similarities and dissimilarities among the reflectance properties of objects, and which are required for optimal conditions to obtain in the actual world. We just happen to be extraordinarily lucky. Call this the **Grand Fluke Thesis**.

The Asymmetry Thesis, then, implies the Grand Fluke Thesis. But the Grand Fluke Thesis is very difficult to believe. The extraordinary may sometimes happen. (An example may be that the universe has, by pure chance, such a fine adjustment of physical parameters to allow for the evolution of life at all.) Nevertheless, the Grand Fluke Thesis is very difficult to swallow.

I conclude that the only reasonable view is that optimal conditions obtain in both the actual world and in $W$. There is, then, some slack between neurobiology and behavior, on the one hand, and optimality, on the other. In the case of Percy and Twin Percy, the difference in neurobiology and behavior is accompanied by a difference in optimality. Not so in the case of Maxwell and Twin Maxwell.

Therefore, in this case, Tye Externalism delivers the wrong verdict, Same Experiences. In Maxwell, the binary opponent channel state $P$ is caused, under *optimal* conditions, by the binary reflectance property F&H. Likewise, in Twin Maxwell, the unitary opponent channel state $Q$ is caused, under *optimal* conditions, by the binary reflectance F&H. I also assume that both $P$ and $Q$ play the S-role. Therefore, in spite of their different levels of $\phi$ and $\psi$ activity and behavioral dispositions, as they view the square, Maxwell and Twin Maxwell bear the Optimal Cause Relation to the same reflectance property, F&H. For each is in some state or other that plays the S-role and that is caused by F&H under optimal conditions. On Tye Externalism, the Optimal Cause Relation is necessarily identical with the awareness relation and F&H is necessarily identical with the color orange. Hence, in spite of their different levels of $\phi$ and $\psi$ activity and behavioral dispositions, Tye Externalism implies that both Maxwell and Twin Maxwell bear the awareness relation (the sensory representation relation) to the color orange. In spite of the differences between them, then, Tye Externalism implies that Maxwell and Twin Maxwell have the same color experience. In other words, in spite of the differences between them, Tye Externalism implies that Maxwell and Twin Maxwell's color experiences are the same in representational content. Therefore it implies that they are also the same in qualitative character. Tye Externalism, then, violates both C-Dependence and the Experience-Behavior Link.

In short, there is some slack between neurobiology and behavior, on the one hand, and optimality on the other. There are possible individuals who differ considerably in neurobiology and behavior, but who nevertheless bear the Optimal Cause Relation to the same external physical properties. Therefore, on Tye Externalism, there is some slack between neurobiology and behavior, on the one hand, and qualitative character, on the other. There are possible individuals who considerably differ in neurobiology and behavior, but who nevertheless have qualitatively identical experiences.

Fred Dretske's version of Externalism does not appeal to the notion of optimal conditions and we can see immediately that it delivers the mistaken verdict of Same Experiences. Dretske in effect identifies the awareness relation with the Indication Relation. This relation is defined in terms of the relation *state S has the implied function of indicating property y*. The relevant notion of 'function' is biological function. How do we determine what some state, $S_1$, has the implied biological function of indicating? Suppose that $S_1$ belongs to representational system $S$, and that representational system $S$ has the function of indicating the $F$ of a certain range of objects, where $F$ is a determinable property (color, location, form, or whatever). According to Dretske, to determine what $S_1$ has the biological function of indicating, we determine what the actual value of $F$ is when $S_1$ is tokened in system $S$ while $S$ is functioning as it was designed (by evolution) to function and in the kind of circumstances for which it was designed. If the actual value of $F$ under these conditions is $F_1$ (where $F_1$ is a determinate value of $F$), then $S_1$ has the biological function of indicating $F_1$ (Dretske 1995, 48–51).

Let us apply Dretske's test to Maxwell and Twin Maxwell. By stipulation, as they view the square, both Maxwell and Twin Maxwell's visual systems are functioning as they were designed by natural selection to function. And, in each case, the actual reflectance property of the square is F&H. In general, when our visual systems are functioning as they were designed to function in the actual world and we are in $P$, the viewed object has reflectance F&H. In general, when our counterparts' visual systems are functioning as they were designed to function in $W$ and they are in $Q$, the viewed object has reflectance property F&H. So, by Dretske's test, it immediately follows that both $P$ in the actual world and $Q$ in $W$ have the *implied biological function of indicating* F&H. Further, both $P$ and $Q$ play the S-role. Therefore, in spite of their different levels of $\phi$ and $\psi$ activity and behavioral dispositions, as they view the square, both Maxwell and Twin Maxwell bear the Indication Relation to F&H. On Dretske Externalism, the Indication Relation is the awareness relation and F&H is the color orange. Therefore, in spite of their different levels of $\phi$ and $\psi$ activity and behavioral dispositions, Dretske Externalism implies that both Maxwell and Twin Maxwell have an orange experience.

William Lycan (2002) prefers Externalist Intentionalism, at least for color experience. But, unlike Tye and Dretske, he does not specify the wide physical relation which he thinks makes for sensory representation, although he does express sympathy with causal and teleological relations (1996, 75). However, every version of Externalism on the list yields the verdict of Same Experiences. On viewing the square, both Maxwell and Twin Maxwell are in states that play the S-role and that asymmetrically depend on the instantiation of F&H. Both are in states that play the S-role and that under optimal conditions track F&H and that in turn enable them to discriminate F&H objects from non-F&H objects and that lead to other behavior appropriate to F&H objects. Both are in states that play the S-role and that lead to successful behavior only on the condition that F&H is instantiated. Both are in states that play the S-role and that are such that if their consumer devices are to perform their functions in a historically normal way, then F&H must be instantiated. In short, in spite of their different levels of $\phi$ and $\psi$ activity and behavioral dispositions, Maxwell and Twin Maxwell's wide relations to reflectance properties are preserved. Hence all versions of Externalism deliver the verdict that they have the same color experiences, thereby violating both C-Dependence and the Experience-Behavior Link.

Consider the following parallel. The states of a mercury thermometer and the corresponding states of a thermoelectric thermometer, though mechanically different, are representationally the same, since they optimally track, and have the function of indicating, the same temperatures. Or again, in different languages, different sounds can mean the same thing. Likewise, on Externalism, the corresponding states of Maxwell and Twin Maxwell's visual systems—for instance, *P* and *Q*—though neurally different, are representationally the same, because they optimally track, and have the biological function of indicating, the same reflectance properties. The neural content-carriers are counterfactually altered, but the externally-determined content carried is held constant. So Externalism delivers the mistaken verdict of Same Experiences.

Of course, Externalism alone does not entail the mistaken verdict of Same Experiences. All Externalism says is that two possible individuals might be the same in neurobiology, and yet differ phenomenally. It does not automatically entail that two possible individuals might differ in neurobiology and behavior, and yet be phenomenally the same. But all versions of Externalist Intentionalism that I know of have this consequence as well. In particular, they deliver the mistaken verdict of Same Experiences in the case of Maxwell and Twin Maxwell. For they explicate representational content and hence qualitative character in terms of wide physical relations; and, in spite of the neurobiological and behavioral differences between them, Maxwell and Twin Maxwell bear the same relevant wide physical relations to the outside world.

This concludes my argument for the First Premise and the Second Premise. From these two premises, it follows that all the versions of Externalism on the list are false.

Note that the argument does not rely on the bare intuition that a certain kind of anti-Externalist "inverted spectrum scenario" is possible (Shoemaker 1994), or on the intuition that internal factors play a role in determining qualitative character (in my view, we have no such intuition). Rather, I argue for C-Dependence on empirical grounds. Then I describe the case of Maxwell and Twin Maxwell in non-phenomenal terms. And I use C-Dependence as well as the Experience-Behavior Link to argue that the case constitutes a counterexample to the relevant versions of Externalism. No appeal to intuition is required.

As I noted in the previous section, Dependence is not limited to color vision. In other sensory modalities, the empirical case for the direct role of internal factors in determining qualitative character is perhaps even stronger. Therefore, counterexamples of the same kind could be multiplied. They have the same structure as the Maxwell and Twin Maxwell case, but they involve different sensory modalities. For instance, consider the following case. Two individuals inhabit different possible worlds and have different evolutionary histories. They are in neural states involving significantly different neuronal discharge frequencies in the cortical areas relevant to pain intensity, and in consequence have significantly different pain-related behavioral dispositions; but these different states are optimally caused by, and have the biological function of indicating, the very same type of bodily disturbance in the leg (Pautz 2005). The correct verdict is Different Experiences but Tye Externalism and Dretske Externalism deliver the mistaken verdict of Same Experiences. So if one finds a defect in example of Maxwell and Twin Maxwell, then it seems that there is bound to be another example which is free from the defect.

The general moral is that the Intentionalist cannot identify the awareness relation with a wide physical relation, for the combination of Intentionalism and a wide account of the awareness relation delivers the mistaken verdict of Same Experiences in such examples. Externalist Intentionalism must be rejected, because it fails to accommodate the empirically-determined role of internal factors in shaping qualitative character.

## 4. Objections and Replies

The argument is simple: it has two premises. Given the stipulated neurobiological and behavioral differences between Maxwell and Twin Maxwell, I take it that the First Premise will be uncontroversial: the correct verdict is Different Experiences. I will consider two objections to the Second Premise: that all the versions of Externalism deliver the mistaken verdict of Same

Experiences. I put each objection in the mouth of an imaginary Externalist objector. (For replies to other objections, see Pautz 2005.)

*Objection*

A simple version of Externalism might deliver the mistaken version of Same Experiences. But a more sophisticated version delivers the correct verdict of Different Experiences, as follows. Opponent channel states represent compositionally. For instance, $P$ is a complex representational state built up in a conjunctive fashion from two simpler representational states: *undergoing positive* $\phi$ activity and *undergoing positive* $\psi$ activity. What each conjunct represents is determined by what hue-magnitude value it is optimally caused by. What the complex representational state $P$ represents is then determined, compositionally, by what each conjunct represents. Likewise for opponent channel states generally. Once we see that opponent channel states represent compositionally, we can see that $P$ and $Q$ differ in what they represent.

*Reply*

I do not see how the appeal to compositionality helps. To explain why, I must be more precise than I have been so far. I have pretended that Maxwell is aware of the determinable color *orange*, which the Tye Externalist identifies with F&H. In fact, Twin Maxwell is aware of a determinate shade of orange. Let us suppose that he is aware of balanced orange, that is, the property of being 50% reddish and 50% yellowish, or 50R50Y for short. The defender of the hue-magnitude account will say that values of the physical magnitudes F, G, H and I also stand in relations of proportion, and he will identify the property of being 50R50Y with the extradermal physical property of being 50F50H (Byrne and Hilbert 2003). I believe that the hue-magnitude account is mistaken (Pautz 2003), but let us suppose that it is right for the sake of argument.

Presumably, a compositional version of Tye Externalism would go as follows. $P$ is the conjunction of a state of the $\phi$ channel and a state of the $\psi$ channel. In particular, it is the conjunction of *undergoing relative* $\phi$ *activity of 0.5* and *undergoing relative* $\psi$ *activity of 0.5*. (Here I ignore complications about how to measure opponent channel activity.) In the actual world, *undergoing relative* $\phi$ *activity of 0.5* is optimally caused by, and thereby represents, the property of being 50F. Likewise, in the actual world, *undergoing relative* $\psi$ *activity of 0.5* is optimally caused by, and thereby represents, the property of being 50H. In this way, on the compositional view, the complex representational state $P$ represents the complex physical property of being 50F50H. Again, I think that this is mistaken. But, for purposes of argument, let us assume that, in both the actual world and in $W$, individual opponent channel states track values of hue-magnitudes.

Now consider *Q* in the counterfactual situation *W*. Like *P*, *Q* is the conjunction of a state of the $\phi$ channel and a state of the $\psi$ channel. In particular, it is the conjunction of *undergoing relative $\phi$ activity of 1.0* and *undergoing no $\psi$ activity*. By stipulation, in *W*, *undergoing relative $\phi$ activity of 1.0* is optimally caused by the property of being 50F. Therefore, on Tye Externalism, it represents the property of being 50F. By stipulation, in *W*, *undergoing no $\psi$ activity* is optimally caused by the property of being 50H. Therefore, on Tye Externalism, it represents being 50H. Hence, on the compositional version of Tye Externalism, in *W*, the complex representational *Q* represents the complex property of being 50F50H, just as *P* does in the actual world.

Even on the compositional view, then, *P* (in the actual world) and *Q* (in *W*) represent the same extradermal physical property, namely being 50F50H. They just do so *via* different levels of $\phi$ and $\psi$ activity. Analogy: in differently calibrated thermometers, different levels of mercury might represent the same temperature. Maxwell and Twin Maxwell are like differently calibrated thermometers. Therefore, even if we take compositionality into account, Tye Externalism delivers the mistaken verdict of Same Experiences.

Moreover, in any case, the compositional response cannot provide a general response to the argument. There are counterexamples to Externalism with the same structure as the example of Maxwell and Twin Maxwell, but involving other sensory modalities, for instance pain or taste (Pautz 2005). The neural states underlying pain intensity or taste are not complex representational states built up from simpler representational states. Therefore, even if the compositionality response worked in the case of Maxwell and Twin Maxwell, it is unavailable in other examples. (For further discussion, see Chalmers 2005.)

*Objection*

The First Premise is true. The correct verdict is Different Experiences. The central forms of Externalism, for instance Tye Externalism and Dretske Externalism, deliver the mistaken verdict of Same Experiences. Therefore, the example of Maxwell and Twin Maxwell shows that the central forms of Externalism are mistaken. But there is a more complicated version of Externalism that delivers the correct verdict of Different Experiences.

Consider belief. What we believe is arguably determined by two factors: what causes our beliefs, and their internally-determined functional roles. Consider, for example, the belief that water is wet and the belief that $H_2O$ is wet. Here the external factor is the same. But the two beliefs are different, because they have different internally-determined functional roles: they interact differently with other beliefs and with behavior. Consider now the belief that water is wet and our counterparts' belief that $H_2O$ is wet on Putnam's Twin Earth. The internal factor is the same. These beliefs interact with other beliefs and mental states in the same ways to produce behavior.

Yet the beliefs are different (at least in a sense), because the external factor is different: the first belief is caused by and is about $H_2O$ while the second is caused by and is about XYZ.

The Intentionalist might give a similar two-factor theory of the awareness relation. In particular, the Intentionalist might identify the two-place awareness relation with a two-place physical relation of the following form:

> **The Two-Factor Relation**: **x** is in some internal state S and S has internal property I and external property E and $f(<I, E>) = \mathbf{y}$

Here I is an "internal property" of a state concerning what other internal states and behaviors it is apt to cause while E is an "external property" of an internal state concerning what it is apt to be caused by in the outside world; and *f* is a function from pairs of internal properties and external properties onto Q-properties.

The Two-Factor Relation is wide. Suppose that two individuals, Bob and Nob, are in the same internal states, and that the internal states are apt to cause the same behaviors. Then the states have the same internal properties. Suppose, however, that there are evolved differences between their receptor systems, so that the internal states are *optimally* caused by different external properties. (Note that this case is importantly different from Block's (1990) Inverted Earth case. In that case, the difference in wiring is due to the insertion of inverting lenses, so that optimal conditions do not obtain. By contrast, in this case, the difference is the result of evolution.) It seems possible that there should be such a case. On a two-factor theory, because of the difference in the external factor, Bob and Nob bear the Two Factor Relation and hence the awareness relation to different properties. Otherwise the external factor drops out as irrelevant and we have a one-factor theory, with sensory awareness being fully determined by the internal factor. Consequently, a two-factor theory implies that Bob and Nob have different experiences, even though they are postreceptoral neural duplicates and also behavioral duplicates. (Note that Tye Externalism also has this consequence, contrary to his claim (2003, 174) that his theory accommodates Internalism for creatures as sophisticated as humans.) Therefore the view that the awareness is identical with the Two-Factor Relation is an Externalist view.

Nevertheless, Two-Factor Externalism delivers the right verdict, Different Experiences, in the case of Maxwell and Twin Maxwell and in other such cases. In the case of Maxwell and Twin Maxwell, the external factor is the same but the internal factor is different. On the Two-Factor Theory, the internal factor matters as well. In consequence, Maxwell and Twin Maxwell bear the Two-Factor Relation to different properties, and hence have different experiences.

*Reply*

I have three objections to Two-Factor Externalism.

First, what is described is not a theory, but a theory schema. To get a full-fledged theory, we must replace '*f*' with a rule for going from pairs of an internal property and an external property to Q-properties. In addition, the rule must be a simple one. The reason I say this is that if the Two-Factor Theory is correct, then the semantic value of 'x is aware of y' in sentences like 'Bob is aware of the color orange' is the Two-Factor Relation. But if the Two-Factor Relation is not codifiable in a fairly simple general rule—if it is not an at least somewhat "natural" relation—then it is difficult to see how it might be the semantic value of 'x is aware of y'. (This is supported by Lewis's theory of content in terms of use plus eligibility (1984), but I think that it is plausible independently of that theory.)

But what is the general rule for calculating what Q-properties an individual is aware of from the internal properties and the external properties of his inner states, analogous to the rule for calculating a resultant force from two component forces? The Externalists we have been examining provide a general rule in terms of the external factor alone. For instance, the Tye Externalist simply says that x is aware of y iff x is in a state that plays the S-role and that optimally tracks y. But the Two-Factor Theory cannot say anything so simple. For instance, in the case of Maxwell and Twin Maxwell, *P* and *Q* optimally track the same external property, namely F&H. But the Two-Factor Externalist will say, rightly in my opinion, that they are aware of different colors, because of the difference in the internal factor. But what is the general rule which takes us from the combination of the internal properties and the external properties of *P* and *Q* to what colors Maxwell and Twin Maxwell are aware of? Or again, suppose that an individual is in a neural state which involves a certain rate of neural firing and which optimally tracks bodily damage of a certain extent. What is the general rule for calculating resultant pain? Call this the No Rule Problem.[6] (Hilbert and Kalderon 2000 appear to endorse a two-factor theory in the sense at issue, but they do not address this problem.)

Second, Two-Factor Externalism delivers the verdict of Different Experiences in the case of Bob and Nob. But Bob and Nob have the same behavioral dispositions. Therefore Two-Factor Externalism violates the Experience-Behavior Link.

Third, in a Putnam Twin Earth case, intuition supports the claim that the correct verdict is Different Beliefs because of the difference in the external factor. But, in the case of Bob and Nob, there is no intuitive or theoretical reason to accept the claim that the correct verdict is Different Experiences because of the difference in the external factor. Thus, while there might be intuitive reason to accept a two-factor theory of belief (or at least those beliefs which are susceptible to Twin Earth examples), there is no intuitive or theoretical reason to accept an analogous two-factor theory of qualitative

character. On the contrary, there is intuitive reason to reject such a theory: it violates the Experience-Behavior Link.

In summary, the case of Maxwell and Twin Maxwell, and other cases of the same kind, show that the Intentionalist cannot identify the awareness relation with any of the wide physical relations in the list featured in §3. For the combination of Intentionalism and the claim that the awareness relation is identical with one of these wide physical relations delivers the mistaken verdict of Same Experiences in such cases. The three considerations presented here show that the Intentionalist also cannot identify the awareness relation with a wide, "two-factor" relation.

## 5. Further Issues

It may be thought that the moral is to give up Intentionalism. This is not my own view. There is nothing in what I have said which forces us to reject standard Intentionalism. In fact, I wholeheartedly agree with the slogan "phenomenology ain't in the head" (Tye 1995, 151). In accordance with the transparency observation, I hold that the qualitative character of color experience is determined by what colors we sensorily represent things as having "out there". Colors, whatever they are, are not properties of our own brains or experiences. If they are properties of anything at all, they are properties of external objects. Yet I insist that the correct verdict in the case of Maxwell and Twin Maxwell is Different Experiences. Therefore I combine these two claims. I hold that, owing to the internal differences between them, Maxwell represents the square as red-yellow, while Twin Maxwell represents it as unitary red. They bear the awareness relation to different colors. I take a similar view of the other sensory modalities. In short, I believe that, while phenomenology is not in the head, it depends on what happens in the head. It may seem mysterious that differences inside the head can result in differences in what properties we sensorily represent outside the head, but there is nothing incoherent about it. Further, there is a very strong motivation for both Intentionalism and Different Experiences. Intentionalism is supported by the transparency observation and other considerations. And Different Experiences is supported by Dependence and the Experience-Behavior Link.

However, the combination of Intentionalism and Different Experiences raises difficult issues about the awareness relation and the Q-properties. As I noted in §1, I believe that this combination of views leads to Primitivism about both the awareness relation and the Q-properties.

In effect, here I have carried out the first step in the argument from Intentionalism and Different Experiences to Primitivism about the awareness relation. In particular, I have shown that the conjunction of Intentionalism and Different Experiences means that the awareness relation is not reducible to a **wide physical relation**. By Different Experiences, as a

result of the internal differences between them, Maxwell and Twin Maxwell have different color experiences. By Intentionalism, they bear the awareness relation to different colors or color-like properties. But they bear all the same relevant wide physical relations to the same (physical) properties. Hence, given Intentionalism and Different Experiences, the awareness relation must be distinct from any such wide physical relation.

The second step in the argument will require examining **narrow physical relations**, where a relation is narrow iff, necessarily, neurobiological duplicates living under the same laws bear the relation to the same items. Maybe there is a narrow physical relation, *R*, such that Maxwell and Twin Maxwell bear *R* to different properties, owing to the internal differences between them. In general, maybe there is a narrow physical relation that tracks our neurobiology and behavioral dispositions. (A crude example is the output-oriented relation: *the hypothesis that* ***x*** *is aware of* ***y*** *is part of the best total explanation of* ***x****'s internally-determined behavioral dispositions*.) Then a version of Reductive Intentionalism according to which the awareness relation is identical with the narrow physical relation *R* would deliver the correct verdict of Different Experiences in cases like that of Maxwell and Twin Maxwell. But elsewhere I argue that the awareness relation cannot be identified with a narrow physical relation on the grounds that no narrow physical relation has the right properties (Pautz 2004, forthcoming a). In this way, Intentionalism and Different Experiences lead to Primitivism about the awareness relation. What properties we bear the awareness relation to depends on the internal physical state of the subject; but the awareness relation is not itself physical. Many think that Intentionalism helps with the mind-body problem. In my view, the opposite is true. Our most developed reductive accounts of intentional relations in general are wide. But, given the empirically-determined role of internal factors in shaping experience, such theories fail to apply to the sensory representation relation (the awareness relation). The sensory representation relation cannot be identified with a wide physical relation. And, once wide physical relations are ruled out, there remain no physical relations which are plausible candidates for identification with the sensory representation relation.

Here I have set aside the Q-properties. But I also believe that the combination of Intentionalism and Different Experiences leads to Primitivism about the colors and other Q-properties. Consider, for example, the colors. Of course, one might combine together standard Intentionalism, Different Experiences, and reflectance Physicalism about colors. This position is consistent. If one accepts this combination of views, then one will have to say that Maxwell and Twin Maxwell sensorily represent the square as having different reflectance properties, and thereby have different color experiences, as a result of the internal differences between them. In this way, one might accommodate the correct verdict of Different Experiences. This leads to an embarrassing question. Who represents the true reflectance

property of the square? Let us suppose that the defender of this combination of views takes the chauvinistic line that Maxwell gets it right and Twin Maxwell gets it wrong. In particular, Maxwell correctly sensorily represents the square as F&H while Twin Maxwell incorrectly sensorily represents it as UF. (Recall that UF is the reflectance property with which the Physicalist would identify the color unitary red.) In general, we in the actual world get it right, while our counterparts in *W* get it wrong.

While this is a consistent position, it faces a number of problems. First, once we accept standard Intentionalism and Different Experiences, I don't see the motivation for reductionism about colors, especially once we are convinced that this combination of views means that reductionism about the awareness relation fails (Pautz 2003). Second, there is a problem about brute representational facts. There is no explanation of why one set of internal factors issues in the sensory representation of one reflectance property rather than another. On this combination of views, necessarily, if x is in opponent channel state *P* and has behavioral dispositions $B_1$, then x bears the sensory representation relation to reflectance property F&H; necessarily, if x is in different opponent channel state *Q* and has different behavioral dispositions $B_2$, then x bears the sensory representation relation to reflectance property UF; and so on. But, as far as I know, such conditionals may not be derived from any theory of sensory representation or, more generally, from any more basic modal truths. Therefore, they must be taken to be brute modal facts. Third, there is the problem of coincidence. Since selection pressures do not determine all the details of our wiring, our wiring could easily have been as it is in *W*. The spectrum could easily have been shifted from the way it actually is. So, on this combination of views, it could have easily been that our color experiences were generally non-veridical. In other words, it is just a fluke that things look to have the colors (on this view, the reflectance properties) that they actually do have. We are extremely lucky. My main objection to this claim is simply that it is difficult to accept such a grand coincidence. But it also has the apparent corollary that our belief that things have the colors they look to have cannot constitute knowledge because it is true by accident. Third, there is the problem of structure: colors stand in certain resemblance relations and are unitary or binary, but arguably the corresponding reflectance properties do not stand in the requisite resemblance relations and are not unitary or binary (Hardin 1988).

If we accept Primitivism about colors, we may not be able to avoid the problem of brute representational facts, but we can avoid the other problems. Once we accept Primitivism, we are free to accept a response-dependent version of Primitivism according to which an object instantiates a primitive color C iff it normally produces experiences of primitive C (McGinn 1996). In the actual world, the square normally produces experiences of primitive red-yellow. In *W*, the square normally

produces experiences of unitary red. Therefore, on Response-Dependent Primitivism, the square instantiates primitive red-yellow in the actual world and it instantiates primitive unitary red in *W*. In consequence, on this view, both Maxwell and Twin Maxwell get it right. Thus Response-Dependent Primitivism avoids the problem of coincidence and the resulting threat to the claim that we know the colors of things. Alternatively, once we accept Primitivism, we might opt for Eliminative Primitivism. On the best version of this view, color properties exist but nothing—including our own color experiences—instantiates them. Colors live only in the intentional contents of our experiences. On this view, both Maxwell and Twin Maxwell get it wrong. Again, this view avoids objectionable coincidences. (For arguments against McGinn's response-dependent version of Primitivism and for Eliminativism, see Pautz forthcoming b.) Finally, Primitivism of any variety avoids the problem of structure. Once we accept Primitivism, we can credit colors with whatever structure they appear to have. For these reasons, once we accept standard Intentionalism and Different Experiences, it seems to me that Response-Dependent Primitivism or Eliminative Primitivism about colors is superior to reflectance Physicalism.[7]

## Notes

[1] A version of this paper was read at the 1998 New England Undergraduate Philosophy Conference. I would like to thank the audience there for helpful comments. I would also like to thank members of the Centre for Consciousness Reading group at the Australian National University for helpful discussion of a more recent version. Thanks to David Barnett, Paul Boghossian, Alex Byrne, Robert C. Coghill, Tyler Doggett, C. L. Hardin, Stephen Schiffer, Susanna Siegel, Daniel Stoljar, Joseph Levine, Brian McLaughlin, Joseph Owens, David Wall and two anonymous referees. I am especially indebted to Ned Block, David Chalmers, and Michael Tye.

[2] We might call the simple form of Intentionalism I will work with here the **Property-Complex Theory**. See Bealer (1982), Dretske (1995, 101–102; 1999, 163–164; 2003, 70–74), McGinn (1999, 319–323) and especially Johnston (2004). Some may say that the Property-Complex Theory is not a form of Intentionalism. I take this to be a verbal issue. I use 'Intentionalism' broadly so that the Property Complex Theory counts as a form of it.

It is more usual for Intentionalists to identify properties of the form *having an experience with qualitative character Q* with relations to propositions rather than properties. For instance, according to **Russellian Propositionalism** they are relations to existentially quantified, general **Russellian propositions** of the form *there is an object at place p with properties* $P_1$, $P_2$, *and* $P_3$. These propositions are taken to be "structures" exclusively built up from the properties attributed (Tye 2000, 48). We might call the special, propositional attitude relation we bear to propositions in experience **sensorily entertaining**. When one sensorily entertains a proposition into which *P* enters, whether the proposition is true or false, one is directly aware of *P*—the property *P* is directly present to one's consciousness. In hallucination, the proposition is false and the property contained in the proposition is not instantiated in the environment. Tye (2000, 48) speaks of being directly aware of properties in experience, even hallucinatory experience; but his view seems closer to Russellian Propositionalism. Here I will argue that the awareness relation is not reducible to a wide physical relation. The argument could be used, *mutatis*

*mutandis*, to show that the same is true of the propositional attitude relation *X sensorily entertains proposition Y*.

Both complex properties and general Russellian propositions are structures built up in some non-mereological way exclusively from properties. I believe that the Property-Complex Theory and Russellian Intentionalism face a version of the Many-Property Problem arising from the nature of these "structures". In my view, a more complicated form of Intentionalism is required. Again, I work with the Property-Complex Theory only for simplicity.

Forms of Intentionalism vary along another dimension: **standard** (for instance, Tye 2000) and **non-standard** (Shoemaker 2005). I accept standard Intentionalism (see note 7). In the text I will use 'Intentionalism' to mean standard Intentionalism.

[3] Byrne and Hilbert (2003) defend a more "subjectivist" version of the hue-magnitude account. According to their version, 'orange is a perceptual mixture of red and yellow' is true iff everything that *looks* orange *looks* reddish and yellowish, where the hue-magnitudes reddishness and yellowishness are identical with the physical magnitudes F and H. By the way, I implicitly assume that reddishness, greenishness, yellowishness and bluishness are hue-magnitudes. This is not accurate. The property of being reddish, for instance, is a determinable color property. By contrast, a magnitude is a set of properties, together with a ratio-scale (Byrne and Hilbert 2003). But such complications will not play a role here and may be ignored.

[4] For the Optimal Cause Relation, see Tye 1995 and Tye 2000. For the Indication Relation, see Dretske 1995. For asymmetric dependence, see Fodor 1990. For the view that both input factors and output factors are relevant to sensory representation, see Armstrong 1999, chapter 12 and Davies 1993, 239–242. For success semantics, see Papineau 1993. On consumer devices, see Millikan 1989, especially 289–290. Fodor and Papineau do not apply their ideas on representation to sensory representation.

[5] Perhaps it will be objected that optimal conditions fail to obtain in Twin Maxwell's situation because his color beliefs and utterances are false. But the Tye Externalist seems committed to the claim that Twin Maxwell has a true color belief with the content *there is an F&H present*, because his color belief is caused under optimal conditions by an F&H object. Nor can the Tye Externalist say that optimal conditions fail to obtain in Twin Maxwell's situation because he incorrectly represents the square as UF (that is, on Physicalism, unitary red). For it would be circular to explain sensory representation in terms of optimal conditions and then explain optimal conditions in terms of the correctness of our sensory representations. Further, we may imagine similar cases concerning pain or taste in which it is indisputable that the Symmetry Thesis holds. Therefore, even if the case of Maxwell and Twin Maxwell does not constitute a counterexample, these other cases certainly do. For a discussion of these points, see Pautz 2005, 2006 and Chalmers 2005.

[6] Jonathan Schaffer suggested something like the following answer to the No Rule Problem on behalf of the defender of Two-Factor Externalism. Suppose we accept an unusual view of the Q-properties according to which they are ordered pairs consisting of (i) an internal property and (ii) a property outside the head tracked by one of our sensory systems. Then the Two-Factor Theorist may solve the No Rule Problem by saying that x is aware of y iff x is in a state S that has some internal property I and that normally tracks some external property *A* and $y = <I, A>$. So, for instance, it follows from this view that, as they view the square, Maxwell is aware of $<I_1, F\&H>$ while Twin Maxwell is aware of $<I_2, F\&H>$, where $I_1$ is the internal property of *P* and $I_2$ is the internal property of *Q*. Consequently, this view delivers the correct verdict of Different Experiences in the case of Maxwell and Twin Maxwell. I agree that the account of the Q-properties as ordered pairs allows a solution to the No Rule Problem. But it is at odds with the transparency observation and is problematic in other ways.

[7] I have suggested that we combine Different Experiences with standard Intentionalism. For instance, we should say that, as a result of the internal differences between them, Maxwell and Twin Maxwell experience the square as having different *colors*. Sydney Shoemaker (2006) would disagree. He would say that they experience the square as having the same color, namely

the color orange. In his scheme, they have different color experiences, because they experience this same color under different "qualitative characters". He describes this view as a non-standard form of Intentionalism. It seems to me that this view is at odds with English usage. Given that Maxwell has a red-yellow experience while Twin Maxwell has a unitary red one, ordinary people would describe them as experiencing different *colors*. They would say the square looks orange to Maxwell and unitary red to Twin Maxwell. Further, Non-Standard Intentionalism is at odds with the transparency observation on which it is meant to be based. For that is the observation that when we focus on the character of a color experience we find only the *color* of which we are directly aware, and that the character of color experience is fully determined by the *color* of which we are directly aware. We do not find a color and a "qualitative character" under which the color is presented. For other criticisms, see Tye 2000.

Further, I believe that there is no strong argument for Shoemaker's non-standard version of Intentionalism. His argument seems to be this. (Pr1) If standard Intentionalism is true, then if two people are spectrum inverted, then at least one must be systematically misperceiving the colors of things (Shoemaker 2003, 255). (Pr2) It is not the case that, if two people are spectrum inverted, then at least one must be systematically misperceiving the colors of things. (C) Therefore, standard Intentionalism is mistaken. But the availability of McGinn's response-dependent version of Primitivism about colors shows that (Pr1) is false. On this view, as we have seen, even if standard Intentionalism is correct, and Maxwell and Twin Maxwell normally experience the square as having different colors, both of them get it right. Furthermore, in my opinion, Shoemaker does not take seriously enough the combination of standard Intentionalism with Eliminativism about colors. This is my own view. According to it, Maxwell and Twin Maxwell experience the square as having different color properties, neither of which the square has. Against this view, or a similar view, which he calls 'Figurative Projectivism', Shoemaker objects that it implies that "our perceptual experience is incurably infected with illusion" (1994, 26). But I do not see what is wrong with incurable illusion. In any case, in my opinion, response-dependence or incurable illusion is preferable to Shoemaker non-standard version of Intentionalism, in view of its problems.

## References

Armstrong, David. (1999) *The Mind-Body Problem*, Boulder: Westview Press.

Bealer, George. (1982) *Quality and Concept*, Oxford: Clarendon Press.

Billock, V., Gleason, G. and Tsou, B. (2001) "Perception of Forbidden Colors in Retinally Stabilized Equiluminant Images: An Indication of Softwire Cortical Opponency?," *Journal of the Optical Society of America* 18: 2398–2403.

Block, Ned. (1990) "Inverted Earth," in J. Tomberlin, ed., *Philosophical Perspectives* Vol. 4: Atascadero, CA: Ridgeview Publishing Company.

Boynton, R., Schafer, W. and Neun, M. (1964) "Hue-wavelength Relation Measured by Color-naming Method for Three Retina Locations," *Vision Research* 29: 115–128.

Bradley, Peter and Michael Tye. (2001) "Of Colors, Kestrels, Caterpillars, and Leaves," *Journal of Philosophy* 98: 469–487.

Byrne, Alex and Hilbert, David. (2003) "Color Realism and Color Science," *Behavioral and Brain Sciences* 26: 3–21.

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

Chalmers, David. (2005) "Representationalism Showdown", [http://fragments.consc.net/djc/2005/09/representationa.html].

Chalmers, David. (2006) "Perception and the Fall from Eden," in T. Szabo Gendler and J. Hawthorne, eds., *Perceptual Experience*, Oxford: Oxford University Press.

Coghill, R., Sang, C., Maisog, J., and Iadarola, M. (1999) "Pain Intensity Processing within the Human Brain: A Bilateral, Distributed Mechanism," *Journal of Neurophysiology* 82: 1934–1943.

Coghill, R., McHaffie, J. and Yen, Y-F. (2003) "Neural Correlates of Interindividual Differences in the Subjective Experience of Pain," *Proceedings of the National Academy of Science* 100: 8538–8542.

Crane, H. and Piantanida, T. P. (1983) "On Seeing Reddish Green and Yellowish Blue," *Science* 221: 1078–1080.

Davies, Martin. (1993) "Aims and Claims of Externalist Arguments," in E. Villanueva, ed., *Philosophical Issues* 4: Atascadero, CA: Ridgeview Publishing.

De Valois, Russell. (2004) "Neural Coding of Color," in Leo M. Chalupa and John S. Werner, eds., *The Visual Neurosciences*, *Volume 2*: Cambridge: MIT Press.

De Valois, Russell and Karen De Valois (1993) "A Multi-Stage Model of Color Vision," *Vision Research* 33: 1053–1065.

Dretske, Fred. (1995) *Naturalizing the Mind*, Cambridge: MIT Press.

Dretske, Fred. (1999) "The Mind's Awareness of Itself," *Philosophical Studies* 95: 103–124. Reprinted in *Knowledge, Perception, and Belief: Selected Essays* (Cambridge: Cambridge University Press).

Dretske, Fred. (2003) "Experience as Representation," *Philosophical Issues* 13: 67–82.

Fodor, Jerry. (1990) *A Theory of Content and Other Essays*, Cambridge: MIT Press.

Gordon, James and Israel Abramov. (2004) "Color Vision," in Bruce Goldstein, ed., *Blackwell Handbook of Sensation and Perception*, Oxford: Basil Blackwell.

Guth, S. L. (1991) "Model for Color Vision and Light Adaptation," *Journal of the Optical Society of America*, 976–993.

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

Hilbert, David and Mark Kalderon. (2000) "Color and the Inverted Spectrum," in S. Davis, ed., *Color Perception: Philosophical, Psychological, Artistic and Computational Perspectives*, Oxford: Oxford University Press.

Hurvich, L. (1981) *Color Vision*, Sunderland, MA: Sinauer Press.

Hurvich, L. and Jameson, D. (1955) "Some Quantitative Aspects of an Opponent-Colors Theory: I. Chromatic Reponses and Spectrum Saturation," *Journal of the Optical Society of America* 45: 546–552.

Johnston, Mark. (2004) "The Obscure Object of Hallucination," *Philosophical Studies* 120: 113–183.

Kay, P., Berlin, B., and Merrifield, W. (1991) "Biocultural Implications of Systems of Color Naming," *Journal of Linguistic Anthropology* 1: 12–25.

Lennie, Peter. (1999) Color Coding in the Cortex, in *Color Vision: From Genes to Perception*, ed. K. Gegenfurtner and L. Sharpe, Cambridge: Cambridge University Press, 235–47.

Lewis, David. (1984) "Putnam's Paradox," *Australasian Journal of Philosophy* 62: 221–236.

Lycan, William. (1996) *Consciousness and Experience*, Cambridge: MIT Press.

Lycan, William. (2002) "The Case for Phenomenal Externalism," in J. Tomberlin, ed., *Philosophical Perspectives*, *Vol. 16*: Atascadero, CA: Ridgeview Publishing.

MacLaury, R. E. (1997) *Color and Cognition in Mesoamerica: Constructing Categories as Vantages*, Austin: University of Texas Press.

McGinn, Colin. (1989) *Mental Content*, Oxford: Blackwell.

McGinn, Colin. (1996) "Another Look at Color," *Journal of Philosophy* 93: 537–553.

McGinn, Colin. (1997) "Missing the Mind: Consciousness in the Swamps", *Nous* 31: 528–537.

McGinn, Colin. (1999) "The Appearance of Colour," in *Knowledge and Reality: Selected Essays*, Oxford: Clarendon Press.

Millikan, Ruth. (1989) "Biosemantics," *Journal of Philosophy* 86: 281–297.

Papineau, David. (1993) *Philosophical Naturalism*, Oxford: Blackwell.

Pautz, Adam. (2003) "Have Byrne and Hilbert Answered Hardin's Challenge?," *Behavioral and Brain Sciences* 26: 44–45. http://philrsss.anu.edu.au/people-defaults/adamp/papers/have_byrne_and_hilbert_answered.pdf

Pautz, Adam. (2004) *The Hard Core of the Mind-Body Problem: Essays on Sensory Consciousness and the Secondary Qualities*, New York University Doctoral Dissertation.

Pautz, Adam. (2005) "Sensory Awareness is not a Wide Physical Relation", a longer version of the present paper. [http://philrsss.anu.edu.au/people-defaults/adamp/papers/sensory_awareness_not_wide.pdf]

Pautz, Adam. (2006) "Externalist Intentionalism and Optimal Conditions," [http://philrsss.anu.edu.au/people-defaults/adamp/papers/optimal_conditions.pdf]

Pautz, Adam. (forthcoming a) "Sensory Awareness as Irreducible: From Internalist Intentionalism to Primitivism," [http://philrsss.anu.edu.au/people-defaults/adamp/papers/sensory_awareness_as_irreducible.pdf]

Pautz, Adam. (forthcoming b) "A Critique of Realist Primitivism about Color" [http://philrsss.anu.edu.au/people-defaults/adamp/papers/critique_primitivism.pdf]

Porro, C., V. Cettolo, M. Francescato, P. Baraldi. (1998) "Temporal and Intensity Coding of Pain in Human Cortex," *Journal of Neurophysiology* 80: 3312–3320.

Shoemaker, Sydney. (1994) "Phenomenal Character," *Nous* 28: 21–38.

Shoemaker, Sydney. (2003) "Content, Character and Color," *Philosophical Issues* 13: 253–278.

Shoemaker, Sydney. (2006) "The Ways Things Appear," in T. Szabo Gendler and J. Hawthorne, eds., *Perceptual Experience*, Oxford: Oxford University Press.

Smith, D., John, S. and Boughter, J. (2000) "Neuronal Cell Types and Taste Quality Coding," *Physiology and Behavior* 69: 77–85.

Stevens, S. (1975) *Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects*, New York: John Wiley and Sons.

Tye, Michael. (1995) *Ten Problems of Consciousness: An Intentional Theory of the Phenomenal Mind*, Cambridge: MIT Press.

Tye, Michael. (2000) *Consciousness, Color and Content*, Cambridge: MIT Press.

Tye, Michael. (2003) *Consciousness and Persons*, Cambridge: MIT Press.