CAN THE PHYSICALIST ABOUT COLOUR EXPLAIN COLOUR STRUCTURE IN TERMS OF COLOUR EXPERIENCE?

Adam Pautz
The University of Texas at Austin
The Australian National University
apautz@mail.utexas.edu

ABSTRACT: I develop and defend the argument, first put forward by C. L. Hardin and others, that Color Physicalism must be rejected because it cannot accommodate color structure. In particular, I answer the response, defended by David Lewis, Brian McLaughlin and others, that the Color Physicalist can accommodate color structure by explaining color structure in terms of color experience. My main objection to this response is that it is at odds with the transparency observation and a Relational View of color experience.

I

Physicalism about colour is the thesis that colours are identical with response-independent, physical properties of surfaces [Armstrong 1997; Byrne and Hilbert 2003; Jackson and Pargetter 1987; Lewis 1997; McLaughlin 2003; Shoemaker 1991; and Tye 2000]. On the most popular form of Physicalism, colours are identical with (disjunctions of) dispositions to reflect various proportions of light at various wavelengths, or reflectance-types for short. I will focus on this version of the view, although what I shall say carries over to other versions. Physicalism has much to recommend it: it is realist and it is reductionist.

But some philosophers have mounted a powerful argument against Physicalism about colour based on claims about colour structure [Campbell 1969; Hardin 1988; Maund 1995; Thompson 1995]. We will focus on two kinds: claims about the resemblance of colours and claims about the unitary/binary character of colours. For instance, consider:

[1] Blue resembles purple more than green
[2] Purple is a binary colour, while green is a unitary colour

[1] is a claim about the resemblance of colours and [2] is a claim about the unitary/binary character colours. The sense in which green is a unitary colour that it does not, as it were, contain a hint of any other colour. It is not at all yellowish or bluish (or reddish, reddish greens never being found). Purple, by contrast, is binary in the sense that it contains a hint of two other colours: every shade of purple is to some degree reddish and bluish. Sometimes this is expressed by saying that shades of purple are “perceptual mixtures” of red and blue. In consequence, while green cannot be readily described in terms of other colours, purple can be described as red-blue. There are exactly four (types of) unitary colours: reds, greens, yellows and blues. All remaining colours are binary combinations of two of the four unitary colours and so can be characterized in terms of them; these include red-yellows, red-blues, greens-yellows, and green-blues. The resemblances among colours, their unitary-binary character, and other such facts about them, make up colour structure.

Here is my preferred way of formulating the argument from colour structure against Physicalism.

---

1 Thanks to Paul Boghossian, David Chalmers, Jonathan Cohen, Frank Jackson, and Daniel Stoljar.
Incompatibilism If Physicalism about colour is true, then claims about colour-resemblance and unitary-binary character are false.

Colour Structure Realism But such claims are evidently true. Therefore, Physicalism is false.²

The case for Colour Structure Realism is based on our experience of the colours. For example, we can just see that blues resemble purples more than greens, and that purples are binary while greens are unitary. The unitary/binary structure of colours may seem surprising and strange to the uninitiated, but many experiments demonstrate that, at least with a little practice, we can come to non-collusive agreement about the unitary/binary structure of colours. For instance, once we have had a little practice, we are able to come to agree that a given shade of orange is 75% red and 25% yellow [Boynton et al. 1964]. We may not know the analysis of such colour structure claims, but that they are true, and that we know that they are true on the basis of visual experience, seems beyond serious doubt. Therefore, any adequate account of colour must accommodate their truth.

The prima facie case for Incompatibilism is based on a natural understanding of colour structure claims. [1] is a claim of overall comparative resemblance among properties. Ordinarily, in such claims, ‘x resembles y more than z’ expresses an internal relation: a relation that supervenes on the intrinsic characters of its relata. For instance, suppose

\[ L_1 \text{ resembles } L_2 \text{ more than } L_3 \]

where \( L_1, L_2 \) and \( L_3 \) are determinate lengths. In [3] ‘x resembles y more than z’ expresses an internal relation. What makes [3] true is the intrinsic characters of the lengths \( L_1, L_2, \) and \( L_3 \). Of course, we might disagree about the nature of the relation: we might say that it is primitive, or that it is analyzable in terms of partial identity, or that being supervenient there really is no such relation [Armstrong 1997]. But that ‘x resembles y more than z’ picks out an internal relation in [3] is quite clear. We express the point in ordinary language by saying that for [3] to be true the lengths themselves must stand in the indicated resemblance-order. It is not enough that they be related to things that fall in the indicated resemblance-order. This seems like a truism hardly worth stating. In general, claims of comparative overall resemblance, if true, are made true by the intrinsic characters of the relata.³ Since it is also claim of comparative overall resemblance, it is

² Boghossian and Velleman [1991] and Johnston [1992] develop an epistemological argument with a similar form: they argue that if Physicalism is true, then we cannot know on the basis of visual experience the truth of colour structure claims, contrary to fact. The argument displayed in the text seems to be a more fundamental objection against Physicalism. So I will focus on it. The response to this argument that I will be considering here also provides a response to the epistemological argument, but I will not discuss this here.

³ Unlike lengths, things typically have many properties and vary along many dimensions. Thus, typically sentences reporting comparative resemblance are subject to a great deal of indeterminacy concerning what respects of similarity and difference matter, and concerning the importance assigned to the respect that do matter [Lewis 1986: 254]. The indeterminacy might be resolved differently in different contexts. But this does nothing to refute my claim that ‘x resembles y more than z’ always picks out an internal relation. For the respects that matter are always intrinsic respects. So, on every acceptable precisification, ‘x resembles y more than z’ expresses an internal relation; and a claim of comparative resemblance is made true, if it is true, by the intrinsic characters of the relata.

By the way, I said that [3] is true in virtue of the ‘intrinsic characters’ of the lengths. But lengths are properties and have themselves no interesting properties (except maybe being a length). So it might be wondered what I could mean by the ‘intrinsic characters’ of the lengths. (i) It might be that resemblances among lengths are analyzable in terms of partial identity [Armstrong 1997]. Then the sense in which lengths
natural to suppose that the same basic semantic account applies to [1]. In other words, it is natural to suppose that in [1] ‘x resembles y more than z’ picks out an internal relation, so that [1], if true, is made true by the intrinsic characters of the colours. As for [2], it is likewise natural suppose that [2] is true only if purple is itself (intrinsically) in some sense binary or composite while green is itself (intrinsically) in some sense unitary or non-composite. Call this the face value construal of [1] and [2].

Now, on Physicalism, the colours blue, purple and green are identical with the reflectance-types $R_b$, $R_p$ and $R_g$. But, going by their intrinsic characters, there is simply no sense in which $R_b$ resembles $R_p$ more than $R_g$. This is evident on viewing reflectance curves for reflectances of these types and is even granted by Physicalists about colour [Byrne and Hilbert 1997, 2003]. Nor, say defenders of Incompatibilism, is there any good sense in which $R_p$ is binary or composite while $R_g$ is unitary or non-composite [pace Tye 2000]. So, given the face value construal of [1] and [2], if Physicalism is true, they are simply false. Thus the prima facie case for Incompatibilism is based on the face value construal plus what we know about reflectance types from psychophysics.

The argument generalizes. In general, reflectances stand in similarity relationships which are very feeble approximations to the similarity relations of the corresponding colours [Thompson 1995; Byrne and Hilbert 1997: note 32). In general, there is no obvious sense in which reflectances are intrinsically unitary or binary. Therefore, given the face value construal of colour structure claims, if Physicalism is true, then many claims about the resemblances of colours, and all claims about their unitary-binary character, are simply false, contrary to Moorean fact. 4

My view is that the argument from colour structure against Physicalism about colour is sound. Instead I defend Primitivism about colour: the view that colours are simple, irreducible properties [Campbell 1993; Maund 1995; McGinn 1996]. For the Primitivist there is no pressure at all to reject colour structure claims. But some Physicalists say that the argument from colour structure is not sound [Shoemaker 1990, 1991; Lewis 1997; McLaughlin 2001, 2003; Cohen 2003]. They defend an account of colour structure claims in terms of colour experience that I’ll call the Experiential Account. On this account, the prima facie case for Incompatibilism is unsound because the face value construal of colour structure claims is a misconstrual. Under this account, there is no conflict between the truth of colour structure claims and the Physicalist view that colours are reflectance-types. Therefore the key premise of the argument from colour structure, Incompatibilism, is mistaken. In response to the argument from colour structure, defenders of the Experiential Account recommend a compatibilist position.

resemble by virtue of their intrinsic characters is that they resemble in virtue of what (shorter) lengths they have as constituents (relative to a ‘decomposition’). (ii) It might be that resemblance among lengths, and more generally resemblance among properties, is not analyzable at all. Then the sense in which lengths resemble by virtue of their ‘intrinsic characters’ is simply that they resemble in virtue of being the lengths that they are. This kind of view may sound nebulous but there is something intuitive about it, especially in the cases, for instance the case of the colours.

A note on metaphysics: I take it that there are properties. In the interest of simplicity, I also assume that properties are universals instead of tropes; that there are determinable universals such as the colours blue, purple, and green as well as determinate universals such as the various maximally specific shades of blue, purple and green; and that colour terms like ‘blue’, ‘green’ and ‘blue’ in statements like [1] and [2] are singular terms referring to such determinable colour universals. All the claims at play in what follows could be stated without loss of content in terms of determinate colours or tropes.

4 In fact the argument from colour structure generalizes to other ‘secondary qualities’: sounds, smells, tastes. In general, these properties stand in certain resemblance relations and have certain ‘structural properties’. But the corresponding extradermal physical properties – sound-waves, chemical properties, and so on – do not stand in these same resemblances relations, and do not have the structural properties in questions. But I will stick to colour here.
It may be that the argument from colour structure is the greatest threat to Physicalism about colour. Unlike other arguments against Physicalism, such as the argument from revelation or the argument from the apparent simplicity of colours, one cannot reject it because it rests on some highly contentious claim about the colours - because it doesn't. It just relies on what we evidently know about the colours through visual experience together with what we know about reflectance-types through psychophysics (plus a very natural view on what it takes for colour structure claims to be true). The Experiential Account provides the most systematic and popular response to this argument. For these reasons, it merits careful evaluation. My purpose here is to provide an evaluation of the Experiential Account. Along the way we will have occasion to examine other interesting issues in the philosophy of colour and colour perception, for instance the nature of the qualitative character of experience and the functionalist model of how colour names refer.

I'll begin by providing an exposition of the Experiential Account. Then I'll examine it critically. Then I'll say what the upshot of the discussion is for the argument from colour structure and the prospects for Physicalism about colour.

II

By an account of claims (statements, propositions, sentences, beliefs) of a certain kind, I simply mean a systematic description of their truth-conditions. We may also 'semantically descend' and speak of accounts of properties and facts. Defenders of the Experiential Account propose a systematic account of colour structure claims and facts in terms of colour experience. Here are some representative passages:

For surface properties of objects to be similar in colour relative to a certain sort of perceptual system is presumably for it to be the case that objects with those properties standardly produce, in creatures having that sort of visual system, experiences that are similar in that way. [Shoemaker 1991: 519]

Our account [Physicalism about colour] provides a correspondence between colours and colour experiences. . . Whatever form it takes, the correspondence yields relations among colours in the image of relations among colour experiences . . . [Lewis 1997: 339]

The comparative claim about red, orange, and blue [red resembles orange more than blue] is thus true in virtue of a comparative fact about the visual experiences in question. [McLaughlin 2003: sect. 9]

facts of [colour-resemblance] and the unique/binary hue distinction are, in the first instance, about the phenomenological character of colour experiences--the what-it's-like aspects of seeing colours; they are only derivatively about colours themselves. [McLaughlin 2001: sect. 13]

[Physicalists about colour] may explain the structural relations among the colours in terms of the structural relations among the experiences colours dispose their bearers to produce [Cohen 2003: 88]

The idea, then, is this. Physicalism about colour is correct. On Physicalism about colour, by contrast to the various error-theories about colour, colours are physical properties of external objects that play a causal role in the generation of our colour experiences. Suppose that this is right. Then we may define a
correspondence between colours and colour experiences. A given colour (on this view, a given reflectance-type) causes different colour experiences in us under different conditions of observation. But for every colour there will be the type of colour experience it would cause if optimal conditions obtained: conditions in which lighting conditions are normal, the subject is a normal state of adaptation and has a normal, properly functioning visual system, and so on. Let us use ‘E[C]’ as short-hand for the colour experience (type) that colour C would cause if optimal conditions obtained. And let us say that the colour experience E[C] corresponds to the colour C.

Now we may state the Experiential Account. According to the Experiential Account, the truth-conditions of [1] and [2] are:

- [1a] E[blue] resembles E[purple] more than E[green]
- [2a] E[purple] is binary, while E[green] is unitary.

[1a] says that the experience corresponding to blue resembles, in respect of its qualitative character, the experience corresponding to purple more than the experience corresponding to green. This should be understandable enough. [2a] requires some explanation. Here the predicates ‘is binary’ and ‘is unitary’ are applied to colour experiences. Ordinarily we apply these predicates not to our colour experiences but to the colours that we experience. So what do these predicates mean when applied to colour experiences? I take it that the idea is this. Colour experiences have a ‘qualitative character’. Some colour experiences are ‘binary’ in respect of their qualitative character, and others are ‘unitary’ in respect of their qualitative character. What [2a] says is that the experience corresponding to purple is binary in respect of its qualitative character while the experience corresponding to green is unitary in respect of its qualitative character. What these facts about the qualitative character of colour experiences come to depends on what theory of qualitative character is right; we will turn to this issue presently.

So, on the Experiential Account, somewhat surprisingly perhaps, when we make colour structure claims such as [1] and [2], we are really talking about our own colour experiences; or more cautiously, the truth or falsity of what we say hangs on facts about our colour experiences - their resemblances and unitary/binary characters. The account may be extended in the obvious way to other colour structure claims. The result is a response-dependent account of colour structure claims, in the sense that it is claimed that facts about observers enter into the truth-conditions of such claims. Thus, while defenders of the Experiential Account give a response-independent account of colours themselves (at least in the sense that they identify them with observer-independent properties, namely reflectance-types), they give a response-dependent account of their second-order, structural properties. But note that this account does not, or need not, entail that the truth of such claims depends on the actual presence of observers. For the relevant correspondence relation between

---

5 In general, from the fact that entities of a certain kind enter into the story of what makes sentences in a certain class true, it does not follow that when we utter those sentences we are talking about those entities. For instance, on many views, properties enter into the story of what makes simple predications true; but this does not mean that when we make simple predications we are talking about properties. Likewise, the defender of the Experiential Account might say that experience-types enter into the story of what makes colour structure statements true, yet when we make such statements we are not talking about colour experiences. Rather, it might be said, they are claims entirely about colours, which are true if those colours are suitably related to colour experiences. But I will sometimes speak as if defenders of the Experiential Account do hold that when we make colour structure statements we are talking about our colour experiences (their language sometimes suggests it: see for example the passage from McLaughlin above). Nothing I will say, and in particular none of the objections I will raise against the Experiential Account, depends on their acceptance of this claim. All my objections will be directed at the claim that [1] and [2] have [1a] and [2a] as their truth-conditions; and defenders of the Experiential Account certainly hold this.
colours and colour experiences is defined in counterfactual/dispositional terms. \( E \) corresponds to \( C \) iff \( C \) would produce \( E \) in us under optimal viewing conditions.

The Experiential Account explains the structural properties of colours in terms of the structural properties of colour experiences. But what is the correct account of the structural properties of colour experiences?

Here is a very natural idea. Where \( E_1, E_2 \) and \( E_3 \) are colour experiences, \( E_1 \) resembles \( E_2 \) more than \( E_3 \) iff in having these colour experiences one experiences the colours \( C_1, C_2, C_3 \), and \( C_1 \) resembles \( C_2 \) more than \( C_3 \). For instance, what makes it the case that \( E[\text{blue}] \) resembles \( E[\text{purple}] \) more than \( E[\text{green}] \) is that in having these colour experiences one experiences, respectively, the colours blue, purple, and green (or certain determinate shades thereof), and the colour blue resembles the colour purple more than the colour green. (Here and in what follows, I use ‘experience’ in what is sometimes called the ‘phenomenological sense’. In this sense, one can be said to experience a colour – one can be aware of a colour, it can look to one as if something has a colour – even if one does not experience a material object that has that colour.) Likewise a colour experience \( E \) is binary iff the colour that one experiences in having \( E \) is binary. And so on. In general, the idea is that the structural properties of colours are simply inherited from (or amount to) the structural properties of the colours that we are sensorily related to in having those colour experiences. Call this Relationism about the structural properties of colour experiences. For myself I believe that it is fundamentally right.

This idea can be developed in different ways, depending on what view one takes on what it is to have an experience as of a colour, or more generally, what it is to have an experience as of a certain ostensible property. Sense Datum theorists hold to experience a colour one must experience a mental object that possesses that colour. Intentionalists (or ‘Representationalists’) deny this [for instance Tye 2000]. They say that to experience a colour is to ‘visually represent’ that something has that colour. One can be perceptually acquainted with a colour – one can visually represent that something has a colour – even if one does not experience anything that has that colour, as it might be, in a case of illusion or hallucination. Other Relationist alternatives to the Sense Datum Theory are the Meinongian Theory [Smith 2002; McGinn 2004], the Property-Complex Theory [Bealer 1982; Johnston 2004], and the Multiple Relation Theory [Alston 1999]. However, for our purposes here it does not matter how the basic idea is developed.\(^6\)

Relationism about the structural properties of colour experiences, however it is developed, is very natural for the following reason. Many philosophers have been impressed by the observation that when we attempt to focus on qualitative character of a visual experience we end up “looking through it” to the familiar properties – colours, shapes, etc. – that we experience in having the visual experience. This is called the transparency observation [see for example Tye 2000]. The observation applies also to the structural properties of colour experiences. This is not surprising, since the structural properties of colour experiences supervene on their qualitative characters. For instance, suppose that you have concurrently three colour experiences, \( E_1, E_2 \) and \( E_3 \), in which you experience colours \( C_1, C_2, C_3 \), respectively. If you try to determine whether \( E_1 \) resembles \( E_2 \) more than \( E_3 \), you inevitably end up focusing on the colours \( C_1, C_2, C_3 \), and trying to determine whether \( C_1 \) resembles \( C_2 \) more than \( C_3 \). In general, if you try to focus on resemblances among your colour experiences, you end up focusing on the resemblances among the colours that you experience. Likewise, if you try to determine whether a colour experience is ‘binary’, you simply

\(^6\) Shoemaker defends a form of Relationism – in particular, a form of Representationalism – which differs from the view characterized in the text in the following way: he says, not that the qualitative character of experience is determined by what ‘colors’ we experience (or in his favored idiom, visually represent things as having), but by what ‘colour appearances’ we experience. (For his most recent version of the view, see [Shoemaker 200X forthcoming]. In the interest of simplicity, I will not discuss this view here. When I speak of ‘Relationism’, I will always mean Relationism as formulated in the text (in terms of ‘colors’).
end up focusing on the colour you experience and trying to determine whether it is binary. And so on. This strongly suggests some form of the Relationist view that the structural properties of colour experiences derive from (or are nothing but) the structural properties of the colours we experience in having those colour experiences. The transparency observation may not show which form of Relationism is correct – that is a matter to be decided on the basis of other considerations. But it does strongly suggest that some form of Relationism is correct.

But the defender of the Experiential Account cannot accept Relationism in any of its various incarnations for a very simple reason. One cannot analyse colour resemblance in terms of resemblance among colour experiences and then analyse resemblance among colour experience in terms of resemblance among colours. That would be circular. Thus Shoemaker writes, “if the similarity of objects with respect to consists in their aptness to produce colour experiences that are phenomenally similar, then the phenomenal similarity of colour experiences cannot consist in the fact that they represent their objects as similar with respect to colour ” [1991: 519].

In consequence, defenders of the Experiential Account need another account of the structural properties of colour experience. Here is one idea. According to a very simplified version of the opponent process theory of colour vision, we have two opponent channels, the R-G channel and the Y-B channel. Each can assume a positive or negative state of activation. When the R-G channel assumes a positive state of activation we have a reddish experience, and when it assumes a negative state of activation we have a greenish experience. (By a ‘reddish experience’ I mean an experience of the qualitative type in fact produced by objects we would call ‘red’. This terminology allows us to remain neutral on the debate over the nature of qualitative character.) Likewise for the Y-B channel. Thus the structural properties of colour experiences are mirrored by, and explained by, the structural properties of certain of our brain states. Some brain states are ‘binary’ (they involve the joint activation in two channels) and some are ‘unitary’ (they involve activation only in one channel). Further, in general, if \( E_1 \) resembles \( E_2 \) more than \( E_3 \), then \( B_1 \) resembles \( B_2 \) more than \( B_3 \), where \( B_1 \), \( B_2 \), and \( B_3 \) are the opponent channel states correlated with \( E_1 \), \( E_2 \), and \( E_3 \), respectively. So the defender of the Experiential Account might identify colour experiences with brain states and the structural properties of colour experiences with the structural properties of those brain states. (McLaughlin [2003: end of sect. 18] express sympathy with this view.) Call this the Neurobiological Account of the structural properties of colour experience.

The Neurobiological Account should appeal to defenders of the Experiential Account of the structural properties of colours for two reasons. First, it doesn’t appeal to the structural properties of colours, so they can accept it without fearing circularity. Second, the defenders of the Experiential Account advocate a reductive view of colours. Presumably, at least some of the reasons they have for being reductionists about

---

7 The opponent process theory may suggest the Neurobiological Account, but it is not the case that the Neurobiological Account is the only account of colour experience compatible with the opponent process theory. Relationism is also compatible with the opponent process theory. It is true that most Relationists are also externalists [for example Tye 2000]. They maintain that what colours we experience, and thereby the structural properties of colour experience, is not determined by what goes on in the brain, but by extrinsic head-world connections. In my view, they are constrained to reject the opponent process theory. But this is not compulsory. The Relationist might also be an internalist Relationist. On this view, what colours we experience things as having at a time, and thereby (given Relationism) the structural properties of our colour experiences, is determined by what brain states we are in at that time. This view would accommodate the opponent process theory. It would make non-accidental the isomorphism which, according to the opponent process theory, holds between our brain states and the colours we experience: for example, the fact that if you are in similar opponent channel states, you experience similar colours, and if you are in a ‘binary’ brain state, you experience a binary colour [Author 200X forthcoming].
colours carry over to colour experiences. So it seems that by parity of reasoning they ought to advocate a reductive view of colour experiences. The Neurobiological Account is just such a view.

Here I will in various places assume that the defenders of the Experiential Account accept the Neurobiological Account of colour experiences and their structural properties, although the points I will make are independent of this assumption and would apply even if they accept some other view. Under this assumption the Experiential Account becomes the view that colour structure claims about the colours are made true by facts about the structural properties of the corresponding brain states.

The payoff of Experiential Account for defenders of Physicalism about colour is that it undercut the argument from colour structure by implying the falsity of its key premise, Incompatibilism. Suppose you have experiences of blue, purple and green, and so utter:

1. Blue resembles purple more than green
2. Purple is a binary colour, while green is a unitary colour

As we have seen, on the Experiential Account, perhaps somewhat surprisingly, in uttering [1] and [2] you are ‘really’ talking about your colour experiences (that is, on the Neurobiological Account, your brain states), and the truth-conditions of your statements may be represented thus:

1a. E[blue] resembles E[purple] more than E[green]
2a. E[purple] is binary, while E[green] is unitary

Therefore, according the Experiential Account, contrary to what in the introduction I called the face value construal [1] and [2], the truth of what you say in uttering [1] and [2] does not require that that the colours blue, purple and green themselves fall in the indicated resemblance order or that they themselves have in any sense the properties of being composite or non-composite; it just requires that the corresponding colour experiences (or brain states) do so.

Now enter Physicalism. On Physicalism, the colours blue, purple and green are identical with the reflectance-types $R_b$, $R_p$ and $R_g$. So, given the Experiential Account of [1] and [2], we obtain the following as the truth-conditions of [1] and [2]:

1b. E[$R_b$] resembles E[$R_p$] more than E[$R_g$]
2b. E[$R_p$] is binary, while E[$R_g$] is unitary

These truth-conditions obtain. As noted previously, it is not the case that the reflectance-types $R_b$, $R_p$ and $R_g$ themselves fall in the right resemblance order or that they themselves in any sense have the properties of being composite or non-composite; but the corresponding colour experiences (or brain states) certainly do. We are so wired up that $R_b$ and $R_p$, although not themselves similar, optimally produce in us very similar brain states, and that $R_p$ and $R_g$, although not themselves binary or unitary, optimally produce in us ‘binary’ and ‘unitary’ brain states. (The ‘binary’ brain state is literally made up of R-activity and B-activity.) So, given that [1] and [2] have the truth-conditions that the Experiential Account assigns to them, in uttering [1] and [2], one comes out as speaking truly, despite the mismatch between colours and the corresponding reflectance-types. As McLaughlin writes, on the Experiential Account, “[n]othing we could learn about the properties that are in fact the colours that is independent of how they dispose things to look would refute any of these claims, for if the comparative claims are true, they are so in virtue of what it is like for things to look the colours in question” [2003: sect. 9].
It is helpful to view the Experiential Account as an instance of a general strategy in philosophy noted by Kripke [1982: 64-5]. A philosopher proposes a philosophical theory of some part of reality. The theory appears at odds with some common sense claims about that part of reality. The philosopher proposes that this appearance is based on a misunderstanding of what it takes for these claims to be true. On the correct understanding of what it takes for these claims to be true – the one proposed by the philosopher – there is no conflict. For instance, Berkeley asserted that his Idealism and our ordinary claims about material objects are not in conflict because the correct account of those claims is a phenomenalistic, response-dependent one in terms of the ideas we would have if we did certain things. It is exactly the same in the present case. On the face value construal of colour structure claims, there is a conflict between their truth and the Physicalists’ claim that colours are reflectance-types. In response, Physicalists say that the face value construal is a misconstrual, and that on the correct understanding of those claims – the one provided by the Experiential Account – there really is no conflict.

Kripke expresses [1982: 65] his opinion that typically the face value construal of the claims in question is correct, and that the real misconstrual comes when the says, “All the ordinary man really means is . . .” and provides a complicated analysis compatible with his own philosophy. This is my view in the present case. There is nothing to say for the Experiential Account and plenty to say against it. To bear my case out, I will raise five objections against the Experiential Account.

* It is important to keep in mind that the Experiential Account is not a reformatory analysis. It is not claimed that, in view of the truth of Physicalism, we should ‘change the meanings’ of the colour structure sentences [1] and [2] so that they have [1a] and [2a] as their truth-conditions, in order that we may speak truly when we utter them. It is claimed that they have had these truth-conditions all along; the Experiential Account is a view about the actual semantic content of a certain class of English sentences. Moreover, it is a complicated analysis which attributes to the sentences in question ‘deep structures’ which are quite different from their surface structures. So it is surprising how little defenders of Experiential Account have to say in its favor. For the most part, they just seem to pull it ‘out of thin air’. Of course, for Physicalists, the Experiential Account has a very nice feature: it allows a compatibilist response to the argument from colour structure. But that does not show that it actually is true. Is there any genuine reason to accept it, aside from the question-beggning motivation that it secures a compatibilist response to the argument from colour structure?

Typically, a non-obvious semantic analysis is supported its ability to account for the validity of certain inferences, to explain ambiguities, to solve semantic puzzles, and to account for other such semantic data. (Think of the evidence Russell cited in favor of a quantificational analysis of definite descriptions.) But defenders of the Experiential Account do not cite any evidence of this kind in favor of their view, and I cannot think of any such evidence in its favor

Brian McLaughlin makes some points which he apparently thinks support the Experiential Account, but I believe that they are either mistaken or do not support the Experiential Account. (i) He says, “We know the facts of unity [resemblance] and the relationships between the unique and binary hues as a result of attending to the phenomenological characters of colour experiences” [2001: end of sect. 14]. In my view, this is not the case. We know these facts by attending to the colours themselves. We look out, not in. (As a Relationist, I would say that except perhaps in rare cases) there is no such thing as ‘attending to the phenomenological characters of colour experiences’ as distinct from attending to the colours that we experience, but this is not essential to my point. See pp. 19-20 of this paper.) It is true that our beliefs about the structural properties of colours are justified by our experiences of colours (or coloured things), but this does not entail that those beliefs are about our colour experiences, as the Experiential Account would have it. (ii) McLaughlin also says, “The unique/binary hue distinction is drawn phenomenologically” [2001: end of sect. 14]. This is not always true. For instance, Boynton et al. [1964: r] write that a unique yellow is “one that cannot co-exist with blue and that is neither reddish nor greenish”. This characterisation does not mention experiences. It is true that the distinction is sometimes drawn phenomenologically. For instance, Hurvich [1981: r] writes that the “unique point” in the hue-circle is the “place where a yellow occurs in which we cannot detect any red or even any green”. But this just seems to be a heuristic and does not show that the Experiential Account is true. In other words, it does not show that claims about the unitary/binary characters of colours are to be analysed in terms of colour experiences. To see this, consider a simple analogy. The distinction between clean shirts and dirty shirts, like the distinction between unitary and binary colours,
My first objection is that the Experiential Account is unavoidably circular. It gets the direction of explanation exactly backwards. It attempts to analyse the structural properties of colours in terms of the structural properties of colour experiences. But intuitively the correct order of explanation is exactly the reverse of this. The structural properties of colour experiences are to be analysed in terms of the structural properties of the experienced colours. In other words, Relationism about the structural properties of colour experiences is correct. As noted before, one cannot have it both ways. If resemblances among colour experiences just amount to resemblances among the experienced colours, there is simply no room to analyse resemblance among colours in terms of resemblance among colour experiences. So the Experiential Account must be false.

David Lewis seems address this objection to the Experiential Account in the following passage:

> We might have had the offhand opinion that these relations [resemblance relations] originated as relations among surface properties [colours]. If so, we were wrong. But I am not sure we had any such option at all; and if we did, we have no business elevating it into a Moorean fact of folk psychophysics. [1997: r]

Lewis here makes it seem as if we have no reason to think that structural properties of colour experience derive from the structural properties of the colours (as Relationism would have it) rather than the reverse (as the Experiential Account would have it). But in my opinion there is reason to think that this is the case: as noted above, the transparency observation provides one argument for Relationism, and I believe that there are others.

I would be content to leave the matter here, because I am strongly committed to Relationism. But I am conscious that others may not be persuaded of the truth of Relationism. However, there are other objections.

My second objection is as follows. The Experiential Account is a view about the actual truth-conditions of our colour structure beliefs and statements. But it seems that we know that it gets the truth-conditions wrong on the basis of simple introspection.

We have some degree of ‘privileged access’ to the truth-conditions of our beliefs and statements. The access may not be complete. This, it might be thought, is shown by unobvious analyticities, externalism about content, and the fact (if it is a fact) that the best semantic theories for certain sentences of English are at variance with our offhand semantic intuitions concerning their truth-conditions. But we do have through introspection some degree of access to what it takes for our beliefs and statements to be true and what objects and properties they are about.

For instance, suppose that I come to believe

---

shows up in experience. So, it can be indirectly characterized in terms of experience: we might say, as a rough and ready characterization, that shirt is clean when we can detect no dirt in it, and dirty otherwise. Yet the distinction between clean shirts and dirty shirts is not correctly analysed in terms of our experiences. Maybe it is the same for the distinction between unitary colours and binary colours.
[3] $L_1$ resembles $L_2$ more than $L_3$,

where as before $L_1$, $L_2$, and $L_3$ are determinate lengths. We know that the Experiential Account is not true in this case. My belief is not true merely if $E[L_1]$ resembles $E[L_2]$ more than $E[L_3]$. Rather, the face value construal is correct in this case. The truth of my belief requires that the lengths themselves stand in the indicated resemblance-order. It is not enough that they optimally cause experiences in us that do so. How do we know this? First, it is introspectively evident that my belief is entirely about what I experience — the lengths $L_1$, $L_2$, and $L_3$ — and has nothing to do with experiences. Second, it is just obvious that in believing [3] what I believe is that the lengths themselves stand in the indicated resemblance-order. (In the jargon: my belief is true if the lengths fall into this resemblance-order by virtue of their intrinsic characters.) So for my belief to be true it is not enough that the lengths optimally cause in us experiences that fall into this resemblance-order. It is required that the lengths themselves fall into that order.

But now suppose I experience a blue marble, a purple marble and a green marble, and focusing on their colours, come to believe

[1] Blue resembles purple more than green

The very same introspective evidence against the Experiential Account and for the face value construal applies here. So, by parity of reasoning, we ought to reject the Experiential Account and accept the face value construal in this case as well. (If we don’t, it seems that we will have to admit that we are not justified in rejecting the Experiential Account in the case of length on the basis of the introspective evidence, for in the case of length introspective evidence against the Experiential Account is exactly parallel. This is hard to believe.) My belief that blue resembles purple more than green is not true merely if $E[blue]$ resembles $E[purple]$ more than $E[green]$. First, it is introspectively evident that my belief is entirely about what I experience — the colours blue, purple and green — and has nothing to do with experiences. Experience-types do not enter into the truth-conditions at all. In forming my belief, my attention was focused out, not in.

---

9 Nominalists would say that my belief is not about lengths (there are none) but about the particulars that we would (misleading, in their view) say ‘have’ lengths. I disagree, but I do not think that the matter can be decided by simple introspection. So I should rather say it is introspectively evident that my belief is about what I experience, lengths or the things that have lengths.

10 Kripke made a well-known objection against Lewis’s counterpart theory [1980]. Sometimes it is interpreted (whether rightly or wrongly) as the following objection: if the counterpart theory is correct, the claim that Humphrey might have won the election is not about Humphrey but about his counterparts. Of course, the obvious reply is that, on the counterpart theory, it remains a claim about Humphrey; it is a claim about Humphrey, which attributes to him the property of having a winning counterpart [Lewis 1986: 196].

My objection to the Experiential Account is not analogous to this one. I do not say that if the Experiential Account is true then colour structure claims are not about colours but about colour experiences, although I sometimes speak as if the Experiential Account has that implication in the text (see note 5). Instead, I say that the Experiential Account has the consequence that colour experiences, in addition to colours, enter into the truth-conditions of our colour structure claims (i.e. it has the consequence that the truth or falsity of our colour structure claims hangs on facts about our colour experiences). The Experiential Account does have this consequence. But it is highly counterintuitive. Intuitively, it is only the colours that we experience that enter into those truth-conditions. This is my objection.

My objection against the Experiential Account may be usefully compared to a distinct objection against the counterpart theory. The counterpart theory does have the consequence that individuals in other worlds enter into the truth-conditions of our de re modal claims. But this is highly counterintuitive. As Salmon writes, “when we say that Humphrey might have won, what we say certainly has nothing to do with the political goings-on in alternative universes . . .” [1988: 239]. Lewis seems to be addressing this distinct objection to the counterpart theory when he writes, “Counterpart theory does say . . . that someone else —
Second, in believing [1] what I believe is that the colours themselves stand in the indicated resemblance-order. I know what I believe, and this is what I believe. So for my belief to be true it is not enough that the colours optimally produce experiences in us that fall into that resemblance-order. It is required that the colours themselves do so. If they do not, my belief is false.

Of course, the same applies to everyone else. It is plain that in believing [3] what one believes is that the colours themselves fall into the indicated resemblance-order. There is a convention in our language to use the sentence ‘blue resembles purple more than green’ to communicate this belief (there is a convention of truthfulness and trust, in the sense of Lewis [1975]). So, ‘blue resembles purple more than green’ simply does not have the truth-conditions that the Experiential Account assigns to it. It has the truth-conditions that the face value construal assigns to it. 11

To say, as the defenders of the Experiential Account do, that our beliefs about length resemblance are about the length themselves, while our beliefs about colour resemblance are ‘really’ about our experiences of the colours, is to read something into our beliefs about colour resemblance that is simply not there. We are what might called ‘naïve’ about colour resemblance. We believe that colours themselves fall into certain resemblance-orders, just as we believe that lengths (and properties in other families) themselves fall into certain resemblance-orders. So, the Experiential Account is simply wrong about the truth-conditions of our

11 Some clarifications. (i) I am not arguing here that colours do themselves (‘intrinsically’) resemble in the way we take them to, or that colours are themselves (‘intrinsically’) unitary or binary. I am making a more modest claim: that this is what we believe, and that this is what is required for the truth of our ordinary colour structure claims. It follows that the Experiential Account of those claims is false and the face value construal true. Since the Experiential Account is false, and since the compatibilist response to the argument from colour structure that we are examining here rests on the Experiential Account, it follows that this compatibilist response fails. (ii) It may seem that my appeal introspection is implausible on the grounds that ordinary people would not claim that in believing (1) what they believe is that the colours stand in the indicated resemblance order by virtue of their intrinsic characters. That may seem too theoretical. My reply is that they would claim this; they just would not express it by using the technical term ‘intrinsic characters’. For they would claim that they believe that the colours themselves stand in the indicated resemblance-order (just as they would claim that they believe that the lengths themselves stand in the resemblance-order indicated in [3]); and this is a way of expressing in ordinary language what philosophers would express by saying that the colours stand in the indicated resemblance-order by virtue of their intrinsic characters. (iii) Nothing I say requires the dubious claim that the property or concept of being intrinsic or intrinsic character enters into the contents of ordinary peoples’ colour structure beliefs. To see this, let us suppose – just for the purposes of illustration – that a Russellian view of the content of belief is correct. Then my claim can be put in this way. There is a three-place internal relation of comparative resemblance, call it Resembles-More. Given a Russellian framework, when one believes [1], the content of one’s belief may be represented as follows: <Resembles-more, <blue, purple, green>>. This does not require that the concept or property of being intrinsic enters into the content of the belief.
colour resemblance beliefs and statements. The face value construal of these beliefs and statements is the correct one.

Parallel remarks apply in the case of beliefs and statements about the unitary/binary character of the colours.

My third objection is that there is no plausible story to be told concerning how colour structure beliefs and statements could have the truth-conditions that the Experiential Account assigns to them.

Some philosophers think that our pretheoretic intuitions about truth-conditions are not sacrosanct and may be overruled by other considerations [Lewis 1986: 240-1]. So the defender of the Experiential Account might stand his ground and maintain, contrary to our intuitions on the matter, that colour structure statements and beliefs have the truth-conditions that the Experiential Account assigns to them. But he would still face the following distinct objection. Somehow the non-intentional facts determine the intentional facts. So if the Experiential Account is correct, and colour structure sentences and beliefs have response-dependent truth-conditions, then there must be some non-intentional facts that make it the case that they have these truth-conditions. But there seem to be no such facts. There seems to be nothing in our use of colour structure sentences, or in our use of such sentences together with the way the world is, which could make it the case that they have these truth-conditions. So the Experiential Account must be false.

Here is a way to underscore the problem. Consider statements about the resemblance of colours and statements about the resemblances of other families of properties, for instance lengths. The underlying non-intentional (use, causal, functional, conceptual role) facts seem relevantly the same. So what could make it the case that they have the radically different kinds of truth-conditions that the defenders of the Experiential Account assign to them?

It is natural to think that the truth-conditions of sentences are determined by the truth-conditions of the beliefs that they are conventionally used to express [Lewis 1975]. So the defender of the Experiential Account might say that the solution to the problem is simple. The beliefs we use colour structure sentences to express have response-dependent truth-conditions. (This may be introspectively implausible – that was the basis of my first objection. But we are now assuming that the defender of the Experiential Account chooses to ignore the introspective evidence against his view.) And colour structure sentences inherit the response-dependent truth-conditions of the beliefs that we use those sentences to express.

But this ‘solution’ doesn’t make the problem go away; it just kicks the problem upstairs from language to thought. The problem now becomes: what non-intentional (causal, functional, conceptual role) facts make it the case that our colour structure beliefs (or the neural realizers of those beliefs) have the response-dependent truth-conditions in question?

It might be replied that this is a problem for everyone. It is just the general problem of saying how the non-intentional facts settle the truth-conditions of our beliefs and sentences. But this is not right. For instance, maybe I can see what might make it the case that ‘water is wet’ is true iff H₂O is wet, and maybe I can see what might make it the case that a quantificational analysis of definite descriptions is correct. But I cannot see what could make it the case that colour structure beliefs and sentences could come to have the response-dependent truth-conditions that defenders of the Experiential Account assign to them.¹²

¹² One might think that the objection can be answered by appealing to some kind of causal or counterfactual-dependence account of content: very crudely, what makes it the case that our colour structure beliefs have the truth-conditions that the Experiential Account assigns to them is that they are caused by our colour experiences when things go right. This strikes me as a non-starter. Many of our beliefs are caused by our experiences, but are not about our experiences. For instance, our beliefs about the resemblances of lengths are typically caused by our experiences of those lengths, but the Experiential Account is not true of our beliefs about resemblances among lengths. Indeed, we can imagine a world, ω, in
It might be thought that the objection can be answered by appealing to a theory of the content of thought and language that makes heavy use of a principle of charity, as follows. What one believes, and what the sentences in one’s language mean, is given by the best total assignment of contents to one’s inner states and sentences. The best total assignment is determined by the functional roles of one’s states plus certain constitutive principles. One such principle is the principle of charity. Roughly speaking, this principle states: all else being equal, maximize the truth of our beliefs – or at any rate, those beliefs of ours which we can reasonably be expected to get right [Lewis 1974]. Colours are identical with reflectance-types (assume). There is no response-independent sense in which \( R_b \) resembles \( R_p \) more than \( R_g \) or in which \( R_p \) is binary while \( R_g \) is unitary. So, given that colours are reflectance-types, our colour structure sentences and beliefs, and in particular [1] and [2], can come out true only if they have the response-dependent truth-conditions that the Experiential Account assigns to them. In consequence, they have those response-dependent truth-conditions. (This may go against our intuitions, but our intuitions are not sacrosanct.) The same is not true of sentences and beliefs about resemblances among lengths. These claims can come out true even if they have response-independent truth-conditions. For lengths really do intrinsically resemble in the ways we think that they do. So they have response-independent truth-conditions, in agreement with our intuitions on the matter. This is what makes it the case that colour structure claims and lengths structure claims have quite different kinds of truth-conditions.

This reply runs the risk of begging the question. We are engaged in the task of evaluating the Experiential Account. The Experiential Account is offered in answer to an objection to Physicalism about colour. I have raised an objection against the Experiential Account: that there is nothing that could make it true. The answer to my challenge just given rests on: (i) a theory of content that makes heavy use of a principle of charity and (ii) Physicalism about colour. So it would appear question-begging in the present context.

There is another problem. Let \( T \) be the correct theory of content. Suppose that this answer to the problem of semantic mystery is correct: \( T \), together with the claim that colours are reflectance-types, entails that colour structure beliefs have response-dependent truth-conditions, because this assignment of truth-conditions is in accordance with the principle of charity (and maximises the satisfaction of the other relevant ‘constitutive’ principles, and so on). Then by parity of reasoning \( T \) should have similar consequences in other cases. For instance, suppose that Berkeley is right: there are only minds and their ideas. If this were the case, \( T \) should have the consequence that an ordinary claim about the material world, \( p \), is true iff \( it \ is \ as \ if \ p \), because given that we live in a Berkeleean world this assignment of truth-conditions is in accordance with charity. But this is intuitively incorrect: if Berkeleean idealism turned out to be true, ordinary claims about material objects would not have these truth-conditions. They would have their face-value truth-conditions, and so would be uniformly false. The point is a general one. If the proposed answer to the problem of semantic mystery is correct, then by parity of reasoning the truth-conditions of our beliefs should in general malleable to a great degree. They should be dependent on how the world turns out to be, so that they are by and large aligned with how the world turns out to be. As this is not the case, the proposed answer must be incorrect. Charity, although it might have some role in determining content, cannot have the very strong role that the response assigns to it.

which we only have visual experiences, and in which do not interact causally with objects in our environments. In \( w \), our beliefs about resemblances among lengths are only ever caused by our visual experiences of those lengths, but even in \( w \) the Experiential Account of our beliefs about the resemblances of lengths would not be correct. So, the fact that beliefs about resemblances among colours are only ever caused by our experiences of colours cannot be sufficient to make it the case that these beliefs are about our colour experiences and have the truth-conditions that the Experiential Account assigns to them.
I conclude that the defender of the Experiential Account has no plausible story concerning what might make it the case that colour structure beliefs and statements have the response-dependent truth-conditions which that account assigns to them. With nothing to make the Experiential Account true we must conclude that it is false.

My fourth objection is that the Experiential Account is incredible given our experience of the colours. Suppose that you are looking at a shade of orange. It looks binary to you. But it does not look disposed to produce binary experiences in people when optimal conditions obtain. Likewise, when colours look similar, they do not look such as to cause similar experiences in people when optimal conditions obtain. This strongly suggests, contrary to the Experiential Account, that colour structure facts are not dispositional facts about what colour experiences colours are disposed to cause in people when optimal conditions obtain.

Here is a more informal way to appeal to phenomenology to argue against the Experiential Account. On the Experiential Account, the fact that orange is binary amounts to the fact that, if optimal conditions were to obtain, it would cause binary colour experiences (or brain states) in human beings. Thus the fact that orange is binary is an extrinsic and relational fact about it. But this just seems wrong on the basis of visual experience. The fact that orange is binary seems to be a wholly intrinsic, 'here-and-now' fact about the colour. To see that it is binary, you only need to focus on the colour itself. Likewise for colour resemblance. On the Experiential Account, the fact that blue resembles purple more than green amounts to the fact that they are related to things – experiences or brain states – that fall into this resemblance-order. Against this, it seems perceptually obvious that the fact that blue resembles purple more than green obtains by virtue of the intrinsic character of the colours. To see that the colours fall into this resemblance-order, you only need to focus on the colours themselves.

We may distinguish between first-order and second-order claims about colours. Roughly, first-order claims about colour attribute colours to objects (the table is red) while second-order claims attribute second-order properties and relations to colours themselves (the colour red is unitary). Many philosophers have advocated response-dependent accounts of first-order claims about colours, holding that for an object to have a certain colour is just for it to be disposed to produce experiences of a certain kind in us. In fact, many philosophers have taken the view that colours are simply identical with dispositions to produce experiences. Call this (reductive) Dispositionalism about colour. One objection against Dispositionalism is phenomenological. In looking red, an apple doesn’t look disposed to produce reddish colour experiences in us [McGinn 1996]. Defenders of the Experiential Account reject Dispositionalism in the first-order case but accept something like it in the second-order case. What I pointing out is that the phenomenological objection applies with equal force in the second-order case. 14

---

13 Here I say that the colour orange – a property (a universal or trope) - looks a certain way. Against this, Byrne [2003] says that only particulars can look (in the phenomenal sense) to us to have properties, never properties themselves. Let’s call this Strong Particularism about perception. I think that this is false. We often say that colours, for instance, look certain ways to us – saturated, or bright, or reddish – and an error theory of such talk seems to me implausible.

There is another view one might take. On this view, statements about how colours look can be true, but they are analyzable in terms of statements about how particulars look. Thus, the statement that purple looks reddish is analysed as: everything that looks purple looks reddish. Call this Weak Particularism. (It may be that this is really Byrne’s view [see Byrne 2003: 31].) What I say is compatible with this view.

14 Several have raised criticisms against the appeal to phenomenology against Dispositionalism in the first-order case, and it might be thought that these criticisms also apply to my appeal to phenomenology in the second-order case. I find the criticisms unconvincing, but cannot discuss the matter here.
My fifth and final objection is that the Experiential Account gets the truth-values of certain modal statements wrong. Grant, as defenders of the Experiential Account maintain, that the colour purple is identical with the reflectance-type $R_p$. Then, because $R_p$ might have caused unitary instead of binary experiences under optimal conditions owing to differences in our neural wiring, the colour purple might have caused unitary instead of binary experiences under optimal conditions [Author 200X forthcoming]. For instance, it might have caused unitary reddish experiences under optimal conditions. This, together with the Experiential Account, implies that purple – that very quality – might have been unitary instead of binary! By similar reasoning, on the Experiential Account, blue might have failed to resemble purple more than green.

In general, since the defenders of the Experiential Account analyse the structural properties of a colour in terms of its perceptual effects on us, and since they are committed to saying that those effects might have been different, they are committed to saying that the structural properties of any given colour might have different than they in fact are. But this is not the case. It is clear that every shade of purple is necessarily a bit reddish and bit bluish, and so is necessarily a binary colour. It is equally clear that blue could not have failed to resemble purple more than green. In general, colours have their structural properties essentially. Call this Necessity. Therefore the Experiential Account must be wrong. Following Jonathan Cohen [2003], let us call this the argument from necessity. (Cohen takes it to be an argument against Physicalism whereas I use it as an argument against the conjunction of Physicalism and the Experiential Account.)

This, I think, is a very strong argument. I know of three replies that defenders of the Experiential Account have made or could make, but I believe that they are unsuccessful.

(i) Cohen [2003: 91–2] agrees, or at least is prepared to agree, that if the Experiential Account is correct, then Necessity is false. He suggests the response of biting the bullet and concluding that Necessity is false.

In my opinion, this is a case of failing to apply modus tollens. We have a very strong intuition that Necessity is true. In fact, I would maintain that we can be more certain of the truth of Necessity than of the truth of any controversial philosophical theory of colour. Therefore, if we are forced to choose between the Experiential Account and Necessity, the proper choice is to hold on to Necessity and to reject the Experiential Account.

Cohen [2003: 91] addresses the point that intuition strongly favors Necessity. In response, he says that the Physicalist about colour (I take it he means Physicalists about colour who accept the Experiential Account) “does not share (or any case does not accept) the intuition in question”. In other words, Cohen points out that the defender of the Experiential Account does not accept the intuition: for Experiential Account implies that Necessity is false and hence that the intuition in favor of Necessity is incorrect. Consequently, he says, “it is question-begging to use that very disputed intuition in an argument against her”.

Is this an adequate response? To begin with, Cohen suggests that defenders of the Experiential Account do not share the intuition in favor of Necessity. In my opinion this is a dubious claim, at least if it means that they do not feel the pull of the intuition. If a defender of the Experiential Account claimed that he does not feel the pull of the intuition in favor of Necessity (not merely that feels the pull of the intuition but rejects it – unreasonably in my opinion – on the basis of his theory), I would doubt his sincerity. The pull of the intuition is undeniable and very strong. This gives us very good reason to think that the intuition is correct, and hence, that the Experiential Account is mistaken.

To turn to Cohen’s principal charge – that of begging the question - Cohen says that it is question-begging to appeal the intuition to argue against the Experiential Account on the grounds that, if the intuition is correct, then the Experiential Account is false, so that those who accept the Experiential Account will reject the intuition. But this is not enough to show that it is question-begging (at least, question-begging in some probative sense) to use the intuition to arrive at the conclusion that the Experiential Account is false. Otherwise, whenever we have an intuition that is in conflict with a philosophical theory, it is
question-begging (and therefore, illegitimate) to use that intuition to argue against the theory, simply because it is in conflict with the theory, and so would be rejected by proponents of the theory. For instance, by this reasoning, we must say that it was question-begging (and therefore, illegitimate) of Kripke to appeal to our modal and semantic intuitions to arrive at the conclusion that the Description Theory of Proper Names is false, because if the Description Theory is correct then those intuitions must be false, so that those who accept the Description Theory do not accept the intuitions in question! Indeed, by this reasoning, any argument against any claim is question-begging, on the grounds that if the argument is sound then the claim is false. This is not the case. Consequently I think that Cohen has not said enough to show that it is question-begging (at least in a sense that would imply that it is illegitimate) to appeal to the intuition. 15

(ii) Cohen also attempts to explain away our intuition in favor of Necessity. Necessity says that colors have their structural properties necessarily. But Cohen suggests that it is mistake to think that intuition supports this claim. Instead, he claims, intuition supports the claim colour experiences have their structural properties necessarily. This intuition, he says, is correct. But we misreport this intuition as the intuition that colors have their structural properties necessarily. That is because we are prone to think of properties of colour experiences as properties of colours. However, Cohen says, the claim that colours have their structural properties necessarily is a distinct claim which is not itself supported by intuition. So, we are duped into thinking that intuition supports Necessity when this is not really the case. 16

In my view, this is not a credible story. To see this, suppose that you are looking at a blue thing, a purple thing, and a green thing. You focus on their colours and report the intuition that that the first colour necessarily resembles the second more than the third, that the second is necessarily binary, and so on. According to the reply, you misreport your intuition: the true subject of your intuition is not the colours you experience but your experiences of them; and if it seems to you that the subject is the colours it is because you are in the grips of some kind of confusion. But this is just not right. Your attention was focused all the while on the colours. It was focused out, not in. It was focused on what you experience, not on your experiences. I think that this is just obvious, but if proof is needed, consider the following. Colour

---

15 Of course, it is possible to imagine conditions in which it would be unreasonable to appeal to an intuition to arrive at the conclusion that a philosophical theory is mistaken. Suppose that a theory \( T \) is incompatible with intuition \( I \), but that it accommodates many of our other intuitions, that it is systematic, and in general, that is overall the best theory available. Under these conditions, it might be reasonable to suppose that \( T \) is true and that our intuition \( I \) is unsound – even if we cannot identify the source of the error. Then, under these conditions, it is not reasonable to use \( I \) to conclude that \( T \) is false, although I would not call this case of ‘begging the question’. (Physicalists about the mind claim, or should claim, that these conditions hold where \( T \) is Physicalism and \( I \) is our intuition that Zombies are possible.) But Cohen has not been shown that these conditions hold where \( T \) is the conjunction of Physicalism and the Experiential Account and \( I \) is our intuition that Necessity is true. (In my option, one cannot show this, for the overall best theory of colour is not Physicalism but a version Primitivism – see my concluding remarks.)

16 Let Revelation be the doctrine that, for every property \( P \), if a colour has \( P \) essentially, then we can know on the basis of visual experience that \( C \) has \( P \) essentially. If Physicalism is true, then the property which is in fact the colour purple has essentially the property of being identical with reflectance-type \( R \). But we don’t know this on the basis of visual experience. So if Revelation is true, then Physicalism is false.

McLaughlin [2003: sect. 2] says about Revelation what Cohen says about Necessity: we are duped into thinking that it holds for colours, because we have the intuition that it holds for colour experiences, and we somehow confuse colours with our experiences of them. (But there is a difference. Cohen says that Necessity is true about colour experiences but not colours. By contrast, McLaughlin says that Revelation is false about both colour experiences and colours.)

It is worth mentioning here that nothing I say depends on Revelation. I claim that, for some properties that colours have essentially, we know on the basis of visual experience that colours have them essentially: namely the structural properties of colours. I do not claim that this holds for all properties, and nothing I say requires this.
experiences are properties of people. But you were certainly not attending to properties of people: you were attending to properties that qualify the objects that you experience or nothing at all. And your intuition concerned these properties. Hence, pace Cohen, intuition does support Necessity: the claim that colours have their structural properties necessarily.

It might be thought that this objection relies on taking the transparency observation seriously and advocating Relationism about the structural properties of colour experience. This is not the case. On Relationism, there is a sense in which there no such thing as focusing on a colour experience as distinct from focusing on colour experienced in having that colour experience. So if Relationism is correct, the objection is immediate. But the objection does not depend on Relationism. Suppose we reject Relationism and say that at least in some cases there is such a thing as focusing on our colour experiences and their structural properties, as distinct from the colours that we experience and their structural properties. Still, even on this view, there is such a thing as focusing on the colours, as distinct from our colour experiences. This is an obvious fact that any theory of colour experience must respect. So the objection may still be made. I claim that, when we form our intuition in favor of Necessity, we are focusing on the colours, not our colour experiences; and hence that the true subject of the intuition is definitely the colours. We are not duped.

I conclude that Cohen has not undermined the argument from Necessity. He has not provided any reason to doubt our intuition in favor of Necessity, and he has not shown that it is unreasonable to appeal to this intuition to reject the Experiential Account.

(iii) Brian McLaughlin [2003: sect. 9] makes some remarks which suggest a more conciliatory response to the argument from Necessity than is given by Cohen. (Cohen [2003: note 23] briefly suggests, but does not develop, an analogous response.) On this response, the first step in that argument is incorrect. It is not the case that, if the Experiential Account is true, then Necessity is false. Even if the Experiential Account is true, Necessity might be true (or, at least, it might be true in a sense, as we will soon see). Consequently defenders of the Experiential Account don’t have to reject Necessity after all. They can have their cake and eat it too. Unlike Cohen’s response, this response does not require the rejection of Necessity.

The response depends on a controversial theory of the semantics of colours terms. (See McLaughlin [2001, 2003]. Many others defend this view, with differences of detail, including Jackson and Pargetter [1987] and Lewis [1997]. I will focus on McLaughlin’s version.) According to this theory, colour terms are synonymous with definite descriptions to do with the production of colour experiences. For instance, ‘purple’ is synonymous with ‘the property that causes purple (that is, reddish-bluish) experiences under normal conditions’. These descriptions ‘give the meaning’ the colour terms; they do not merely fix their reference [McLaughlin 2001: sect. 4]. This view about colour terms is obviously analogous to Lewis-style functionalism about the meaning of mental terms. Let us call it the Description Theory of Colour Terms. The Description Theory is used as a premise in an argument for Physicalism about colour, parallel to Lewis’s famous argument for the identity theory of mental states: since by conceptual analysis colour terms refer to the properties that optimally cause our colour experiences, and since as a matter of empirical fact reflectance-types are those properties, it follows that colours are reflectance-types.

To see how the Description Theory might provide a response to the argument from Necessity, consider an instance of the claim that colours have their structural properties necessarily, for example:

[5] Necessarily, purple is reddish-bluish

By the Experiential Account, for a colour to be reddish-bluish is for it to produce reddish-bluish experiences under optimal conditions. So [5] is equivalent to
[5a] Necessarily, purple causes reddish-bluish experiences under optimal conditions.

Now enter the Description Theory. By the Description Theory, ‘purple’ is semantically equivalent to the description ‘the property that causes reddish-bluish experiences under optimal conditions’. So we may replace the occurrence of ‘purple’ in [5a] with this description, to obtain:

[5b] Necessarily, the property that causes reddish-bluish experiences under normal conditions, causes reddish-bluish experiences under normal conditions

Thus the Description Theory and the Experiential Account jointly imply that [5] is equivalent to [5b]. Hence, given this combination of views, the question of whether [5] is true becomes the question of whether [5b] is true.

Now [5b] has two readings. It has a wide scope or ‘de re’ reading which we may gloss as follows: the property that in the actual world causes reddish-bluish experiences under normal conditions (namely, \( R_p \)) is such that it in every world \( w \), it causes reddish-bluish experiences under normal conditions in \( w \). This reading makes [5b] false. Psychophysics is contingent. But [5b] also has a narrow scope or ‘de dicto’ reading which we may gloss as follows: for every world \( w \), the property that causes reddish-bluish experiences under normal conditions in \( w \), causes reddish-bluish experiences in \( w \). This reading makes [5b] true (modulo worlds where there is no property that causes reddish-bluish experiences under normal conditions).

Since the Description Theory and the Experiential Account together imply that [5] and [5b] are equivalent, and since [5b] has a reading that makes it true, these views together guarantee that [5] has a reading that makes it true – namely the alleged de dicto reading.

Here is a more informal way to put the point. According to the Description Theory, the semantic value of ‘purple’ relative to a world \( w \) is the property that causes reddish-bluish experiences in \( w \) – for short, the property that plays the purple-role in \( w \). (At least, this is the case if we assume that ‘purple’ is understood ‘non-rigidly’.) According to the Experiential Account, necessarily, if a property plays the purple-role, then it is reddish-bluish. That is just what it is for a property to be reddish-bluish. So ‘purple is reddish-bluish’ is guaranteed to come out true with respect to every world. Different properties play the purple role, and hence deserve the name ‘purple’, with respect to different worlds; but, by the Experiential Account, just by virtue of playing the purple-role, each is reddish-bluish. So, no matter what world we evaluate it at, ‘purple is reddish-bluish’ comes out true.

According to the present response, the same applies to all sentences of the form of [5] which give expression to the intuition that colours have their structural properties necessarily. The Experiential Account and the Description Theory together guarantee that they have a reading, the alleged de dicto reading, that makes them true. But, as McLaughlin says, the necessity is de dicto or de conceptu, not de re [2003: sect. 9]. Call this the de dicto response to the argument from Necessity against the Experiential Account.

Could the de dicto response to the argument from Necessity be correct? I don’t think so, for a couple of reasons. First, it depends essentially on the Description Theory of Colour Names, but I would argue that the Description Theory is false. The objections parallel the modal, semantic and epistemic objections that Kripke raised against the Description Theory of Proper Names [Kripke 1981]. (i) The Description Theory implies that colour terms refer to different properties with respect to different worlds or counterfactual situations: that they are ‘non-rigid’ designators (or least that they have a reading on which this is so). But, intuitively, ‘red’ refers to the same property in every world. The intuition of rigidity applies just as strongly to ‘red’ as it does to ‘Godel’, ‘pain’ and so on [Kripke 1981]. (ii) The Description Theory implies that whether or not
‘purple’ refers in the actual world depends on whether or not there is a property that plays the purple-role in the actual world. If there is such a property, ‘purple’ refers to it; otherwise it does not refer. But, again, intuitively this is not the case. For instance, suppose that it turned out that (‘consider a world as actual in which’) we are brains in vats and our colour experiences are not caused by the properties of external objects, but by the operation of an evil scientist’s computer program. In that case, there would be no property that plays the purple-role, for there would be no such property as the surface property of objects that causes reddish-bluish experiences in us under optimal conditions. So by the Description Theory ‘purple’ would not refer. But, intuitively, ‘purple’, no less than names of primary qualities such as ‘square’, would still refer: it would refer to a property that we are acquainted with in experience that things look to have but do not. This intuition does not arise out of the acceptance of a controversial theory of the nature of colour experience (such as a form of Relationism); rather it is a quite theory-neutral intuition which I think ordinary folk would claim to have. If mere intuition is not enough, and proof is needed of the claim that ‘purple’ would still refer if such a scenario turned out to be actual, consider the following. Intuitively, if it turned out that we are brains in a vat, certain sentences whose truth requires that ‘purple’ refers would still be true. For instance, ‘red resembles purple more than green’ would still be true. But then, contrary to the Description Theory, it must be that ‘purple’ would still refer even if it turned out that we are brains in a vat. (iii) The Description Theory of Colour Names implies that it is a priori that if the colour purple exists, then it is the cause of ‘reddish-bluish’ experiences under optimal conditions, just as a simple version of the Description Theory of Proper Names implies that it is a priori if Aristotle exists then he was the teacher of Alexander. Thus, the Description Theory implies that a certain version of Eliminativism about colours, according to which the colour purple exists but is not a property of physical objects that causes reddish-bluish experiences under optimal conditions (maybe it is a property of nothing at all), is a priori false. Against this, while such a version of Eliminativism might be false, we cannot rule it out a priori. I can think of responses to all of these objections, which I cannot go into here; but in my view the objections are ultimately successful against the Description Theory of Colour Names. If the Description Theory is false, the de dicto response to the argument from Necessity against the Experiential Account cannot even get off the ground.

But there is a more fundamental objection against the present response to the argument from Necessity. Our intuition that colours have their structural properties necessarily is de re. So, Necessity is properly formulated in terms of de re modality. For instance, suppose that we are having a conversation about the colour purple (or a certain determinate shade of purple) and someone says

[6] It could not have failed to be reddish-bluish

Intuition supports the claim that [6] is true in the context. Intuition also supports the following de re modal claims:

[7] The property that is in fact the colour purple is necessarily binary
[8] The colour purple is such that it is necessarily binary

But even if the Description Theory of Colour Names is correct, the Experiential Account implies that [6]-[8] are false. [6] doesn’t even contain the name ‘purple’. So even if The Description Theory is correct, and ‘purple’ is short for the description ‘the property that causes reddish-bluish experiences under optimal conditions’, the de dicto response does not apply in this case. The Experiential Account inescapably implies that [6] is false. (I take the trick of using ‘it’ to make the de re reading compulsory from Lewis [1986: 250], who uses it in a different connection.) Likewise for [7] and [8]. [7] contains the rigidifier ‘actually’, and [8] is so formulated as to favor the de re reading. So even if The Description Theory is true, the Experiential
Account implies that [7] and [8] are false. Suppose that you are looking at a shade of purple. The Experiential Account implies that it – this very same quality - might have failed to be reddish-bluish. But, intuitively, this is not the case.

Although it is a bit more cumbersome, we can also formulate de re modal claims about colour resemblance, for instance

[9] The properties that are in fact blue, purple and green could not have fallen into a different resemblance-order.

Even if The Description Theory is true, the Experiential Account implies that [9] is false. But it is obviously true.

To sum up. Given the Description Theory of Colour Names, the defender of the Experiential Account may be able to secure the truth of [5] (at least, relative to the alleged de dicto reading). But, intuition also supports such de re modal claims about the colours as [6]-[9]. And, even if the Description Theory is true, the Experiential Account implies that these claims are false. So even if the Description Theory is true, the Experiential Account does not accommodate Necessity, properly understood.

IV

I hope to have done two things. First, I have pointed out that the case for Incompatibilism, the key premise of the argument from colour structure, is based on a certain natural understanding of what it takes for colour structure claims to be true – the face value construal. Second, I have argued against a rival account of the truth-conditions of those claims, the Experiential Account. On the Experiential Account, the truth-conditions of colour structure claims are such that those truth-conditions can obtain even if colours are reflectance-types. So, the Experiential Account, if true, shows that Incompatibilism is false. But the Experiential Account is false. Colour structure statements and beliefs simply do not have the truth-conditions which it assigns them. The Experiential Account is circular, it is introspectively incredible, there is no plausible story to be told about how it might be true, it is phenomenologically implausible, and it has patently false modal consequences. Since the defense of Compatibilism we have examined here rests on the Experiential Account, that defense fails.

It might be thought that there is a more plausible compatibilist response to the argument from colour structure. But, although I cannot make the case here, in my view there is not. I believe that the face value construal of colour structure claims is the right construal, and that, given the face value construal, Incompatibilism unavoidably follows. So, the argument from colour structure identifies one major (I would say fatal) cost of Physicalism about colours: an error theory of nearly all of our colour structure claims. I believe that Physicalism about colour has other major costs.

Of course, every philosophical position on colour has its costs. The best we can do is to draw up a balance sheet and see where the balance of consideration tilts. It might be suggested that, when all is said and done, Physicalism about colour, in spite of its problems, will still emerge as the best overall theory of colour. I disagree. In my view, the overall best theory of colour and colour experience has the following components.

(i) Primitivism Colours are simple properties which are not identical either with reflectance-types or dispositions to produce experiences [Campbell 1993; Maund 1995; McGinn 1996]. (Primitivism is,
however, compatible with the view that colours are supervenient on such properties. Campbell and McGinn accept forms of Primitivism of this kind.) On this view of colours, colour structure poses no problem. Colour structure claims may be accommodated at face value (perhaps as reporting primitive facts about the colours).

(ii) Relationism The qualitative characters of our colour experiences, and so their structural properties, are determined by what colours we experience in having those colour experiences. This accommodates the transparency observation. As we saw, defenders of the Experiential Account are constrained to reject Relationism, and so do not respect the transparency observation.

(iii) Internalism While colour experience is essentially relational, it is fixed by the internal physical state of the subject. What colours we experience at a time, and thereby (given Relationism) the structural properties of our experiences, is fixed by our total brain state at that time. This accommodates the opponent process theory (see note 7 and [Author forthcoming]).

(iv) Eliminativism Colours are not properties of our own brains, properties of our own experiences, or properties of external objects. Colours exist but are nowhere instantiated. They live only in the intentional contents of our experiences. (For a defense of a view of this general kind, see [Maund 1995].) 18

But the positive task of defending these claims must wait for another occasion.

REFERENCES

Author Unpublished ms.

18 I should mention that the views (i) – (iii) may also be combined with Realism about colour. For instance, McGinn [1996] holds that colours are simple, irreducible properties; but he also holds that a simple colour, C, supervene with metaphysical necessity on the disposition produce experiences of C in normal subjects under normal conditions. Colours, while not themselves dispositional, supervene on dispositions. Thus, suppose that, while a certain object produces experiences of the primitive colour property red in normal humans under normal conditions, that same object produces experiences of the primitive colour property green in normal Martians under normal conditions. Then, according to McGinn’s view, the object instantiates both the primitive colour property red and the primitive colour property green. That way, both we and the Martians come out right. McGinn is a colour liberal. However, for reasons that I will not go into here, I believe that an Eliminativist version of Primitivism is preferable to McGinn’s Realist version of Primitivism.