

# Colour and consciousness: Untying the metaphysical knot\*

Pär Sundström

Colours and consciousness both present us with metaphysical problems. But what exactly are the problems? According to standard accounts, they are roughly the following. On the one hand, we have reason to believe, about both colour and consciousness, that they are identical with some familiar natural phenomena. But on the other hand, it is hard to see how these identities could obtain. I argue that this is an adequate characterisation of our metaphysical problem of colour, but a mischaracterisation of the problem of consciousness. It mischaracterises the problem by presenting consciousness as more "colour-like" than we have reason to take it to be. The real problem of consciousness is, I suggest, that almost nothing theoretically useful is known about this phenomenon at present. I also explore some implications of this perspective on the problem of consciousness. Given the shape of the problem, I argue that we can't rule out all forms of *eliminativism* about consciousness. Nor can we rule out that future research will close the "explanatory gap" that consciousness gives rise to.

## 1. Introduction

Problems come in different kinds. Consider for example the following two cases.

*The Bank Robbery Case.* A detective investigates a shooting during a bank robbery. The bank robbers were all arrested at the site of the crime. Moreover, the event took place in broad daylight and in the presence of many witnesses, so almost all our questions about it have been answered. In fact, it remains to settle only one thing: the identity of

---

\* I presented a version of this paper at a CUNY Cognitive Science Symposium, and I thank the participants for a rewarding discussion. I recall some especially useful comments from David Rosenthal and Josh Weisberg. I am also grateful to Ned Block, Larry Hardin, Ingvar Johansson, Sten Lindström, Peter Nilsson, Barry Stroud and an anonymous referee for helpful comments and discussions. Research on the paper was funded by The Bank of Sweden Tercentenary Foundation and The Swedish Foundation for International Cooperation in Research and Education (STINT).

the shooter. The puzzle for our detective is that, although we are confident that we have identified every suspect, each one of them fails to match the profile of the shooter. Suspect A is tall whereas the shooter was short. Suspect B is a woman while there is evidence that the shooter was a man. And so on.

*The Case of the Found Heart.* A human heart is found in a public trashcan. Nothing more is known. A detective is assigned the case.

The cases are both puzzling, but in different ways. In the first case, the problem is that a contradictory picture seems to emerge from the available information. The information suggests that the shooter had certain properties, that he is among the arrested individuals, but also that none of these has the properties of the shooter. In the second case, the problem is that there is very little useful information. It is hard to get the case off the ground, because practically nothing is suggested or ruled out by what is known.

If we misunderstand the nature of a problem, we increase the risk of going astray on it. For example, if in the Case of the Found Heart we should take some groundless speculation to heart and mistake it for reliable data, we may rule out possibilities that ought not to be ruled out, or deem certain hypotheses highly likely that are in fact no more likely than many others, given what we know.

In the following, I shall discuss two metaphysical problems: the problem of colour and the problem of consciousness. These problems are often presented in very similar terms. Thus, consider, first, Mark Johnston's account of the problem of colour.

According to Johnston, we know from psychophysics that

the external explanatory causes of our color experiences .... are either non-dispositional microphysical properties, light-dispositions (reflectance or Edwin Land's designator dispositions or something of that sort) or psychological dispositions (dispositions to appear coloured) with microphysical or light-dispositional bases (1992, 139).

Assuming that colours sometimes causally explain our colour experiences, it follows that "we must look among these properties if we are to find the colors" (139). However, Johnston argues that the colours, at least as we pre-theoretically take them to be, can't be

identical with any such properties. Pre-theoretically, we take the natures of colours to be "fully revealed" or "laid bare" in colour experience (138, 139). But no microphysical property or light-disposition is fully revealed in visual experience, so colours can't be identical with any such property (140). Psychological dispositions are, according to Johnston, sometimes thus revealed. However, it is not apparent in visual experience that colours are such dispositions, as it ought to be, if they were, and if experience reveals their entire nature. So colours can't be identical with such dispositions either (140-2).

Nagel gives a similar account of the metaphysical problem of consciousness. Causal considerations leave Nagel convinced that consciousness must be an integral part of the physical world. At the same time, he perceives a striking difference between consciousness and the physical. Anything physical is "objective": it "can be observed and understood from many points of view and by individuals with differing perceptual systems" (1974, 442). But consciousness, which is a matter of what things are like for an organism, is not thus objective: "facts about what it is like *for* an experiencing organism - - are accessible only from one point of view" (ibid.). Because of this difference, Nagel thinks we are unable at present to understand how consciousness could fit into the physical world.

In the same vein, Levine thinks that consciousness presents us with a sort of "Kantian antinomy":

We have excellent reason for thinking that mental phenomena, including conscious experience, must be a species of physical/natural phenomena. On the other hand, we also have excellent reason for thinking conscious experience cannot be captured in physical/natural terms (2001, 9-10).

The root of this puzzle is, according to Levine, that we have a peculiar "substantive and determinate conception" of our own conscious experiences. And when we compare experiences thus conceived with physical/natural phenomena of the brain, it hard to see how they could be one and the same thing (84).<sup>1</sup>

---

<sup>1</sup> Similar presentations of the problem of consciousness can be found in, e.g., McGinn (1989), Chalmers (1996), Kim (1998) and Papineau (2002).

If we trust these accounts, the problems of colour and consciousness are both rather like the Bank Robbery Case above. On the one hand, we have reason to believe, about both phenomena, that they are identical with some familiar "natural" phenomena. But on the other hand, we are puzzled about how these identities could hold. The alleged puzzles are that we have antecedent "profiles" of both colour and consciousness -- we take colours to be something like revealed, non-dispositional, visual qualities, and consciousness to be in some sense "subjective" and intimately connected with our first-person point of view -- and these profiles fail to match any of the phenomena that, according to the causal considerations, colours and consciousness should be identical with.

In my view, this picture is correct only in the case of the colours. The problem of colour is indeed rather like the Bank Robbery Case. But the problem of consciousness is much more like the Case of the Found Heart. The real problem of consciousness is that, although there is plenty of speculation around, there are very few *data* to guide our inquiries. Those who take the problem of consciousness to be like the Bank Robbery Case have, I suspect, already taken some speculation about consciousness on board.

It might of course be a *coincidence* that consciousness is, mistakenly, taken to pose just the same sort of metaphysical problem that colour in fact poses. But I doubt that it is. There is a long tradition of thought according to which colour and consciousness are closely related phenomena. The misperception of the problem of consciousness is, I suspect, a product of this tradition, and that is why I in this paper discuss the two phenomena together. My hope is that the grip of tradition may be broken if we place colour and consciousness side by side, and carefully consider their similarities and differences.

In the next section, I say a little more about the historical background of the topic. I also introduce one crucial assumption for the remaining discussion -- The Separation Thesis. In section 3, I present the problem of colour, and in section 4 the problem of consciousness.

If we improve our understanding of the problem of consciousness, we will be better equipped to assess what may be ruled out and not ruled out in the study of consciousness. In section 5, I consider two such issues. I argue that we can't rule out all forms of

eliminativism about consciousness. Moreover, for all we know at present, the "explanatory gap" between consciousness on the one hand, and physical and functional properties on the other, may well be closed by future research.

## 2. The Galilean legacy and The Separation Thesis

On the face of it, it may seem that colour and consciousness don't have much to do with one another. Colours are properties of things like tomatoes, cucumbers and lemons, while consciousness is a property of us or our experiences. Or so we tend to assume as we go about our everyday lives.

Nevertheless, there is a tradition in science and philosophy of taking colour and consciousness to have quite a lot to do with one another. This tradition goes back at least to Galileo, who thought that, in reality, it is consciousness or the mind rather than tomatoes and lemons that possesses colours:

I think that tastes, odors, colours, and so on are no more than mere names so far as the objects in which we place them is concerned, and that they reside only in the consciousness (*The Assayer*, quoted from Drake 1957, 274).

I believe it's fair to say that the Galilean view of the relation between colour and the mind is the historically received one. It seems to have been embraced by Locke, for example. True, Locke sometimes asserts that colours are properties of objects outside the mind. They are -- like the other secondary qualities -- "powers to produce various sensations in us" (1995, II.viii. 10). However, on a natural interpretation, any disagreement Locke may have with Galileo on this issue concerns *semantics* rather than metaphysics. On this reading, Locke thinks that the *qualities* that tomatoes and cucumbers visually seem to have are instantiated in the mind, but he takes the colour *terms* to stand for the "powers" of tomatoes and cucumbers to cause these qualities to be

so instantiated.<sup>2</sup> In the 20th century, philosophers from Russell (1912) to Frank Jackson (1977) similarly took the colours we are confronted with in perception to be properties of mind-dependent "sense-data", allegedly distinct from objects like tomatoes and lemons and their surfaces. At present, the Galilean view is often voiced by colour scientists.<sup>3</sup>

However, in recent analytic philosophy, there has been quite a bit of dissent from the Galilean view. A number of philosophers have argued that Galileo was wrong, not just to use colour *terms* exclusively for properties of the mind, but in taking the *qualities* we experience in physical things to be properties of the mind. In the words of David Armstrong, who has been a trailblazer for this view:

I think that the only plausible way that a Materialist [like Armstrong] can deal with the secondary qualities is completely to reverse the whole programme started by Galileo, a programme that has persisted for so long. What we should do is put these qualities back into the physical world again (1999, 124).

The same reversal of the Galilean program has been argued for by Hilbert (1987), Harman (1990), Tye (1995) and (2000), Dretske (1995) and Byrne and Hilbert (2003a), and adopted by many others. Meanwhile, the Galilean view has enjoyed rather little defence in recent analytic philosophy.<sup>4</sup>

One can identify at least three convictions that have contributed to sway many philosophers from Galileanism. *One* is that phenomenological evidence tells strongly against the view. Consider a visual experience of a ripe tomato. If I examine this experience, it seems that the only place where I can ever find redness is on the tomato. I'm certainly not *aware* of anything other than the tomato being red. And that seems true quite generally. Anything that ever *seems* red to me will be a physical object or region of space outside the mind, not an experience (see e.g. Harman 1990, 39, Tye 1995, 30f., and 2000, 45f.) *Another* conviction is that Galileanism can't be established by any "argument

---

<sup>2</sup> This point is made by Shoemaker (1990, 97-8). Locke does not completely depart from Galileo's use of colour terms however. He seems to use the terms *both* for the qualities in the mind and for the powers in objects to produce these qualities.

<sup>3</sup> See, e.g., Palmer: "Color is a *psychological* property of our visual experiences when we look at objects and lights, not a *physical* property of those objects or lights" (1999, 95). For several other references, see Byrne and Hilbert (2003a, sect. 1).

from illusion". The fact that I may experience a certain quality even though nothing in my presence has that quality does not show that the quality is a property of my experience. I can have an illusory experience of an *oasis*; that does not imply that any experience has the property of being an oasis. Similarly, it does not follow that any experience has a perceived *quality* from the fact that experiences of that quality may be illusory (see, e.g., Harman 1990, 35-9, Byrne and Hilbert 2003a, sect. 1.3.). A *third* conviction is that sentient creatures are just as thoroughly physical objects as are tomatoes and cucumbers. Therefore, no argument based on the idea that thoroughly physical objects can't have colour qualities can show that those qualities are instantiated in the mind rather than outside of it (see, e.g., Armstrong 1993, 272, 1999, 75).

I think these are excellent reasons for abandoning Galileanism. In this paper, I will not elaborate on them, but simply *assume* a non-Galilean metaphysics of colour and consciousness. More precisely, I shall assume the following:

*The Separation Thesis:* The qualities that are salient in our colour experience and that we experience physical things like tomatoes and cucumbers as having are, if they are properties of anything at all, properties of such physical things.<sup>5</sup>

A clarificatory remark: The thesis, and the motivations for it that I just mentioned, may call to the reader's mind the so-called "representationalist" or "intentionalist" theory of consciousness, according to which facts about what experiences *are like* are identical with or supervene on facts about what experiences *represent*.<sup>6</sup> It is true that contemporary representationalists tend to accept The Separation Thesis, and take this to be an important component of their view. But it should be noted that The Separation

---

<sup>4</sup> One exception is Boghossian and Velleman's (1989), where a version of the Galilean view is defended.

<sup>5</sup> I'm not sure to what degree The Separation Thesis is controversial among contemporary philosophers. One reason I'm uncertain is the following. Somebody's assent to a sentence like 'only physical things outside the mind are coloured' is compatible with the person holding either (a) the "Lockean" view that the qualities that are salient in colour experience are (somehow) *in* the mind while colour *terms* refer to properties of things *outside* the mind, or (b) the Armstrongian view that the very *qualities* are outside the mind. Regrettably, many philosophers don't disambiguate their views on this point.

<sup>6</sup> Representationalism is defended by, e.g., Harman (1990), Shoemaker (1994), Tye (1995) and (2000), Dretske (1995), Byrne (2001b) and Jackson (2003). Lycan (1996) defends a qualified form of the view. An important precursor of the view was Moore (1903).

Thesis carries no commitment to representationalism. You can endorse The Separation Thesis while denying that what an experience is like is a function of what it represents. To see this, consider the hypothesis that differences with respect to *attention* may determine differences in what an experience is like while leaving its representational content unaffected.<sup>7</sup> That hypothesis is clearly compatible with The Separation Thesis, but it is incompatible with representationalism.<sup>8</sup>

The Separation Thesis will be important to my discussion. I will make another, insignificant assumption as well, namely that Galileo's use of colour *terms* is OK. Thus, I will take terms like 'red', 'yellow' and 'colour' to refer to the qualities that are salient in colour experience, and to no other properties, regardless of the verdict of what, if anything, instantiates these qualities. This practise will be harmless. Any reader who feels uncomfortable with it can at any point read 'red' as shorthand for (say) '*phenomenal red*', 'yellow' as shorthand for '*phenomenal yellow*', and so on.

The Separation Thesis deserves its name because it says that colour and consciousness are, in a certain sense, *separate*. According to the thesis, colours are instantiated either nowhere or in physical things outside the mind. Either way, they are not properties or aspects or residents of consciousness or of anything else mental.

But even if one accepts this thesis, one may still think that colour and consciousness are *similar* to one another in some salient respect. And this seems to be what many in fact think. It is often suggested -- with varying degrees of explicitness -- that consciousness is made up, at least in part, by a family of qualities that are in some phenomenologically salient respect similar to the colours and the other qualities we experience objects outside the mind to have. I will call this the *Two Sets of Qualities View* -- or the *2Q-View* for short.

Such a view is suggested by Levine, for example. According to Levine, physical things outside the mind have, or at least appear to have "objective colours", and it is a genuine problem to see how these might fit into the physical world. But in addition, an

---

<sup>7</sup> Chalmers (2004, sect. 3) develops such a hypothesis (without endorsing it).

<sup>8</sup> In fact, there is independence in the other direction as well: representationalism does not entail The Separation Thesis. To see this, consider the view that colours are properties of mind-dependent sense-data, distinct from physical objects and their surfaces. That view is compatible with representationalism but incompatible with The Separation Thesis. What an experience is like may be determined by its



*experience* of objective colour has itself a "colour-like property", which Levine calls a "subjective colour" or a "colour *quale*", and this property, he thinks, poses yet another significant problem for physicalism (1998, 419, 426).

Some passages in Nagel suggest a 2Q-View as well. According to Nagel, progress in science allows us to understand the qualities that seem to characterise our environment as mere ways in which the world presents itself to us. Once this is realised, we can adopt a less qualitative, more objective conception of the environment. However, when we direct our inquiry towards our own minds, we find, Nagel seems to say, *yet another* set of readily identifiable qualities, and these qualities are harder to understand from an objective point of view than are the qualities of the external environment:

It is impossible to exclude the *phenomenological* features of *experience* from a reduction in the same way that one excludes the *phenomenal* features of an *ordinary substance* from a physical or chemical reduction of it (1974, 437; italics added).<sup>9</sup>

The 2Q-View should not be confused with Galileanism. Galileanism says that the qualities that are salient in our colour experience and that we experience physical things like tomatoes and cucumbers as having, belong to the mind and not to tomatoes or cucumbers. The 2Q-View is that, in addition to those qualities -- which I here call 'colours' -- we are or can be aware in colour experience of *yet another* set of qualities, which belong to experience. I shall understand the 2Q-View to be compatible both with the truth and the falsity of Galileanism. The 2Q-View is silent on whether, and if so where, *the colours* are instantiated. It only insists that colour experience presents us with *two* sets of qualities. Likewise, Galileanism is compatible both with the truth and the falsity of the 2Q-View. Galileanism is silent on whether colour experience presents us with any experiential qualities other than the colours; it only insists that *the colours* belong to the mind.

---

representational content even if experiences represent such sense-data and their properties. (This point is made by Byrne 2001b, sect. 6.)

<sup>9</sup> Shoemaker (1994, 25-6) describes -- but does not endorse -- a version of the 2Q-View under the name "figurative projectivism".

If consciousness were, as the 2Q-View says, (partly) constituted by "colour-like" qualities, it would be unsurprising if it posed a metaphysical problem very similar to the problem of colour. However, I shall argue in the following that we should at least not *assume* any such view of consciousness. It is not a *datum* that conscious experiences possess of family a qualities distinct from and yet in some salient way like the colours and the other qualities that we experience as belonging to external things. Once this is recognised, I think we are in a position to see that the real problem of consciousness is different from the problem formulated by Nagel and Levine. The most immediate task, however, is to bring the problem of colour into sharp focus.

### 3. The problem of colour

Consider the following three claims.

(C-Realism) Some physical things outside the mind are coloured.

(C-Parsimony) If a physical thing outside the mind is coloured, then its colour is identical with one of the physical or dispositional properties the thing uncontroversially has.

(C-Nonidentity) No colour is identical with any physical or dispositional property that a thing outside the mind uncontroversially has.

The claims are jointly inconsistent. That wouldn't be much of a problem if at least one of them were obviously false. However, none of them is obviously false. For each one of them, there is one or more than one reason to believe it is true. That is the -- or at least a - - metaphysical problem of colour.<sup>10</sup>

---

<sup>10</sup> Each choice we can make here corresponds to a well-known position in colour metaphysics.

Sacrificing (C-Realism) amounts to what is sometimes called *eliminativism* about colour. Galileanism, according to which colours in some sense belong to the mind, is one version of this view. For references, see section 2 above. According to another version of eliminativism, *nothing* -- either inside or outside the realm of the mental -- is coloured. As I understand them, this view is held by Mackie (1976) and Hardin (1993).

What are the arguments for the claims? (C-Parsimony) is often defended on causal grounds. Thus, for example, we saw above that Johnston thinks colours must be "non-dispositional microphysical properties, light-dispositions ... or psychological dispositions" on the ground that no other properties are among "the explanatory causes of our colour experiences (1992, 139).<sup>11</sup> It is less clear why philosophers so often endorse (C-Realism); defences of this claim tend to be unarticulated, if they are provided at all. But one sometimes cited consideration is this: *that some physical things are coloured* is a proposition of the "Moorean" kind that is epistemically more secure than any of the theoretical premises needed to contest it. This argument may be what Lewis has in mind:

it is a Moorean fact that there are colours rightly so-called. Deny it, and the most credible explanation of your denial is that you are in the grip of some philosophical (or scientific) error (1997, 325).<sup>12</sup>

---

Sacrificing (C-Parsimony) amounts to what is sometimes called *primitivism* or *the simple view* of colour. For defence, see Campbell (1993), Yablo (1995), and McGinn (1996). Stroud (2000) offers a qualified defence of this view.

Sacrificing (C-Nonidentity) amounts to either *physicalism* or *dispositionalism* about colour. For defences of physicalism, see Smart (1975), Armstrong (1993, chap. 12; 1999, chap. 11), Hilbert (1987), Byrne and Hilbert (1997b; 2003a), Dretske (1995, chap. 3), Tye (1995, sect. 5.3; 2000, chap. 7), Bradley and Tye (2001), Lewis (1997), and Jackson (1998, chap. 4). For defence of dispositionalism, see McGinn (1983), McDowell (1985), Johnston (1992), Shoemaker (1994) and Harman (1996). Cohen (2003) and McLaughlin (2003) have recently proposed "functionalist" views of the colours. I shall not, however, give these views separate treatment here.

Note that *dispositionalism*, as I use the term, is different from Locke's view, which is often called by that name. On the interpretation offered in section 2 above, Locke held that the qualities that are salient in colour vision belong to the mind, while dispositional properties of objects deserve the colour *names*. As I use terms here, this is a form of eliminativism about colour. Contrary to Locke, the "dispositionalists" mentioned above -- as I understand them -- accept (C-Realism) and hold that the very *qualities* in question are identical with dispositions of physical objects outside the mind to cause in us certain kinds of experiences.

<sup>11</sup> Causal arguments for (C-Parsimony) are also given in Mackie (1976, sect. 1.2), Jackson (1977, sect. 5.3-5.4; 1998, chap. 4), Shoemaker (1990, 104), and Bradley and Tye (2001, 470-1). For critical discussions of such arguments, see Broackes (1992), Campbell (1993), Yablo (1995), and Byrne and Hilbert (forthcoming, sect. 4.1). The debate about such arguments in the case of colour is embedded in a more general debate about whether higher-level or supervenient properties can be causally efficient or relevant to causal explanation. For some contributions to this more general issue, see e.g., Steward (1997) and Kim (1998).

<sup>12</sup> An alternative line of argument in favour of (C-Realism) is that we can make sense of our *representations* of physical things outside the mind as having colour qualities only if we take these qualities to be sometimes instantiated in these things. A version of such an argument is given by Stroud (2000, chap. 7).

I shall remain neutral on what force the arguments for (C-Parsimony) and (C-Realism) have, and instead focus on the -- in my mind rather considerable -- case that can be made for (C-Nonidentity). Before I present the case, let me take a moment to explain why it is of interest in the present context.

Given a little reflection or prompting, almost anyone will sense that there is *something* -- something "qualitative" -- in the vicinity of colour, perception and/or consciousness that seems to be extraordinarily hard to identify with any of the "naturalistically respectable" properties that things or states or events in the world uncontroversially have. Often, this qualitative something is taken to be an aspect of *consciousness*. Qualitative aspects of consciousness, it is felt, pose an incomparably hard problem for a naturalistic, reductionist account of the world.

I shall try to bring out, however, that it is quite hard to see how *colours* could be identical with any of the physical or dispositional properties that things outside the mind uncontroversially have. And if one becomes, like me, impressed by how hard *this* problem is, then one may begin to doubt that we face an *additional* problem -- which is of the same kind, just harder -- of seeing how any qualitative aspects of *consciousness* could be identical with any uncontroversially instantiated, naturalistically respectable properties. In my view, we face no such additional problem. This is not to say that consciousness is not mysterious, but it is to say that consciousness is not mysterious *in that way*.<sup>13</sup> This is to be elaborated in the next section, however. Right now, we should keep our focus on the case for (C-Nonidentity).

What, then, is the obstacle to identifying the colours with some set of physical or dispositional properties that things in our environment uncontroversially have? We saw above that, according to Johnston, the root of the difficulty is that colours seem to be "fully revealed" to visual experience. I agree that this is the -- or at least a -- central difficulty, and I shall try to spell out somewhat more fully what I take it to consist in.

---

<sup>13</sup> Alex Byrne (2006) has recently argued that colours present a serious metaphysical difficulty and that once this is appreciated, consciousness ceases to be problematic: "there is a 'hard problem' of colour ... [and] once we recognize the source of the puzzlement, the mind-body problem disappears" (ibid., last paragraph of the paper). Byrne's line of argument is somewhat similar to mine, but I don't share his conclusion. I think consciousness remains a problem; it's just not the *kind* of problem it is often taken to be.

As Byrne (2001a, sect. 3) points out, the idea that colours are "fully revealed" in visual experience may usefully be dissected into two theses, which I propose to formulate thus:<sup>14</sup>

(C-Infallibility) For any proposition  $p$  that concerns the nature of a set of colours  $C1...Cn$ : if it is reasonable to judge that  $p$  on the basis of visual acquaintance with  $C1...Cn$ , then  $p$  is true.

(C-Self-intimation) For any proposition  $p$  that concerns the nature of a set of colours  $C1...Cn$ : if  $p$  is true, then  $p$  can be confirmed merely on the basis of visual acquaintance with  $C1...Cn$ .

A note about the "nature" of colours: I assume that we have a shared, intuitive sense of a distinction between propositions that concern the "nature" of a set of colours and propositions that don't. Thus, we presumably share a sense that the proposition *shades of orange differ from and resemble one another in three dimensions (hue, saturation and brightness)* concerns the nature of these shades, while the proposition *a given shade of orange is the colour of the Dutch national soccer team* does not concern the nature of that shade. I don't know how to improve on this intuition, so I won't try.

I shall make a partial defence of (C-Self-intimation) here. I hope to convince the reader that the claim deserves serious consideration, at least. For two reasons, I won't bother to defend (C-Infallibility). First, (C-Infallibility) should be somewhat less controversial than (C-Self-intimation).<sup>15</sup> So if we find reason to believe the latter, we have some reason to believe the former. Second and more importantly, I will rely less heavily on (C-Infallibility) than on (C-Self-intimation) in exploring difficulties for reductionism about colour.

---

<sup>14</sup> The wording of this compound formulation differs from Byrne's, but the content is very similar. For a discussion of a slightly stronger pair of theses, see Byrne and Hilbert (forthcoming, sect. 2.2).

<sup>15</sup> The only scenario in which (C-Self-intimation) is true but (C-Infallibility) false is where everything that is essentially true about colours is available to us in perception, and where, in addition, we are presented with some persistent illusions about their natures. That would be somewhat surprising. The scenario in which (C-Infallibility) is true but (C-Self-intimation) false would be less surprising. That is the scenario in which we don't suffer any illusions about the natures of the colours, but where some truths about their natures are unavailable to us in perception.

The case in favour of (C-Self-intimation) is a rather shameless appeal to intuition. I ask you to focus your gaze on one of those qualities that are salient in colour vision. Now consider whether there may be something in the nature of *that very quality* that is hidden from you. Remember that it's a separate, and for our purposes irrelevant, question whether the quality is a referent of any colour *term*. In this paper, I have *stipulated* that it is, but that stipulation is dispensable and insignificant. Whatever the quality is called, the question is whether there may be any essential aspect of *it* that could not be detected on the basis of visual experience alone. The intuition in favour of (C-Self-intimation) is that the quality itself could not have such a hidden nature.

I myself find this intuition rather powerful, and many others philosophers do too.<sup>16</sup> I also think the intuition survives quite a bit of critical reflection. I will illustrate this by considering two attempts to deflate it.

Armstrong responds as follows to the claim that colours are fully revealed in experiences of them:

Against this, I would once again advance the idea ... that everything in the world, *everything*, every event, every property of things and events, every relation that things and events have to each other, are each one of them an epistemological *iceberg*. Our knowledge and rational belief about all these things, though real, is selective and limited. If you take this view then it becomes easier to accept that the secondary qualities might have hidden depths to which we cannot penetrate in perception (1999, 129).

Of course, if you take Armstrong's view, it is not just "easier to accept" that colours have hidden depths. You are *forced* to accept this, since it *follows* from the view that *all* properties have hidden depths.

Armstrong does not, in this passage, tell us why we should conclude, from the truth of his general assertion, that (C-Self-intimation) is false rather than concluding, on the basis of the truth of (C-Self-intimation) that his general assertion is false. But there is some

---

<sup>16</sup> See, e.g., Russell, (1912, 46-7), Strawson (1989, 224), Johnston (1992, 138), McGinn (1996, 541), Maund (2002), Campbell (1993, 178) and Chalmers (2006, sect. 6).

indication that he has in mind an inductive argument of the following kind.<sup>17</sup> In the course of the history of science, all sorts of events, properties and relations have turned out to have natures that are partly hidden from our perceptions of them. Therefore, colours too will turn out to have such partly hidden natures.

I am not moved by this inductive argument, for two reasons. First, it is not clear to me that the historical evidence on which the induction is based is all that impressive. It may be that progress in science has revealed hidden natures in many things, including water, gold and lightning. But there still seem to be a number of properties and relations where we have not found, and are unlikely to find in the future, a nature hidden from our perception. Think of properties and relations like *being thin*, *being hollow* and *being shorter than*. To be sure, such properties may present us with various problems --like vagueness -- but it is not clear that they may turn out to have some nature that goes beyond what is accessible to us given our perceptual acquaintance with them, let alone that they *have* been shown to have such hidden natures.

Second, even if the historical evidence should be more impressive than I just suggested, I have reservations about the induction made on its basis. The intuition in favour of (C-Self-intimation) is surely that there is something *exceptional* about colours - - perhaps together with other sensible qualities. Even if many other properties and relations should have hidden natures, colours -- and perhaps other sensible qualities -- are *special* in the way they wear their natures on their sleeves. And that intuition is unshaken by the historical evidence of the induction.

Brian McLaughlin (2003) attempts in a different way to deflate the intuition that colours are fully revealed in experiences of them. He points out that we need to distinguish two revelation doctrines, one concerning colour, the other concerning experiences, or *what it is like* to have them:

The doctrine of Revelation for colours should be distinguished from the doctrine of Revelation for what it's like to see colours -- the doctrine that the nature of the

---

<sup>17</sup> The indication is that such an argument is given in Armstrong (1999, 29) to show that *the mind* may turn out to have a partly hidden nature.

phenomenal character of a colour experience is revealed to us when we have the experience (99).

McLaughlin suggests that the appeal -- or apparent appeal -- of the doctrine of Revelation about *colour* is generated through a confusion or association of the two. The fundamentally appealing view is Revelation about what it is like to see colour, and the (apparent) appeal of Revelation about colour is generated by being mixed up or associated with that view. How does this happen? McLaughlin's explanation invokes his "functionalist" view of colour, according to which colours earn their status as colours by producing certain phenomenal experiences in us. Because colours and phenomenal experiences are thus intimately related to one another, we are apt to associate the two Revelation doctrines with one another.

Whatever intuitive appeal the doctrine of Revelation for colours enjoys is, I believe, due to two factors: our recognition that knowledge of what it's like to see a colour figures centrally in our understanding of what the colour itself as such is, and the powerful intuitive appeal of the doctrine of Revelation for what it's like to see colours (99).

On McLaughlin's account, many of us are, at some level, functionalists -- or at least relationists of a certain kind -- about colour, even if we don't know it. I have doubts about that. But even setting that issue aside, I have a problem with the account. Contrary to McLaughlin, I don't think we have a very resilient intuition that the natures of what experiences are like are fully revealed to us. This is something I shall try to justify in the next section. I shall argue there that the Revelation doctrine for phenomenal consciousness is *less*, not more, resilient to critical reflection than is the Revelation doctrine for colours. Hence, it is implausible that the apparent intuitive appeal of the latter should be due to the more genuine intuitive appeal of the former plus the mixing up or associating of the two. In fact, as far as I'm concerned, the reverse may well be the case.



I'm not a die-hard defender of (C-Self-intimation) myself. I may be willing to abandon the claim at the end of the day. But I need better reasons for doing so than I find with either Armstrong or McLaughlin. Given the intuitive nature of the case, I can only invite the reader to share this assessment.<sup>18</sup>

I now turn to illustrate how (C-Self-intimation) and (C-Infallibility), respectively, can be used in support of (C-Nonidentity).

(C-Self-intimation) is a powerful dialectical tool. In fact, if (C-Self-intimation) is granted, it seems that (C-Nonidentity) can be established conclusively, or almost conclusively. Consider for illustration the following argument against dispositionalism:

---

<sup>18</sup> (C-Self-intimation) is *one* articulation of Johnston's intuition that colours are "revealed" in experiences of them. It should be noted that the literature contains objections to doctrines called 'Revelation' that are quite different from (C-Self-intimation). Thus for example, Lewis recently attacked the doctrine -- which he named 'revelation' -- that each colour has a "simple, ineffable, unique essence that is instantly revealed to each beholder, or anyway to each beholder with normal visual capacities in normal light" (1997, 338). This doctrine is different from (C-Self-intimation) in two respects. First, (C-Self-intimation) is quiet about the natures of the colours. It says that these natures are accessible in colour experiences, but it says nothing about what the natures are. In contrast, the doctrine attacked by Lewis does seem to say something substantial about the natures of colours, namely that they are "simple" and "ineffable". Second, the doctrine attacked by Lewis says that colours are "instantly revealed" to beholders, but (C-Self-intimation) makes the weaker claim that true propositions concerning the natures of colours *can be confirmed* by beholders. Nothing is said about such confirmation being "instant". In fact, (C-Self-intimation) is compatible with such confirmation requiring quite a bit of attentiveness and conceptual sophistication.

- (1) [Assume for reductio] A given colour, *C*, is identical with the disposition, *D*, to cause, in such-and-such observers under such-and-such circumstances, an experience of such-and-such kind.
- (2) It's in the nature of any disposition to be a disposition.
- (3) Hence, it's in the nature of *C* to be a disposition (from (1) and (2)).
- (4) Hence, it should be possible to confirm that *C* is a disposition merely on the basis of visual acquaintance with *C* (from (3) and (C-Self-intimation)).
- (5) But it is not possible to confirm that *C* is a disposition merely on the basis of visual acquaintance with *C*.
- (6) Hence, not-(1).<sup>19</sup>

If (C-Self-intimation) is granted, the argument is very forceful. It is valid, and its only premises are (2) and (5). (5) seems secure: whatever else may be said for the view that colours are dispositions of some sort, it is not a view that can be confirmed solely on the basis of visual acquaintance with colours. And (2) is highly attractive -- at least if one grants any version of the distinction between what's in the nature of a property and what just happens to be true of it. A property that is in fact a disposition would not, it seems, be the very property it is were it not a disposition.<sup>20</sup>

---

<sup>19</sup> This argument is similar to the argument, mentioned above, that Johnston formulates against dispositionalism (1992, 140-2). In that paper, Johnston argues that dispositionalism survives the argument, but only because we may choose to speak of colours "less inclusively" than we normally do. Dispositions may deserve to be called by the colour terms, Johnston argues, even though they are not all that we normally take the colours to be (ibid., 142). In a later, unpublished manuscript, Johnston rejects dispositionalism and adopts, as far as I understand him, a primitivist view of colour (Johnston, unpublished).

<sup>20</sup> There is some wiggling room on this point, but I doubt it is very significant. I think (2) is safe if dispositions are distinct from whatever categorical bases they may have. However, if, as Armstrong has argued, dispositions are *identical* with their categorical bases, then it may not be in the nature of a disposition to be a disposition (see Armstrong 1999, chap. 5, and his contributions in Armstrong et al 1996). In the present context, this option is of limited significance, though. This is because most dispositionalists about colour (perhaps *all* who think that dispositions *have* categorical bases) would take the categorical base of the disposition *D* to be some microphysical property, *P*. And that would -- even assuming Armstrong's view of dispositions -- leave the argument essentially unaffected. It is presumably in the nature of *P* -- and hence, by hypothesis, *D* -- to be whatever microphysical property it is. But it's not possible to confirm that *C* is the way *P* (or *D*) essentially is on the basis of visual acquaintance with *C*. Hence, given (C-Self-Intimation), *C* is not identical with *P* (or *D*). (True, there is logical space for a different view. Take Armstrong's view of dispositions, and suppose further that the categorical base of *D* is, not a microphysical property, but simply the visually given quality, which -- let's assume -- lacks a hidden nature. This view, which is a sort of marriage of dispositionalism and primitivism, is not threatened by (C-Self-intimation). If *C=D* and *D*=this quality, then all we can conclude is that it's in the nature of *C* to

It seems clear, moreover, that any view to the effect that colours are light-dispositional properties or microphysical properties is refuted by a parallel argument, provided that (C-Self-intimation) is granted. Thus, given (C-Self-intimation), (C-Nonidentity) is almost impossible to resist.

(C-Infallibility) is not quite as powerful a tool as (C-Self-intimation) for the purpose of defending (C-Nonidentity), but it can be used to present considerable challenges for physicalists and dispositionalists, at least. I will illustrate with a case against physicalism about colour.

Consider to begin with the following propositions, which is a very small sample of what we may with some confidence assert regarding the colours.

There is a red that is to no degree bluish, yellowish or greenish.<sup>21</sup>

Any orange is to some degree reddish and to some degree yellowish.

Reds differ from and resemble one another in three dimensions (hue, saturation and brightness).

There are blues that are darker than the darkest yellow.

Blues (in some sense) "recede" while reds "come out".

For the purposes of the argument to follow, I want to consider *colour properties* in a broad sense that includes not only properties like red and yellow and specific shades of red and yellow, but also properties and relations like reddish, yellowish, dark, bright, darker than, and so on. I will assume that (C-Infallibility) is true even if we take it to concern colour properties in this broad sense. I don't think that is more controversial than taking it to be true if restricted to properties like red and yellow and specific shades of red and yellow. It seems no more controversial that visual experience provides us with correct information about the natures of *reddish* and *dark* than that it provides us with correct information about the nature of *red*.

---

be the way this quality is. And it *can* be confirmed merely on the basis of visual acquaintance with *C* that *C* is this way. But one may wonder what reason there could be to hold this view; it seems to me to have no advantage over simple, i.e. non-dispositionalist, primitivism. And I don't know anyone who has held it.)

<sup>21</sup> What is bluish? An ostensive definition should suffice. Bluish is what you experience whenever you experience either purple, blue or turquoise. For more on bluish and other "hue magnitudes", see Byrne and

Now here is an argument against colour physicalism:

- (i) For each of the above propositions (and many others like them), it is reasonable to judge that proposition to be true merely on the basis of visual acquaintance with the colour properties it is about.
- (ii) Further, each of these propositions concerns the natures of the colour properties it is about.
- (iii) Hence, the propositions are all true (from (i), (ii) and (C-Infallibility)).
- (iv) However, for any set of physical properties  $P1...Pn$  with which these colour properties may be identified, it is not part of the *natures* of  $P1...Pn$  to be the way that the (true) propositions say that the colour properties are.
- (v) Hence, the colours are not identical with any set of physical properties (from (ii), (iii), (iv) and Leibniz' law).

This argument is again valid. The premises are (i), (ii) and (iv). For my part, I can't see how (i) and (ii) could fail to hold. Nobody should be accused of jumping to conclusions for holding true any of the propositions at issue merely on the basis of visual experiences of the relevant colours. So (i) would seem to be in good standing.<sup>22</sup> (ii) seems secure as well -- or secure enough. There may be room for dispute about individual propositions. For example, perhaps it can be questioned whether it is part of the nature of blues to "recede". But provided that some version of the distinction between what's in the nature of a property and what just happens to be true of it is granted, it seems safe to assume that a significant number of propositions like the ones above concern the natures of the colour properties they are about. Surely it isn't accidentally true of a given shade of orange that it is to some degree reddish and to some degree

---

Hilbert (1997b, sect. 3.2; 2003a, sect. 3.2), Byrne (2003, sect. 7), Tye (2000, sect. 7.4), and Bradley and Tye (2001).

<sup>22</sup> I add one qualification to this. I believe there is room for debate about exactly how we should specify what we are entitled to judge on the basis of visual experience. For example, it is not obvious whether visual experience entitles us to hold that any orange is *identical with* some combination of reddish and yellowish, or rather that any orange *has the property* of being reddish and yellowish (for discussion of this distinction, see Byrne and Hilbert 2003b, sect. 2.3). This also means that there is room for a physicalist to specify the relevant claims in different ways in trying to refute the premise (iv); see discussion of this premise below.

yellowish, in the way that it is accidentally true of it that it was chosen to be the colour of the Dutch national football team.<sup>23</sup>

If I'm right about premises (i) and (ii), the only response available to a physicalist is to challenge premise (iv). This premise is less obviously compelling. But at a minimum, it presents the physicalist with a considerable challenge. I think it's fair to say that no physicalist has, to date, come up with a set of properties that refutes (iv), so nobody knows whether the challenge can be met.<sup>24,25</sup>

I shall by-pass a discussion about what force a parallel argument against dispositionalism may have.<sup>26</sup> My ambition here is not to *refute* reductionism about colour, but just to illustrate the -- in my mind considerable -- difficulties any such view faces. I hope I have accomplished that.

As I have mentioned, many have supposed that consciousness poses a similar problem for reductionism, only harder. My next task is to challenge that view.

---

<sup>23</sup> As I understand him, McLaughlin (2003, 115-6) suggests precisely this: that it *is* accidentally true of orange that it is to some degree reddish and to some degree yellowish -- or at least, that we may find out that this is so. But I fail to regard this as a live possibility.

<sup>24</sup> Bradley and Tye (2001) have made the most impressive attempt to meet this challenge that I'm aware of. They try to show that there are physical properties with which red, orange, reddish, yellowish, bluish, etc., might be identified such that it is true of them -- perhaps even essentially true of them -- that any orange is to some degree reddish and to some degree yellowish while there is a red that is to no degree yellowish and to no degree bluish. I find their account rather persuasive, as far as it goes. But as it stands, it does not provide a physicalistic interpretation of *all* the propositions listed. They say nothing about the relative darkness of blues compared with yellows, for example.

<sup>25</sup> I said above that I would rely less heavily on (C-Infallibility) than on (C-Self-intimation) in exploring the difficulties for reductionism about colour, and this can now be explained. The role of (C-Infallibility) in the above argument is to facilitate the drawing of conclusion (iii). There is a point to running the argument that way. It brings out what sort of ground we have for (iii). If we have reason to believe in (iii) at all, it is because our colour vision informs us that the colours are the way the relevant propositions say they are, and because we may trust the deliverances of vision in this case. However, even though invoking (C-Infallibility) highlights this fact, the claim is rather dispensable for the purposes of the argument. We could also introduce (iii) as an underived premise. And it is not clear that the argument would be less convincing if we did so, because (iii) is arguably no less plausible than (C-Infallibility), as that claim stands: we might remain convinced of (iii) even if we should find some exception to (C-Infallibility).

<sup>26</sup> For an exchange on this issue, see Johnston (1992, 152-3), who argues that parallel arguments don't threaten dispositionalism, and Byrne (2003, sect. 2) who responds that whatever force these sorts of arguments have against physicalism they have against dispositionalism as well.

#### 4. The problem of consciousness

The kind of consciousness I'm interested in here is the kind that philosophers sometimes call 'phenomenal consciousness'. For a person, or state (of a person) or event (involving a person) to be phenomenally conscious is for there to be something it is *like* to be that person, be in that state or be the subject of that event. I shall also talk about "phenomenal consciousness properties". This is a generic term applying both to the determinable property of being phenomenally conscious at all and to determinates of this property -- if there are such. Thus, it may be that what it is like to consciously perceive red is not the same as what it is like to consciously feel pain, but that there is something it is like in both cases. If so, the cases both instantiate the determinable phenomenal consciousness property but different determinate phenomenal consciousness properties. Often, I shall use the terms 'consciousness' or 'phenomenal consciousness' as shorthands for 'phenomenal consciousness properties'.

There are similarities between the metaphysical problem of phenomenal consciousness and the metaphysical problem of colour. Consider again three claims:

- (P-Realism) Some physical "things" (persons, states of persons, or events involving persons) are phenomenally conscious.
- (P-Parsimony) If a physical thing has a phenomenal consciousness property, then that property is identical with one of the physical or functional properties that thing uncontroversially has.
- (P-Nonidentity) Phenomenal consciousness properties are not identical with any of the physical or functional properties that physical things uncontroversially have.

As in the case of colour, there is a problem here, because the claims are jointly inconsistent, yet none of them is obviously false.<sup>27</sup>

---

<sup>27</sup> Again, each choice we can make corresponds to a well-known metaphysical view of consciousness.

Sacrificing (P-Nonidentity) amounts to either *physicalism* or *functionalism* about phenomenal consciousness properties. For defences of *physicalism*, see Searle (1992), Armstrong (1993) and (1999), Lewis (1994), Perry (2001), and Block (2003). *Functionalism* is defended in Harman (1990), Dennett

But if one looks closer at the two problems, one finds important differences as well. The most striking difference may be the status of the two realism claims. (C-Realism) is clearly in need of discursive defence. It is not an indisputable datum of colour metaphysics that some physical things outside the mind are coloured. It is indisputable that we *experience* objects as having colours. But the world isn't always as we experience it to be, so why should we have faith that any physical thing really is coloured? For (C-Realism) to deserve our reflected loyalty, this question needs to be answered.<sup>28</sup> In contrast, it has seemed to many that the existence of phenomenal consciousness is an indisputable datum.

I will return to this issue later. My main occupation, however, will be with differences that concern the non-identity claims. A reductionist about consciousness faces, I claim, a situation very different in kind from the situation faced by a reductionist about colour. There are at least three, partly related differences between the situations.

*First difference: The pre-theoretical data.* In the case of colours, we agree to a significant extent about how they are. For example, we agree that *reds differ from and resemble one another in three dimensions (hue, saturation and brightness)*, that *some blues are darker than the darkest yellow*, and that *any orange is to some degree reddish and to some degree yellowish*. This constrains our theorising about colours, because if these things are true, as we agree they are, then the colour properties can't be identical with any set of properties that fail to be as these propositions say that the colour properties are.

In contrast, we don't have a consensus of a corresponding significance about phenomenal consciousness. There just isn't a set of propositions that we all agree on and that constrain our theorising about consciousness in any comparable way.

---

(1991), Dretske (1995), Tye (1995) and (2000), and Lycan (1996). Shoemaker (1982) defends a hybrid of physicalism and functionalism. Some theorists identify consciousness with other *mental* phenomena while remaining -- at least officially -- neutral about whether these phenomena in turn can be identified with physical or functional properties; see, e.g., Rosenthal (2002).

Sacrificing (P-Parsimony) amounts to *dualism*. For defence, see Kripke (1972), Jackson (1982), and Chalmers (1996).

Sacrificing (P-Realism) amounts to *eliminativism*. It is defended by Rey (1988) and (1995).

<sup>28</sup> Note that the "Moorean" argument offered by, e.g., Lewis (see section 3 above) *is* an answer to this question. It tells us that we should have faith that physical things are coloured, because this kind of common-sense opinion is more secure than any theoretical principle required to refute it.

I don't deny that there are *some* facts about consciousness that we agree on. For example, we can agree on some paradigmatic examples of its presence. Thus, if I'm awake and attentive to a pain I'm having, presumably we all agree that there is something it is like for me to have that pain. There are also some very general characterisations of phenomenal consciousness properties that we by and large agree on, for example that they are properties of persons, states of persons or events involving persons.<sup>29</sup> But such a characterisation is too general to constrain our theorising about these properties in any significant way. It is true of so many properties other than phenomenal consciousness properties that they are properties of either persons, states of persons or events involving persons.

As soon as we move beyond paradigmatic examples and such highly general characterisations, controversy begins. It is a striking feature of the literature on consciousness that it displays such persistent disagreements about what the facts about consciousness are. To mention just a couple of examples: philosophers disagree on whether it is like something to hear a noise even if one pays no attention to it,<sup>30</sup> and, again, whether the phenomenology of *thinking* is always exhausted by sensations and imagery *accompanying* thinking or whether there is such a thing as a purely "cognitive phenomenology".<sup>31</sup>

It has been suggested to me that there *is*, nevertheless, a theoretically significant consensus about how consciousness is -- this consensus being expressible by a set of propositions that exactly match the propositions that constrain theorising about colour. Thus, just as reds differ from and resemble one another in three dimensions (hue, saturation and brightness), so there are three corresponding dimensions in which *what it is like* to see red may vary. And just as any orange is to some degree reddish and to some degree greenish, so *what it is like* to see orange is always in some sense a mix of *what it is like* to see green and *what it is like* to see red. And so on.

---

<sup>29</sup> Even on this seemingly uncontroversial issue, however, agreement appears not to be universal. Thus, Tye denies that the "phenomenal character" of an experience, which is the same as what the experience is like to have, is a *property* of the experience. He insists that it is, rather, a certain kind of representational *content* (1995, 137; 2000, 45; and especially (2002/3, point 2).

<sup>30</sup> See Block (1995, 234) for the view that this is like something, and O'Regan and Noë (2001, sect. 7.2) for the view that it is not.

<sup>31</sup> For the former view, see Tye (2003, sect. 2) and Lormand (1996); for the latter, see Strawson (1994, sect. 1.3-1.4) and Siewert (1998, chap. 8).



This suggestion is a version of the *Two Sets of Qualities View* -- or *2Q-View* -- which I mentioned at the end of section 2. According to this view, consciousness is, at least in part, made up by a family of qualities that are distinct from and yet in some phenomenologically salient way like the colours and the other qualities that we experience external things as having. The present suggestion is one version of this view. It says that the two sets of qualities resemble one another phenomenologically by being *isomorphic*.<sup>32</sup>

Now, I doubt that this, or any other, 2Q-View is correct. I am unable to find, in my own colour experience, any colour-like qualities other than those that seem to characterise objects outside the mind.<sup>33</sup>

However, the issue at present is not the truth of the 2Q-View. My argument is compatible with this view. The present issue is whether we have a body of shared convictions about the -- alleged -- experiential qualities comparable in theoretical significance to our convictions about the colour qualities. And it seems clear to me that we don't. Even those who endorse the 2Q-View should, I think, agree with this assessment. The reason is the following.

---

<sup>32</sup> One can imagine other versions of the view. For example, it could be held that experiential qualities and colour qualities resemble one another in something like the way that the chromatic colours (red, green, blue, yellow, orange, etc.) resemble the achromatic colours (white, black and grey). The chromatic and achromatic colours are not isomorphic. Yet, they seem to resemble one another in a phenomenologically salient way. It seems that black is more similar to red than to the scent of a rose, for example.

<sup>33</sup> In considering this issue, it is important to keep clearly in mind the difference between the 2Q-View and Galileanism. This is important because some phenomena or considerations that might *appear* to support the 2Q-View provide only, as far as I can see, (defeasible) support for Galileanism. Consider, for example, Ned Block's discussion of "phosphene experiences" (2003, 177-8). You can generate phosphene experiences by closing your eyes and pressing lightly on your eye balls for about a minute. You will experience colours floating around in a dark space. Block, as I understand him, takes this experience to support the 2Q-View. While the experience represents, as he says, "colored moving expanses", he suggests that "one can be aware of something more" (177). (And I take it that this "something more" is not the darkness of the space, or the shape or motion of the expanses, or any other quality that we sometimes represent physical objects outside the mind as having. After all, whether "something more" is experienced is supposed to be, even by Block's lights, at least *slightly* controversial.) However, I don't find an *extra* set of qualities in this experience any more than I find it in ordinary experiences of tomatoes and lemons. In both cases, colours (and shapes and motions) are represented. In the normal case, they seem to characterise physical objects outside the mind. In the phosphene experience, they appear to belong to some insubstantial expanses. But in neither case am I, as far as I can tell, acquainted with any qualities other than the ones that sometimes seem to characterise objects outside the mind. If the phosphene experience is of interest to the philosophy of perception at all, it is, I suggest, because it lends (defeasible) support for Galileanism. As Block notes (177), the phosphene experience is unusual in not obviously ascribing colours (or shapes or motions) to mind-independent objects (like tomatoes and lemons). But this would seem to support at most Galileanism, the view that *colours* belong to the mind.

Even if there are experiential qualities of the kind that the 2Q-View supposes, it seems clear that they are pretty elusive, even from the first-person point of view. To illustrate this, notice how difficult it seems to be to answer some fairly straightforward questions about them. Thus, let's suppose an experience of a red tomato (as red) involves an introspectible experiential quality distinct from the colour quality. Where is this alleged quality experienced or felt to be located? Is it experienced as located on the tomato, so that colour vision presents the tomato as having two qualities? Or is it experienced as being somewhere between me and the tomato, or as being somewhere in myself? Does the experienced location of the quality shift with, say, shifts in attention? Or is the quality experienced as not spatially located at all? Again, supposing there is an isomorphism between the colour qualities and the experiential qualities involved in colour vision, to what degree if any do these different qualities resemble one another beyond being isomorphic? Take the experiential quality involved in seeing red -- call it, borrowing some terminology from Peacocke (1983, chap. 1), red'. Is there any similarity between red and red' over and above the similarity they have in virtue of occupying the same "positions" in isomorphic structures? If one were acquainted with only red', red and the scent of a rose, would one be able to tell that red' was the experiential quality associated with red rather than with the scent of a rose?

I assume that the answers to these questions are unclear to anybody, including those who hold some 2Q-View. At least, I have not come across a detailed enough account of experiential qualities to know how anybody would answer them. And this suggests what I take to be true: that, if there are at all "experiential qualities", they are very elusive, even to introspection, and we don't at present know much about them.

But if this is so, then we can hardly be assured that the supposed experiential qualities *are* isomorphic with the colour qualities. (Nor, I suppose, can we be assured of any other particular version of the 2Q-View.) For all anybody can tell, the experiential qualities associated with seeing shades of red might differ from and resemble one another in (many) more dimensions than the reds differ from one another -- or, for that matter, in (considerably) fewer dimensions than the reds differ from one another.<sup>34</sup>

---

<sup>34</sup> At the extreme end of this last suggestion is the view that the property of being phenomenally conscious lacks determinates. On this view, an experience of green never differs from an experience of red with

I take the claim I'm making here -- that we have very little by way of useful pre-theoretical data to go by in the study of consciousness -- to be, if not entirely uncontroversial, at least unoriginal. It seems to match the complaint, often voiced by philosophers of consciousness, that it is impossible to give any informative and uncontroversial definition or characterisation of consciousness, and that all one can do by way of homing in on the subject is to point to examples of its presence. Thus, Block admits that he "cannot define P[henomenal] consciousness in any remotely noncircular way". And, while this predicament may not be unique for consciousness, the problems of identifying this subject matter is "in some respects worse than for many other concepts ... because really all one can do is *point* to the phenomenon" (1995, 230). In a similar vein, Chalmers complains that consciousness

can be frustratingly diaphanous ... it is notoriously difficult to pin down the subject matter. ... Trying to define conscious experience in terms of more primitive notions is fruitless ... The best one can do is give illustrations and characterizations that lie at the same level. ... I presume that every reader has conscious experiences of his or her own. If all goes well, these characterizations will help establish that it is just *those* that we are talking about (1996, 3-4).

However, I shall argue in the following that these observations should lead us to reconsider some *other* things that are often said about consciousness. For one thing, they should lead us to reconsider the plausibility of *infallibility* and *self-intimation* doctrines about consciousness.

---

respect to phenomenal consciousness. Of course, these experiences differ in what they represent, but according to the view, this representational difference does not determine or constitute a difference in "what it is like". Such a view is suggested by Moore (1903). Moore holds that, with respect to consciousness "all sensations are alike" (444). Sensations differ from one another with respect to their "objects", examples of which are blue and green, but these objects are *not* aspects of the sensations themselves: "'blue' is as much an object, and as little a mere content, of my experience, when I experience it, as the most exalted and independent real thing of which I am ever aware. There is, therefore, no question of how we are to 'get outside the circle of our own ideas and sensations'. Merely to have a sensation is already to *be* outside that circle. It is to know something which is as truly and really *not* a part of *my* experience, as anything which I can ever know" (451). Thus, according to Moore, what distinguishes a sensation of blue from a sensation of green is something that does not belong to the experience itself. Perhaps Byrne (2006) should be read as expressing sympathy with something like this view, but I'm not sure. I'm somewhat sympathetic with the

*Second difference: The status of infallibility and self-intimation doctrines.* As in the case of colours, there is a widely shared intuition that the nature of consciousness is reliably and in its entirety revealed to us through our first-person acquaintance with it. Since consciousness is just a matter of what things are like from a first-person point of view, it may seem that any carefully measured judgment about its nature, based on that first-person acquaintance, must be true,<sup>35</sup> and further, that anything that is essential to consciousness should be accessible to that point of view.<sup>36</sup>

However, given the observations just made, I don't see that it can be reasonable to have any faith in these infallibility and self-intimation doctrines. If practically nothing is uncontroversial when it comes to the nature of phenomenal consciousness, how could we be justified in thinking that the nature of consciousness is nevertheless reliably and in its entirety revealed to us? How could we be justified in thinking about *anything* that its nature is reliably and its entirety revealed to us if we are unable to agree on any reasonably informative claim about *what* that nature is?

Please note: I don't claim that the present lack of agreement about the nature of consciousness *refutes* any infallibility or self-intimation doctrine; it remains possible that we just haven't figured out how to "tap" our acquaintance with consciousness for reliable information even though the information is there, in some sense. I only claim that it cannot be reasonable to be *confident* in this, as long as practically everything about the nature of consciousness is controversial.

The case of colours is different. We have a rather rich and agreed upon set of ideas about what the natures of the colours are. So the infallibility and self-intimation doctrines about the colours are not subject to the kind of doubt I'm raising here.<sup>37</sup>

---

view myself. However, my aim in this paper is not to defend any *view* of consciousness, but rather to clarify what sort of metaphysical *problem* consciousness poses.

<sup>35</sup> This seems to be McLaughlin's view, for example: "we are acquainted with the phenomenal characters of our experiences in a way that leaves no room for illusion" (2003, 143).

<sup>36</sup> This may be part of what Nagel has in mind when he says: "It is difficult to understand what could be meant by the *objective* character of an experience, apart from the particular point of view from which its subject apprehends it. After all, what would be left of what it was like to be a bat if one removed the viewpoint of the bat?" (1974, 443).

<sup>37</sup> This is my objection, premised in section 3, to McLaughlin's claim that the apparent appeal of the Revelation doctrine about colour derives from the more genuine appeal of the Revelation doctrine about

*Third difference: The "potential reduction bases"*. Let us say that a property Q is in the "potential reduction base" (PRB) for a property P at a time t iff we have at t *some* reason to think either that P is Q or that P is some conjunction or disjunction of properties involving Q.

It then seems that the (PRB) for the colours is at present (a) relatively small, (b) relatively well understood and (c) plausibly complete. It is relatively *small* because it includes the following, limited number of properties: dispositions to cause experiences of certain types under certain conditions, dispositions to reflect light in certain ways, and the physical bases of these dispositions. This (PRB) is relatively *well understood* because its properties are. And it is *plausibly complete* in that we have reason to believe that colours, if they are identical with any set of physical or dispositional properties at all, are identical with some properties from either of these three sets. Frank Jackson sums up the situation:

There is an important sense in which we know the live possibilities as far as colour is concerned. We know that objects have dispositions to look one or another colour, that they have dispositions to modify incident and transmitted light in ways that underlie their dispositions to look one or another colour, that they have physical properties that are responsible for both these dispositions, and that subjects have experiences as of things looking one or another colour. We also know that this list includes all the possibly relevant properties (1998, 87).

In contrast, it cannot be said that we "know the live possibilities" as far as consciousness is concerned. The (PRB) for consciousness is (a) large, (b) in large parts little understood, and (c) not plausibly complete; if anything, it is plausibly incomplete.

To get some appreciation for the variety of actual and potential theoretical options here, consider the fact that phenomenal consciousness might reasonably be thought to (constitutively) have something to do with one or more than one *mental* phenomenon like attention, self-consciousness, representation, introspection, volition, memory or thought.

---

consciousness. That can hardly be right if, as I have argued, the Revelation doctrine about consciousness is less resilient to critical reflection than is the Revelation doctrine about colours.

Further, one may reasonably hold any number of views about what the functional essences are of these mental phenomena. These two choice points alone provide for numerous *functionalist* hypotheses about the nature of consciousness. And it's very likely that our current set of such hypotheses can be significantly improved upon. In addition, one might deny that consciousness is functional in nature, and hold that it is essentially a physical phenomenon. And then again, there is a variety of options about which phenomenon it might be identified with. Current proposals include ventral stream activation (see Block, 2001) and semisynchronised firings of relevant groups of neurons in the 40-70 Hz range (see Crick and Koch, 1990). But hardly anybody would rule out that radically new physicalist hypotheses may prove more promising than the ones we have right now.

Given these three differences -- concerning the pre-theoretical data, the status of infallibility and self-intimation doctrines, and the potential reduction bases -- it is clear that the situation facing a colour reductionist is very different from the situation of a consciousness reductionist.

The reductionist about *colour* faces a problem that is in several respects like the Bank Robbery Case. *First*, we have, on the basis of visual experience, a rather rich profile of the colours (just as the detective has a rich profile of the shooter). *Second*, the relative plausibility of infallibility and self-intimation doctrines about the colours limits the reductionist's liberty to "compromise" with this profile. Insofar as (C-Infallibility) is plausible, the colours should not be identified with any set of properties that *fail* to be the way we take the colours to be on the basis of experience.<sup>38</sup> And insofar as (C-Self-intimation) is plausible, the colours should not be identified with any set of properties that *are* in some way that we *do not* take the colours to be on the basis of experience. (Similarly, the detective has a well-confirmed and thus "non-negotiable" profile of the shooter.) *Third*, there is reason to think that the "live possibilities" for the colour reductionist are identified: causal considerations seem to show that that if colours are identical with any physical or dispositional properties at all, they are identical with either

---

<sup>38</sup> That is, so long as we restrict ourselves to carefully measured judgments that are about the *natures* of the colours. The same proviso goes for the next sentence.

dispositions to cause experiences, light-dispositions, or the physical bases of either of those dispositions. (Similarly, the detective has reason to believe that the shooter is among the arrested suspects). But *fourth*, there is at least some doubt that there is a good enough match between the colours and these live possibilities for us to identify the former with either of the latter (just as, apparently, none of the suspects matches the profile well enough for the detective to identify either of them as the shooter).

The reductionist about *consciousness*, in contrast, faces nothing like this problem. The problem *can't* be that we have a rich and well-established profile of consciousness that fails to find a match among a well-delimited and well-understood set of functional and physical properties, because we *don't* have a rich profile of consciousness, nor a well-delimited and well-understood potential reduction base for consciousness.

Rather, the problem for a reductionist about consciousness -- and, indeed, for *any* theorist of the subject -- is that there are so few uncontroversial data that constrain theorising. Hilary Putnam once said about the topic of *meaning* that it is "the one topic discussed in philosophy in which there is literally nothing but 'theory' -- literally nothing that can be labelled or even ridiculed as the 'common sense view'" (1975, 216). This is exactly what I claim to be true of the topic of *consciousness*: it is also a topic in which there is "literally nothing but theory".

I believe this diagnosis captures rather well the sense of *mystery* that consciousness tends to occasion. I find this sense of mystery slightly misarticulated by philosophers like Nagel and Levine. On their accounts, consciousness is mysterious when considered in the light of *reductionism* specifically. But we should recognise that consciousness is mysterious *independently* of the issue of reductionism. Consciousness is utterly familiar to all of us. In some sense, it seems to be more intimately known to us than anything else. At the same time, we don't have any informative and uncontroversial ways of characterising it. That is something of a mystery, independently of any commitment we may have to reductionism.

## 5. Some implications

I will conclude by spelling out two implications of what I have said about the problem of consciousness.

*Eliminativism and the undeniable existence of consciousness.* According to Georges Rey, "consciousness may be no more real than the simple soul exorcised by Hume" (1988, 461). Rey suggests that our belief in the existence of consciousness arises from a projection of a ghostly but morally significant property -- a projection grounded in our instinctive concern for the well-being of ourselves, our conspecifics and things that look like us (see also Rey 1995).

I find this suggestion intriguing, but here I shall make the more common assumption that consciousness just couldn't be the sort of illusion that Rey takes it to be. I wish to consider what follows and does not follow from this fact -- or alleged fact.

It is often suggested that the indubitable existence of consciousness rules out "eliminativism" about the phenomenon. Thus for example, Chalmers characterises eliminativism as the view that, "Nobody is conscious in the phenomenal sense" (1996, 161), and dismisses it as "manifestly false" (162).

This strikes me as misleading. The label 'eliminativism' is used for a rather heterogeneous variety of views, and not all of them are ruled out even if we assume -- contra Rey -- that there is some peculiar, indubitable existence datum in the philosophy of consciousness.

To illustrate this, consider the following three claims:

- (a) There is a difference with respect to phenomenal consciousness between dreaming and being in dreamless sleep.
- (b) There is a difference with respect to phenomenal consciousness between being awake and being in dreamless sleep.
- (c) There is a difference with respect to phenomenal consciousness between subliminal perception and normal perception.



If you are a "Rey-style eliminativist", you will regard these claims as expressing nothing but illusions. You will take them to be, in this respect, like the claim:

- (d) There is a difference with respect to witchhood between woman A and woman B.

Presumably, (d) describes a difference "no more real than the simple soul exorcised by Hume", and according to Rey-style eliminativism, so do (a)-(c).<sup>39</sup>

Rey-style eliminativism is not the only way of being sceptical of (a)-(c), however. One may also hold that, although each claim describes a real enough difference, we are deluded insofar as we think they describe differences in one and the same respect. If one has such doubts, one may regard the claims on the analogy of the following:<sup>40</sup>

- (e) If an object x warms up an object y when the two are put in contact with one another, then x is warmer than y.
- (f) If an object x warms up an object z when the two are put in contact with one another more than an object y warms up z when they are put in contact with one another, then x is warmer than y.
- (g) If an object x feels warmer than an object y to normal observers in normal circumstances, then x is warmer than y.

We may be inclined to think that these claims concern one and the same relation, being *warmer than*. But empirical investigation has revealed that they don't. (e) describes a difference in *degree* of heat, or temperature, whereas (f) describes a difference in *amount* of heat, or heat capacity, and (g) a difference in rate of heat *transfer*, or conductivity. And it is important to distinguish these, because even if an object x is "warmer than" an object y in one of these senses, x may not be warmer than y in either of the other two

---

<sup>39</sup> This is not to deny that there are differences between, say, dreaming and being in dreamless sleep, just as there are numerous differences between any two women. But according to the view in question, there is nothing remotely like the differences we imagine there to be when or if we embrace claims (a) or (d).

<sup>40</sup> The following example is borrowed from Paul Churchland (1979), 23-4.

senses. Similarly, it might turn out that (a)-(c) describe different distinctions, and that it is sometimes important to keep them separate.

We would hardly express such a discovery in words like: "nobody is conscious in the phenomenal sense", or "consciousness is no more real than the simple soul exorcised by Hume". (After all, we are not inclined to say such things about the relation of being *warmer than*.) But the idea that (a)-(c) describe different distinctions is still, it seems to me, *eliminativist* in one established sense of the term. It seems clear, for example, that philosophers like Paul and Patricia Churchland, who are characterised as 'eliminativists' by both themselves and others, would regard this kind of discovery as a confirmation of their views.<sup>41</sup>

Is such "Churchland-style eliminativism" ruled out by what is manifest to us concerning phenomenal consciousness? I don't see that it is. It may *seem* to be ruled out, given a certain idea of our epistemic predicament with respect to consciousness. I have in mind the idea that we have a rich and clear conception of what consciousness is and that, since consciousness just is a matter of what things are like for us, it cannot fail to be as it appears to us to be, nor be in any way that it doesn't appear to us to be. If our present grasp of consciousness were like that, we should be able to rule out that (a)-(c) describe different distinctions, because then there would be no room for any significant advances in our understanding of the phenomenon. But our present predicament with respect to consciousness is not like that. On the contrary, I have argued, we know close to nothing with certainty about consciousness, and, for that reason, we should not be confident that consciousness is reliably and in its entirety revealed to us. Given this predicament, I see no ground for ruling out significant, "Churchland-style" advances in our understanding of what we presently conceive of as "phenomenal consciousness". This scenario is a live possibility even if we are convinced that consciousness *is* "more real than the simple soul exorcised by Hume".

---

<sup>41</sup> See, e.g., Patricia Churchland (1986, especially chapters 5, 7 and 9), and Paul Churchland (1979, especially chapter 4).

*The explanatory gap.* The possibility of radical advances in our understanding of consciousness has implications for the issue of whether the "explanatory gap" between phenomenal consciousness and physical and functional properties can ever be closed.

McGinn (1989) argues that we will never be able to make consciousness fully intelligible to ourselves in physical or functional terms. This is a bold prediction. According to McGinn, we can *now* ascertain that we won't in the future find any physical or functional properties to be the way we *then* find consciousness to be. To be justified in predicting this one would have to know fairly well, now, how we will then understand consciousness. It may seem plausible that we have this knowledge, if one believes that we presently have a grasp of consciousness that leaves little room for revision or improvement. And this seems to be McGinn's view:

Our acquaintance with consciousness could hardly be more direct; phenomenological description thus comes (relatively) easily. 'Introspection' is the name of the faculty through which we catch consciousness in its naked vividness. By virtue of possessing this cognitive faculty we ascribe concepts of consciousness to ourselves; we thus have 'immediate access' to the properties of consciousness (1989, 354).

However, if what I have said above is correct, the assurance expressed in this passage is unfounded. Introspection does not provide us -- at the moment -- with a very substantial or clear view of the properties of consciousness. And, I have argued, since that is so, our understanding of consciousness may well develop in quite dramatic ways from now on. That is to say, we don't know very much now about how we will in the future understand consciousness. But then I don't see how we could rule out that what we will later on conceive of in terms of 'consciousness', or descendent notions, will *transparently* be a set of physical or functional properties.

## 6. Summary

I have argued that if we pay due respect to the distinction between colour properties and phenomenal consciousness properties, we will see that the two sets of properties pose very different sorts of metaphysical problem. The most important insight to be gained from this comparison is that the problem -- or mystery -- of consciousness is different from what it is often taken to be. The real problem of consciousness is that we have almost no useful data that constrain our theorising on the subject. This new perspective on the problem of consciousness enables us to reassess some views on, and diagnoses of, the problem. Thus, I have argued that we can't rule out all forms of *eliminativism* about consciousness. Nor can we rule out that future research will close the explanatory gap between consciousness and physical and functional properties.

## References

- Armstrong, David (1993) *A Materialist Theory of the Mind*, revised. ed., London: Routledge.
- Armstrong, David (1997) *A World of States of Affairs*, Cambridge: Cambridge UP.
- Armstrong, David (1999) *The Mind-Body Problem: An Opinionated Introduction*, Boulder, Colorado: Westview Pres.
- Armstrong, David, C.B. Martin and U.T. Place (1996) *Dispositions: A Debate*, ed. T. Crane, London: Routledge.
- Block, Ned (1995) "On a confusion about a function of consciousness", *Behavioral and Brain Sciences*, 18, 227-87.
- Block, Ned (2001) "Paradox and cross purposes in recent work on consciousness", *Cognition*, 79, 197-219.
- Block, Ned (2003) "Mental paint", *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, ed. M. Hahn and B. Ramberg, Cambridge, Mass.: MIT Press.
- Block, Ned, Owen Flanagan and Güven Güzeldere (eds.) (1997) *The Nature of Consciousness: Philosophical Debates*, Cambridge, Mass.: MIT Press.
- Bradley, Peter, and Michael Tye (2001) "Of colors, kestrels, caterpillars, and leaves", *The Journal of Philosophy*, 98, 469-87.
- Boghossian, Paul, and David Velleman (1989) "Colour as a secondary quality", *Mind*, 98, 81-103.
- Broackes, Justin (1992) "The autonomy of colour", *Reduction, Explanation, and Realism*, ed. D. Charles and K. Lennon, Oxford: Oxford UP.
- Byrne, Alex (2001a) "Do colours look like dispositions? Reply to Langsam and others". *Philosophical Quarterly*, 51, 238-45.
- Byrne, Alex (2001b) "Intentionalism defended". *Philosophical Review*, 110, 199-240.
- Byrne, Alex (2003) "Color and similarity", *Philosophy and Phenomenological Research*, 66: 641-65.
- Byrne, Alex (2006) "Color and the mind-body problem", *Dialectica*, 60, 223-44.
- Byrne, Alex, and David Hilbert (eds.) (1997a), *Readings on Color, I: The Philosophy of Color*, Cambridge, Mass: MIT Press.

- Byrne, Alex, and David Hilbert (1997b) "Colors and reflectances", in Byrne and Hilbert 1997a.
- Byrne, Alex, and David Hilbert (2003a) "Color realism and color science", *Behavioral and Brain Sciences*, 26, 3-21.
- Byrne, Alex, and David Hilbert (2003b) "Color realism redux", *Behavioral and Brain Sciences*, 26, 52-9.
- Byrne, Alex, and David Hilbert (forthcoming) "Color primitivism", online at <http://mit.edu/abyrne/www/colorprimitivism.pdf>.
- Campbell, John (1993) "A simple view of colour", *Reality, Representation, and Projection*, ed. J. Haldane and C. Wright, Oxford: Oxford UP. Page references to reprinting in Byrne and Hilbert 1997a.
- Chalmers, David (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford UP.
- Chalmers, David (2004) "The representational character of experience", *The Future for Philosophy*, ed. B. Leiter, Oxford: Oxford UP.
- Chalmers, David (2006) "Perception and the Fall from Eden", *Perceptual Experience*, ed. T.S. Gendler and J. Hawthorne, Oxford: Oxford UP.
- Churchland, Patricia (1986) *Neurophilosophy: Toward a Unified Science of the Mind-Brain*, Cambridge, Mass.: MIT Press.
- Churchland, Paul (1979) *Scientific Realism and the Plasticity of Mind*, Cambridge: Cambridge UP.
- Cohen, Jonathan (2003) "Color: A Functionalist Proposal". *Philosophical Studies*, 113, 1-42.
- Crick, Francis and Christof Koch (1990) "Towards a neurobiological theory of consciousness", *Seminars in the Neurosciences*, 2, 263-75.
- Dennett, Daniel (1991) *Consciousness Explained*, Boston: Little, Brown.
- Drake, Stillman (1957) *Discoveries and Opinions of Galileo*, New York: Anchor Books.
- Dretske, Fred (1995) *Naturalizing the Mind*, Cambridge, Mass.: MIT Press.
- Hardin, C.L. (1993) *Colors for Philosophers: Unweaving the Rainbow*, Indianapolis: Hackett.

- Harman, Gilbert (1990) "The intrinsic quality of experience", *Philosophical Perspectives* 4, ed. J. Tomberlin, Atascadero, California: Ridgeview.
- Harman, Gilbert (1996) "Explaining objective color in terms of subjective reactions", ed. E. Villanueva, *Philosophical Issues* 7, Atascadero, California: Ridgeview.
- Hilbert, David. (1987) *Color and Color Perception: A Study in Antropomorphic Realism*, Stanford: CSLI.
- Jackson, Frank (1977) *Perception: A Representative Theory*, Cambridge: Cambridge UP.
- Jackson, Frank (1982) "Epiphenomenal qualia", *Philosophical Quarterly*, 32, 127-36.
- Jackson, Frank (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford: Oxford UP.
- Jackson, Frank (2003) "Mind and illusion", *Minds and Persons*, ed. A. O'Hear, Cambridge, Cambridge UP.
- Johnston, Mark (1992) "How to speak of the colors", *Philosophical Studies*, 68, 221-63.  
Page references to reprint in Byrne and Hilbert 1997a.
- Johnston, Mark (unpublished) "The manifest".
- Kim, Jaegwon (1998) *Mind in a Physical World: An Essay on the Mind-Body Problem and Mental Causation*, Cambridge, Mass.: MIT Press.
- Kripke, Saul (1972) "Naming and necessity", *The Semantics of Natural Language*, ed. G. Harman and D. Davidson, Dordrecht: Reidel. Reprinted as *Naming and Necessity*, Cambridge, Mass.: Harvard UP, 1980.
- Levine, Joseph (1998) "Colour and qualia", *Routledge Encyclopedia of Philosophy*, ed. E. Craig, London: Routledge.
- Levine, Joseph (2001) *Purple Haze: The Puzzle of Consciousness*, Oxford: Oxford UP.
- Lewis, David (1994) "Reduction of Mind", *A Companion to Philosophy of Mind*, ed. S. Guttenplan, Oxford: Blackwell.
- Lewis, David (1997) "Naming the colors", *Australasian Journal of Philosophy*, 75, 325-42.
- Locke, John (1693/1995) *An Essay Concerning Human Understanding*, Amherst, NY: Prometheus Books.
- Lormand, Eric (1996) "Nonphenomenal consciousness", *Noûs*, 30, 242-61.
- Lycan, William (1996) *Consciousness and Experience*, Cambridge, Mass: MIT Press.

- Mackie, John. (1976) *Problems from Locke*, Oxford: Oxford UP.
- Maud, Barry (2002) "Color", *The Stanford Encyclopedia of Philosophy*, ed. E. Zalta, online at <<http://plato.stanford.edu/archives/fall2002/entries/color/>>.
- McDowell, John (1985) "Values and secondary qualities", *Morality and Objectivity*, ed. T. Honderich, London: Routledge.
- McGinn, Colin (1983) *The Subjective View: Secondary Qualities and Indexical Thoughts*, Oxford: Oxford UP.
- McGinn, Colin (1989) "Can we solve the mind-body problem", *Mind*, 98, 349-66.
- McGinn, Colin (1996) "Another look at color", *The Journal of Philosophy*, 93, 537-53.
- McLaughlin, Brian (2003), "Colour, consciousness and colour consciousness", *Consciousness: New Philosophical Perspectives*, ed. Q. Smith and A. Jokic, Oxford: Oxford UP.
- Moore, G.E. (1903) "The refutation of idealism", *Mind*, 12, 433-53.
- Nagel, Thomas (1974) "What is it like to be a bat?", *Philosophical Review*, 83, 435-50.
- O'Regan, Kevin, and Alva Noë (2001) "A sensorimotor account of vision and visual consciousness", *Behavioral and Brain Sciences*, 24, 939-973.
- Palmer, Stephen (1999) *Vision Science: Photons to Phenomenology*, Cambridge, Mass.: MIT Press.
- Papineau, David (2002) *Thinking about Consciousness*, Oxford: Oxford UP.
- Peacocke, Christopher (1983) *Sense and Content: Experience, Thought and their Relations*, Oxford: Oxford UP.
- Perry, John (2001) *Knowledge, Possibility and Consciousness*, Cambridge, Mass.: MIT Press.
- Putnam, Hilary (1975) "The meaning of 'meaning'", *Language, Mind and Knowledge*, ed. K. Gunderson, Minneapolis: University of Minnesota Press. Page references to reprinting in Putnam, *Philosophical Papers, vol. 2: Mind, Language and Reality*, Cambridge: Cambridge UP, 1975.
- Rey, Georges (1988) "A question about consciousness", *Perspectives on Mind*, ed. H. Otto and J. Tueidio, Norwell: Kluwer Academic Publishers. Page references to reprinting in Block et al 1997.



- Rey, Georges (1995) "Towards a projectivist account of conscious experience", *Conscious Experience*, ed. T. Metzinger, Paderborn: Schöningh-Verlag.
- Rosenthal, David (2002) "Explaining Consciousness", *Philosophy of Mind: Classical and Contemporary Readings*, ed. D. Chalmers, Oxford: Oxford UP.
- Russell, Bertrand (1912) *The Problems of Philosophy*, Oxford: Oxford UP.
- Searle, John (1992) *The Rediscovery of the Mind*, Cambridge, Mass.: MIT Press.
- Shoemaker, Sydney (1982) "The inverted spectrum", *The Journal of Philosophy*, 79, 357-81.
- Shoemaker, Sydney (1990) "Qualities and qualia: What's in the mind", *Philosophy and Phenomenological Research*, 50, Supplement, 109-31. Page references to reprinting in Shoemaker 1996.
- Shoemaker, Sydney (1994) "Phenomenal character", *Noûs*, 28, 21-38.
- Shoemaker, Sydney (1996) *The First-Person Perspective and Other Essays*, Cambridge: Cambridge UP.
- Siewert, Charles (1998) *The Significance of Consciousness*, Princeton: Princeton UP.
- Smart, J.J.C. (1975) "On some criticisms of a physicalist theory of colors", *Philosophical Aspects of the Mind-Body Problem*, ed. C. Cheng, Honolulu: University Press of Hawaii.
- Steward, Helen (1997) *The Ontology of Mind: Events, Processes and States*, Oxford: Oxford UP.
- Strawson, Galen (1992) "'Red' and red", *Synthese*, 78, 193-232.
- Strawson, Galen (1994) *Mental Reality*, Cambridge, Mass: MIT Press.
- Stroud, Barry (2000) *The Quest for Reality: Subjectivism and the Metaphysics of Colour*, Oxford: Oxford UP.
- Tye, Michael (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, Cambridge, Mass.: MIT Press.
- Tye, Michael (2000) *Consciousness, Color and Content*, Cambridge, Mass.: MIT Press.
- Tye, Michael (2002/3) "To panic or not to panic? -- Reply to Byrne", *A Field Guide to the Philosophy of Mind*, ed. M. Nani and M. Marraffa, online at <http://host.uniroma3.it/progetti/kant/field/tyesymp.htm>.

Tye, Michael (2003) "Qualia", *Stanford Encyclopedia of Philosophy*, ed. E. Zalta, online at <<http://plato.stanford.edu/archives/sum2003/entries/qualia/>>.

Yablo, Stephen (1995) "Singling out properties", *Philosophical Perspectives* 9, ed. J. Tomberlin, Atascadero, California: Ridgeview.

*Department of Philosophy and Linguistics*

*Umeå University*

*SE-901 87 Umeå*

*Sweden*

*Email: [par.sundstrom@philos.umu.se](mailto:par.sundstrom@philos.umu.se)*