Final published version in R. Johnson and M. Smith (eds.), *Passions & Projections: Themes from the Philosophy of Simon Blackburn*. Oxford: Oxford University Press, 2015, pp. 116-133. Citations should be to the published version: http://ukcatalogue.oup.com/product/9780198723172.do

What is Colour? A Defence of Colour Primitivism

Cynthia Macdonald

> The particular bulk, number, and motion of the parts of fire, or snow, are really in them, whether anyone's senses perceive them or no: and therefore they may be called real qualities, because they really exist in those bodies. But light, heat, whiteness, or coldness, are no more really in them, than sickness or pain is in manna. Take away the sensation of them; let not the eyes see light, or colors, nor the ears hear sounds; let the palate not taste, nor the nose smell, and all colors, tastes, odors, and sounds… vanish and cease… (Locke 1690/1975: Book 2, chapter 8, Section 17: 137-8)

In this familiar passage, Locke contrasts primary with secondary qualities, claiming that whereas primary qualities really are in the objects of perception, colours, sounds, and the like are no more really in perceptible objects than sensations such as pain are in them. Locke famously vacillates between the view that colours are 'nothing but' powers in the objects due to their primary qualities to produce ideas of colours in the mind and the view that colours are ideas in the minds of perceivers, and he does so for reasons that still preoccupy philosophical thought about the nature of colours. On the one hand, we are disinclined to say that colours are only in the mind, as sensations such as pain are thought to be. On the other hand, we are reluctant to say that colours are 'in' objects in the same way that primary qualities are – they seem to be relational, or extrinsic, rather than intrinsic properties of the surfaces of objects (hence Locke's reference to them as 'powers').[1] Yet the phenomenology of colour

[1] The view that dispositions are relational, or extrinsic, rather than intrinsic properties of the objects that have them is notoriously controversial (for some who think that at least some are, see McKitrick 2003, Yablo 1999, Nolan 2005; for some who think not

experience, which presents colours as features the surfaces of perceptible objects that are seen in colour experience, does not seem to be adequately accounted for by a dispositional view.

Simon Blackburn approaches this problem in his usual subtle and insightful way. While he has some sympathy for those who think that colour properties are analogous to moral ones in that neither sorts of properties are really in the objects that apparently have them, he thinks that there is a crucial disanalogy between these two types of property. Whereas judgements about the colours of objects are descriptive or fact-stating, moral judgements are not. Such judgements are, rather, expressions of moral sentiments. Blackburn is a cognitivist projectivist about colour properties and a non-cognitivist expressivist projectivist about moral ones.

Blackburn's reasons for siding with projectivism about colours can be garnered from his discussion of response-dependence accounts of colours and moral properties and his skeptical conclusions about what they can hope to achieve in the way of explanatory value (Blackburn 1993/2010). Since projectivism about colours is a view that many (including myself) find deeply unattractive, I want to focus on Blackburn's discussion and on the intuitions that form the basis for his dissatisfaction with response-dependence accounts of colours. I share these intuitions, but believe

---

see Armstrong 1973, Bird 1998, Harré 1970, Mackie 1973 and Mellor 1974). My aim here is not to adjudicate between opposing sides of this view (nor to adjudicate on the matter of relationality vs extrinsicness, if they are dispositions) but to articulate the idea, which fuels response-dependence accounts of colours, that colours and other secondary qualities of objects of perception are in some sense less fully objective features of them than their primary qualities such as mass and shape. This distinction might be coherently made out even if colours are intrinsic properties of objects, as some dispositionalists maintain - if, for example, colours supervene on the microphysical properties of objects and their causal powers are dependent on or derived from those microphysical properties.

that they can be respected by a realist primitivism about colour properties.[2] My aim is to argue that this type of view is the preferable one by defending it against two prominent objections.

Blackburn's discussion of response-dependence accounts of colours focuses on accounts of colour concepts, and Blackburn himself is inclined to favour this approach over 'substantial' accounts of the nature of colours, which are concerned with the metaphysics of colour properties. Since my interest is in the latter, I begin in Section 1 by providing some motivation for the more substantial approach. Section 2 sets out some plausible constraints on a theory of colours which appeal to some of the intuitions that Blackburn and I share concerning what a theory of colours should set out to explain and then briefly rehearses the difficulties Blackburn and others think response-dependence accounts face attempting to meet those constraints. Although Blackburn's objection is to concept versions of the account (specifically, to 'weak analyses'), my discussion focuses on the whole on property versions of it. Section 3 explains how the primitivist view meets the relevant constraints, defending it against two of the most prominent objections voiced in the literature.

1. Motivating the Substantial Account

An interesting fact about colours is that most normal people can acquire an understanding of them, and acquire colour concepts, without knowing anything about colour science or about what the physical or dispositional properties are of the

[2] Primitivism is neutral as between two views, realist primitivism and eliminativist primitivism (cf. Chalmers 2006). My aim here is to defend the realist view (proponents of which include Gert 2008, Watkins 2010, Campbell 1997, 2005, and Yablo 1995) rather than the eliminativist one (Pautz 2006).

surfaces of objects or entities that have colours. One can typically master colour concepts long before one knows (a) anything about the propagation of light or about surface reflectance frequencies and (b) anything about the physiology of colour vision in the normal human retina.

Characteristically, colour concepts are acquired on the basis of normal human visual experience and colour terms are taught ostensively on the basis of the structure of largely shared phenomenal experience of how colours appear or seem. This structure forms the basis for two sorts of colour structure claims about determinate colours:

1. Resemblance claims. E.g., 'Blue resembles purple more than green'.
2. Binary/unitary claims. E.g., 'Purple is a binary colur, whereas green is a unitary colour'.

Ostensive teaching of colour terms enables one to apply such terms to objects just on the basis of normal perceptual experience of them, and this connects the use of such terms with phenomenal experience of colours. But this in itself falls short of a response-dependence account of colour concepts. Response-dependence views are motivated by the thought, articulated in Locke's quotation cited above, that such properties are less objective in some intuitive sense than primary qualities like shape or mass. One way in which they are thought to be so is that, whereas truths about the shape or mass of an object of perception seem to hold in virtue of how that object is in itself, independently of its effects on normal perceivers in normal circumstances, truths about its colour do not. Response-dependence accounts exploit this intuitive sense in which truths about an object's colour are less objective than truths about its shape. Such truths, they maintain, are so in virtue of a coloured object's relation to

normal perceivers in normal circumstances, specifically, in virtue of certain subjective responses it characteristically induces in them.

As Blackburn notes, response-dependence accounts are committed to giving necessary and sufficient conditions for an object's having a certain colour in terms of the subjective responses that it is apt to cause in normal subjects in normal circumstances, but they differ from reductive accounts in that they are committed to bi-conditionals of the form:

> X is φ ≡ X is such as to induce the subjective response R in persons P in circumstances C,

Where X ranges over objects of perception, φ over colours, R over subjective responses on the part of thinking subjects that essentially involve a mental phenomenon of some type, P over persons, and C over relevant circumstances.[3] Such conditionals are held to be both necessary and *a priori*.

These bi-conditionals are not reductive because they are to be read both from 'left to right' and from 'right to left'; but it is the reading from right to left that lends support to the view that it is *in virtue of* an object's being such as to elicit the subjective response R, say, the experience as of red, in normal persons in normal circumstances that it is red. And it is this reading that lends support to the view that colour properties are less objective than other properties of perceptible objects. It is the subjective response that all red things have in common, in virtue of which they are all red, rather than some natural intrinsic feature that objects of perception have independently of human experience, which explains why they are red.

---

[3] Blackburn's (1993/2010) formulation takes the form:

> X is φ ≡ X is such as to elicit the judgement that it is φ from [P,C].

I am assuming here that the judgement is elicited in virtue of the evoking of a subjective response of some type on the part of a thinking subject, where this essentially involves some sort of perceptual experience as of X's being φ.

On the face of it, bi-conditionals of the form given above seem to offer, not an analysis of concepts or of the meaning of terms, but rather, a metaphysical account of what it is for things to *be* φ. That is, they seem to specify what it is for a property to be the property φ, not of what it is for a concept to be the concept *φ*, by giving necessary and sufficient conditions for something to be that property. A response-dependent property is one whose nature it is to stand in a relation to the responses of something else. Accordingly, a more perspicuous way of stating the response-dependence account of colour properties would quantify over properties in something like the following way:

> A colour property φ is a response-dependent property if and only if part of what it is to be φ is to stand in relation R to a certain type of subjective mental response in persons P in circumstances C.[4]

The intuition that colours are less fully objective, less fully intrinsic features of the objects that have them, than are other properties of those objects, is not made explicit by concept formulations. If colour properties are response-dependent, then of course they are correctly thought of as response-dependent. But demonstrating that colour concepts are response-dependent is not sufficient to demonstrate that the properties that these concepts stand for are response-dependent. For there may be more than one concept of a given property, one of which is response-dependent and the other of which is not, and this may be true of colour properties.

Consider, for example, Pettit's discussion of response-dependent (1991)/response-privileging (1993) concepts. Pettit follows Johnston's characterization of response-dependent concepts as ones that stand for dispositions

[4] See Wedgewood 1997, from whom this characterization is adapted. The relevant type of response involves colour experiences.

that are manifested in certain responses in normal subjects in normal circumstances. Although he sometimes speaks of response-dependence as being part of the nature of colours,[5] he insists that his account is "not an assertion about the property or object or operation in question" (1993: 202-3), and he explicitly rejects the suggestion that the properties that response-privileging concepts stand for are, in virtue of this, themselves dependent in some way on our subjective responses (see also Holton 1991). As he sees it, concepts of natural kind properties such as the concept, *water*, are plausibly thought of as response-dependent but are not plausibly thought of as concepts of response-dependent properties. This being so, although response-dependence accounts of concepts define them in terms of the subjective responses that they are apt to induce in us, this is compatible with taking "the predicate 'red' to direct us to the realizer-property that makes things look red, not to the role-property" (1998: 122), and the realizer property may be a perfectly objective physical property (or a disjunction of properties) that is not in any way dependent on us.[6]

Like many others, Pettit regards response-dependence accounts of secondary quality concepts as being tied to the epistemology and semantics of secondary quality terms but not in and of themselves to the metaphysics of secondary qualities.[7] But this is

---

[5] So, for example, he says "something is red because it looks red to normal observers: the capacity to look red to such observers is what marks off red things" (1991: 615)

[6] Thus, he says,

> Take any basic term or concept, ' T ', that is used in common amongst a community of speakers to refer to something, T, where T may be a perfectly objective entity: like a spectral reflectance, it may be the sort of thing that can exist in the absence of the community and in the absence of any thinking creatures. ' T ' will be response-dependent just in case an ordinary speaker 's competence in the use of the term goes hand in hand with their believing of anything they encounter that it is T if it seems T and there is no evidence of unfavourable influences; and with their believing that it is not T if it seems non-T and there is no evidence of unfavourable influences. (Jackson and Pettit 2002: 99)

[7] Cf. for example, Yablo:

precisely why such accounts are ill suited to satisfy the intuitions that motivate Locke and those who subscribe to response-dependence accounts of secondary qualities. Their concerns are with the nature of secondary qualities themselves. And this is what seems also to motivate Johnston's discussion of response-dependence of colours and the bi-conditionals associated with them (cf. Johnston 1998). So even if we agree that the bi-conditionals are necessary and *a priori* and provide a satisfactory account of the (folk) colour concepts, there would be a need for the more substantial account of the nature of colour properties.

2. Does the Response-Dependence View Satisfy Plausible Constraints on a Theory of Colours?

Blackburn's principal objection to response-dependence accounts of colour concepts is that they are no improvement on expressivist ones. Even taken as theories of colour properties, they do not explain what they should set out to explain, namely, the reactions of those who are judging the colours of perceptible objects.

> The crucial problem for secondary quality perception is now apparent. Colours are seen, sounds heard, smells smelled. …Our reactions *do* present themselves as perceptual awareness, yet we have no stable conception of their *right* to so present themselves. The true situation is probably much more easily understood with smell and taste than with colour: we are more easily led by Berkeley to think that the nose or palate tells us nothing about the world, than that chromatic vision tells us nothing either. … Dispositional accounts solve the problem of right, but make colours and the rest essentially imperceptible. Rebounding from that, we confront the problem of how the bare subjectivity of response transforms itself into a genuine awareness of a property. (Blackburn 1993/2010: 243)

> To call a property "subjective" is to comment in an ontological vein about what it is. But to say that it is not adequately conceived except (e.g.) in terms of how it makes things look is to applaud certain ways of thinking of the property. Unless standards of adequate conception are dictated by the property and it alone, no ontological conclusions follow. (Yablo 1995: 491)

Blackburn's claim is that the main point of a theory of colours is to explain our perceptual experiences of them, and because response-dependence accounts (as a species of dispositional ones) make colours imperceptible, it cannot explain our perceptual experiences of them.[8] In effect, he is gesturing at a familiar and plausible phenomenological constraint on a theory of colours that many appeal to as a desideratum on an adequate theory of colours (and more generally, of secondary qualities).[9] According to this, colour experience presents colours as genuine features of the surfaces of perceptible objects that are seen in colour perception. Because it does, an important constraint on a theory of colours is a principle which we might call *Transparency*:

(1) *Transparency*: Colours really are as they are presented as being when presented to normal subjects in normal circumstances (i.e., they are they way they look to be).[10]

---

[8] Cf. also Johnston (1998: 17-18), who suggests that it is a consequence of the idea of sensing a family of qualities being 'form of receptivity' to how things are with respect to colours that an object's being red can explain why standard subjects under normal conditions are disposed to see it as red, and argues compellingly that because response-dependence theories cannot meet this explanatory constraint they make response-dependence features essentially imperceptible.

[9] For appeal to phenomenological constraints on a theory of sounds, for example, see O'Callaghan:

> From the outset, two initial kinds of constraint bear on the theory of sounds. The first is phenomenological. Given that sounds are among the things we hear, how we hear them to be is relevant, prima facie, to theorizing about what sounds are. All else equal, an account that captures the phenomenology of auditory perception is preferable to one that does not. (O'Callaghan 2007: 14)

And for an appeal in the area of meta-ethics, see McNaughton 1998: 40.

[10] This statement of the thesis of Transparency is similar to that stated by Campbell, which reads:

> TRANSPARENCY: "The real nature [of, e.g., redness] is transparent to us" (Campbell 1993: 178); "colours are...properties with which ordinary observation directly acquaints us...ordinary colour vision is enough for us to know *which* property blueness is, for example" (Campbell 1993:186).

More recently, he has given a slightly different formulation:

This constraint is said to be difficult for response-dependence theories to satisfy alongside a second important constraint on a theory of the nature of colours:

(2) *Explanation*: Sometimes something's being a given colour (e.g., red) explains why it looks to be that colour (e.g., looks red). (Johnston 1992)

The reason is essentially that stated by Blackburn: colours are not presented as dispositions, or as response-dependent, relational properties, to normal subjects in normal circumstances. Rather, they are presented as intrinsic properties of the surfaces of objects, possessed by them independently of the effects they might have on normal perceivers. One might of course object that if colours *are* in fact dispositions, then some dispositions *do* present themselves to perceivers as intrinsic properties of the surfaces of objects. But this is a difficult claim to sustain, since in visual perception they are not seen in the surfaces of objects that possess them. Even if they are intrinsic properties, they present differently than primary qualities such as shape and mass.

> TRANSPARENCY: Experience of color provides knowledge of the categorical color property intervention on which changes the experiences of observers. (Campbell 2005: 111).

Both versions of the thesis are weaker than Johnston's REVELATION, which states that experience of colour provides knowledge of truths about the nature of colour. According to Campbell, REVELATION claims that colour experience provides propositional knowledge about the essence of colours rather than knowledge of colours themselves and for this reason does not capture the intuition that colour experience directly acquaints us with colours. Versions of *Transparency* are intended to capture the idea that visual experience is all that is needed to know which properties the colours are, but without the commitment that the complete essential nature of colours is 'laid bare' by visual experience of them. And this is compatible with the view that, for example, it is part of the nature of the colours that they supervene on and are realized by microphysical properties of the objects that have them.

Response-dependence theories are also thought by many to have difficulties accommodating the first of the following two further plausible constraints on a theory of the nature of colours:

(3) *Causality*: Colours are properties of things that bear causal relations to experiences as of them (e.g., red is a property that typically causes experiences as of red).

(4) *Commonality*: For any given colour, F, F is a property of objects that paradigmatic instances of F share. (cf. Yablo 1995)

Causality is thought by some to be problematic for such theories because colours constitute a species of disposition, and dispositions are only causally relevant because they are relevant to causal explanations of their manifestations which cite the categorical microphysical properties that realize those dispositions and are causally efficacious, properties that are not identical with the dispositions they realize (Jackson 1998). Others maintain that dispositions are causally irrelevant altogether either on the grounds that the meanings of dispositional terms connect them necessarily and *a priori* to their manifestations whereas causes are contingently connected to their effects (Block 1990; Dardis 1993; Jackson 1996) or on the grounds that the microphysical realizers of dispositions are complete and independent causes of their manifestations, and so there is no causal/explanatory work for dispositions to do (Kim 1990; Prior, Pargetter, and Jackson 1982).

Since our focus in on response-dependence accounts of colour properties rather than on analyses of colour concepts, objections based on the meanings of dispositional terms is less relevant to the discussion than ones concerning dispositions themselves. However, there are many who argue that dispositions *are* causally efficacious and causally relevant to their manifestations. So, although *Causality* is an important

constraint on a theory of colours, the case against response-dependence accounts based on failure to meet it is not decisive. It is true, too, that primitivism is vulnerable to the charge that it makes colours causally inefficacious and causally irrelevant, and we have yet to see whether it can effectively meet that charge.

The main objection to response-dependence accounts, then, is the one gestured at by Blackburn - that it does not meet the phenomenological constraint on a theory of colours, namely, that such a theory should explain the fact that colour experience presents colours as genuine features of the surfaces of perceptible objects that are seen in colour perception, and the Transparency thesis associated with that constraint. If primitivism can account for this and give an effective response to the causal inefficacy/causal relevance charge, then it will emerge as the preferable account. At the same time it will offer a suitable alternative to the expressivist position favoured by Blackburn.

3. A Defence of Primitivism

Primitivist theories of the nature of colour properties occupy a space somewhere in between response-dependence theories and reductive physicalist ones. Reductive physicalist theories take colours to be physical properties (either dispositional ones, such as spectral reflectance frequencies, or non-dispositional ones, such as microphysical categorical properties – which may be highly disjunctive - of the objects that realize them), such physical properties being the *grounds* of dispositions that objects have to cause certain types of responses in normal perceivers in normal circumstances. Response-dependence theories, as we have seen, take colours themselves to be a species of dispositions, whose grounds are either the dispositional or the non-dispositional

microphysical properties that realize those dispositions. Primitivism, in constrast with both of these theories, construes colours as simple irreducible or *sui generis*, perceiver independent, properties that dispose the objects that have them to produce certain types of responses in normal subjects in normal circumstances, which satisfy (1) *Transparency*, (2) *Explanation*, (3) *Causality*, (4) *Commonality*, and

(5) *Nonrelationality*: Colours are intrinsic, nonrelational grounds of dispositions of objects to cause experiences of them.

That is, according to primitivism colours *have*, rather than *are*, dispositions. Thus, colours are the *sui generis* intrinsic, nonrelational grounds of dispositions of objects to cause experiences of them, such grounds themselves being grounded in, by being realized by, microphysical properties of the objects that are coloured.[11] This being so, they really are 'in' the objects of perception, independently of perceivers; they are features of the surfaces of objects that are seen in them and typically are as they look to be; they explain why coloured objects look to be coloured; they are properties that typically cause experiences as of them; and they are the what all things that share particular determinate colour properties have in common – e.g., red is the property that all red things have in common.

Despite its intuitive appeal, primitivism is a minority view amongst theorists of the nature of colours. Critics maintain that there is no simple, *sui generis* property of the

[11] The view that colours have rather than are dispositions is also held by functionalist theories (see, for example, McLaughlin 2003). However, on the functionalist view, it is not the higher-level or role properties of perceptible objects that are realized by microphysical properties that are the colours, but rather, the realizer properties themselves - the ones that play the role of disposing objects that have them to look coloured. Further, it is the realizer property or properties that all red things have in common, in virtue of which red things are red.

kind posited by primitivism that can answer to all of these constraints. Two specific claims are that:

(1) Primivism cannot capture truths about the structure of colours (their resemblances to one another and their unitary/binary natures in particular) (Maund 1995, 2008). So, claims like 'blue resembles purple more than green', which phenomenal experience represents as true about the structure of colours, is not true of the colours and *Transparency* is violated.

and

(2) If primitivism is true, colours are causally inefficacious and/or causally irrelevant properties of the objects that have them. So *Causality* is violated.

Let's consider these objections in turn.

Barry Maund (1995, 2008) and others have argued that the main most serious objection to primitivism is that no objective properties of the kind it supposes there to be can capture truths about the structure of colours. Maund maintains that this objection applies to reductive physicalist accounts as well as to primitivist ones. His objection to the reductive physicalist ones is that, given the radically disjunctive nature of microphysical properties that they identify the colours with, and given the problem of metamers (i.e., that many different microphysical properties can produce experiences as of the same colour), physical science attributes vastly more resemblance relations to the realizers of colours than visual experience can detect on the basis of hue, saturation, and brightness. So it could turn out that one shade of blue is more similar to yellow than it is to another shade of blue on the reductive physicalist account, thereby falsifying truths about colour structure resemblances.

McLaughlin (2003) and others respond to this problem by advocating the experiential account of colour structure, according to which colour structure is not a feature of the microphysical properties that are the colours. Rather, it is a feature of colour experiences themselves. However, this response is considered by some to be problematic because it requires rejecting the view that the phenomenal characters of our colour experiences are fixed by what colours the experiences are experiences of, which violates *Transparency* (Pautz 2006).

Primitivism does not suffer from this problem since it does not identify colours with the highly disjunctive microphysical properties of objects that realize colours. So why is it not capable of capturing truths about the structure of colours? One problem that the view faces is that animals other than humans seem to respond by means of similar neural machinery to microphysical properties of objects in a way that indicates that they are responding to colours. This suggests that facts about unitary and binary colours, and about colour resemblances, are fixed by the neurobiology of *perceivers* rather than by the objective colours themselves, which threatens *Nonrelationality* as well as *Explanation*. For example, Byrne and Hilbert (2007) claim that primitivism cannot give a satisfactory account of the discriminatory capacities of animals such as goldfish, who share many of the mechanisms of colour vision of humans, but can see into the near ultraviolet range, which humans cannot. If, as *Transparency* says, colours are as they look to be, primitivists are at a loss to explain what goldfish are reacting to, since if they are reacting to colours, they are not doing so by reacting to the colours we humans see. And it is exceedingly plausible that they are reacting to colours.

But primitivists are not without resources for responding to this. They could, for instance, say that whatever goldfish are responding to, they are not colours, or at least not the colours that are the objects of normal visual experience – the colours that are

taught ostensively, concepts of which are acquired by normal visual experience, namely the chromatic colours with the features of hue, saturation, and brightness. Or, there might be two 'types' of colours; those that vision detects just by being presented in normal everyday experience, and those that are detected by analogy with the neural mechanisms that are similar to those that ground or by virtue of whose functioning humans detect colours.

One problem with this response is that it is in tension with the spirit of primitivism, since it makes what colours there are dependent on evolution, which presumably could categorise electromagnetic radiation in a potential infinity of ways. In making primitivism a hostage to evolution, the response undermines the robust realism to which primitivism is committed. For this reason primitivists may well be reluctant to appeal to it.

There is, however, an alternative. Primitivists could maintain that goldfish are indeed responding to colours, but that there are more colours than what human beings can detect (Gert 2008). Perhaps colour science can tell us more about the colours, for example, that there are more than those identified in normal visual experience of humans. It is plausible that if goldfish are distinguishing between and responding to colours, they are not doing so by identifying the colours that we humans see. This raises the question of what colours they *are* responding to. Byrne and Hilbert argue that because primitivists cannot plausibly claim that goldfish are responding to the colours that we humans see, they are committed to denying that goldfish are responding to colours at all. But *Transparency* does not commit primitivists to the view that there are no colours outside human ones - only that if there are, we humans cannot represent them. So primitivists are not committed to the implausible claim that reflection on

human colour experience reveals that there are no colours outside the colour space that humans perceive.

It might be argued that this response is anti-occamist in allowing for a plethora of primitive properties that happen to correspond to possible ways of perceiving the world. But it is unclear how forceful this argument is. We know, for example, that there are many sounds that we humans cannot hear but non-human animals such as dogs can. Why should not the same be true of colours? Further, it is an empirical question how many sounds there are; equally it is an empirical question how many colours there are. Whether there are creatures capable of perceiving them is another matter.

It seems, then, that the first objection to primitivism, based on its apparent failure to handle truths about colour structure claims, can be countered in plausible ways. The heart of much of the resistance to colour primitivism seems rather to be based on the view that these supposed higher-level, *sui generis* properties of perceptible objects are, if real, causally inefficacious and/or causally irrelevant to the effects colour properties have in normal subjects in normal circumstances. If this is right, such properties could not explain what a theory of colour should set out to explain, namely our perceptual experiences of colours. What explains our experiences are not these higher-level properties of perceptible objects, but rather, their microphysical realizers.

One principal source of objections of this kind stems from an argument advanced by Jaegwon Kim, originally against the causal efficacy and relevance of mental properties on the nonreductive physicalist/monist view, known as the Exclusion argument (Kim 1990). Briefly, assuming that mental properties are irreducible to but supervene on and are realized by microphysical properties of the events that have them, the question arises whether those events bring about the effects they do in the behaviour of their subjects in virtue of their physical properties or in virtue of their mental ones

(the so-called *qua* problem). These effects are physical ones; mental events causally interact with physical ones and so are causes (as well as effects) of physical events, including the body movements that constitute behaviour. Kim appeals to two principles, one known as *Closure* (according to which every physical event, if it has a cause, has a complete sufficient physical cause), and the other known as *Exclusion* (according to which there cannot be two complete and independent sufficient causes of a given effect), to argue that the only way that the mental property of a mental/physical cause could be causally efficacious (and so causally relevant) to its physical effect in behaviour (overdetermination excluded), would be by breaching *Exclusion*. Since the physical property of the mental/physical event is already guaranteed by *Closure* to be a complete sufficient cause, *Exclusion* rules out the mental one. The argument generalizes to any higher-level property that supervenes on and is realized by lower-level physical ones. So if it works, it also applies to colour properties construed along the lines of the primitivist account.

Elsewhere I've argued that the Exclusion argument equivocates between the causal efficacy of mental events (/instances of mental properties) and causal relevance of mental properties (properties whose instancings are those events) (Macdonald and Macdonald 1986, 2008). On the reading of the argument that concerns events/instances of properties, the conclusion does not follow, since, on a version of the property-exemplification account of events that construes events as instancings of properties at or during intervals of times in objects, the claim that mental events are identical with physical events can be understood as the claim that instancings of mental properties are identical with instancings of physical ones. On this 'co-instantiation' thesis, given the extensionality of the causal relation, mental events are not ruled out by *Exclusion* from being causally efficacious to their physical effects. A similar response can be given for

other higher-level properties that supervene on and are realized by microphysical ones, specifically, for colour properties on the primitivist view.

But the deep worry expressed by the expiphenomenalism charge isn't solved just by addressing the causal efficacy issue. The worry is that even if higher-level properties can be shown to be causally efficacious in virtue of their instances being identical with instances of lower-level ones, this does not by itself show that it is the higher-level property whose causal power is being exercised in bringing about the effects its instances cause. So, what about the reading of the argument that concerns properties? According to this, *Exclusion* says that if a property, *P*, of a cause, *c*, is causally sufficient for an effect, *e*, then no other property, *Q*, distinct from and independent of *P*, is causally relevant for *e* (where *P*'s being causally sufficient for *e* means that an instance of *P* in an event *c* is causally sufficient for *e*, and where, by 'causally relevant' is meant that properties of events are such that their instances are causally effective in bringing about effects of those events) (cf. Macdonald and Macdonald 2008). Accordingly, the property version of *Closure* now reads: if a physical event has any cause, it has a sufficient physical cause, whose physical properties are causally sufficient for its effect. It follows from this reading of the argument that higher-level properties and microphysical properties of physical events cannot both be completely and independently causally relevant to a single effect because they cannot both be completely and independently causally sufficient for that effect. Given the property-version of *Closure* it might look like the higher-level ones are ruled out as causally relevant.

Turn now to colour properties on the primitivist view. Effectively, colour primitivism is a version of non-reductive physicalism. This being so, the property versions of *Closure* and *Exclusion* seem equally to lead to the conclusion that colour

properties are causally irrelevant, given that they are not reducible to lower-level microphysical ones and given the plausible view that it is the lower-level ones that ground the causal efficacy and sufficiency of the higher-level ones. In the face of this, others working on the causal relevance of higher-level properties have attempted to salvage the causal relevance of the higher-level ones in one or the other of two ways.

One is to deny physicalism, claiming that instances of higher-level properties are wholly distinct from instances of lower-level microphysical properties (i.e., properties of the sort quantified over by empirical physical theories) (Yablo 1992). According to this approach, higher-level properties can be causally effective and relevant because their instances are. The strategy assumes that a necessary condition on the causal relevance of a property is causal efficacy of its instance. But in order to ensure the causal relevance of higher-level properties, it requires denying both the property-instance and the property-type version of *Closure* and leaves unresolved exactly what the relationship is between instances of the lower-level properties and instances of the higher-level one.

A second strategy is to deny that higher-level properties are causally effective, insisting that they are nonetheless causally relevant because they 'program for' the instantiation of lower-level physical properties which are causally effective (Pettit 1993). Instantiations of the higher-level properties non-causally 'ensure' that the lower-level properties will be instantiated. This approach rejects the view that a necessary condition on the causal relevance of a property is the causal efficacy of its instance, but in doing so leaves it mysterious how an instance of a causally ineffective property can non-causally 'ensure' that a physical property is instantiated.[12]

[12] For further elaboration and discussion of Pettit's 'program' proposal regarding of the causal and explanatory relevance of higher-level properties, see Macdonald and Macdonald 2010, section 5.

Both strategies are unappealing in maintaining the distinctness not only of higher-level properties and the lower-level ones on which they supervene but also of their instantiations, thereby making it difficult to see how the lower-level properties could realize the higher-level ones. When properties are related as realizer to realized, however, by far the most plausible explanation is that an object's/event's instancing the higher-level one (e.g., being red, being camouflaged) just is its instancing the lower-level one (e.g., having a certain spectral reflectance frequency, being green), despite the irreducibility of the higher-level property to the lower-level one. As in the psychophysical case, the co-instantiation hypothesis is much the more attractive and metaphysically plausible one.

Why not simply respond to the causal irrelevance charge, then, by insisting that *Exclusion* is not compromised because mental and other higher-level properties of objects/events are not, given supervenience and the realization relation, independent of their microphysical realizer properties? Since supervenience itself is a co-variation relation, not a dependency one, the claim here would be that the fact that higher-level properties are realized by lower-level microphysical ones introduces an asymmetric dependency. This might seem to play into the hands of those who claim that such properties are causally impotent because their causal powers are exhausted by the causal powers of their realizers (cf. Kim 1998). However, on a basic difference-making account of causation (Menzies 2008, List and Menzies 2009, Woodward 2003), which focuses on properties rather than on their instances, mental properties, and higher-level properties more generally, can be seen to be causally relevant (in the sense set out by the property version of *Exclusion* alluded to above), despite supervening on and being realized by lower-level ones.

The 'difference-making' account, used by List and Menzies to argue that mental and other higher-level properties are causally efficacious, is borrowed from Pearl (2000), Hitchcock (2001), and Woodward (2003), though List and Menzies primarily focus on Woodward's 'interventionist' version and so the discussion here will be confined to that. However, the account is fairly widely shared and the specifics of the 'interventionist' version do not make a difference to the basic account.[13]

The bare essentials of a difference-making account of causation are encapsulated in the following bi-conditional:

> A variable *F* makes a causal difference to a variable *G* if and only if changes in (interventions on) the values of the *F* variable produce (are correlated with) changes in the values of the *G* variable.

However, much more needs to be said to flesh out the account. Without further elaboration, this is a kind of enhanced correlationist account, the enhancement being due to the idea that if we manipulate the values of the *F*-variable and find that this is accompanied by systematic changes in the values of the *G*-variable, then we can conclude that there is a causal relation between the values of the variables.[14]

List and Menzies flesh out the basic account by adding the following condition for difference-making:

> *Truth conditions for making a difference*: The presence of *F* makes a difference to the presence of *G* in the actual world if and only if it is true in the actual world that (i) *F* is present $\Box\rightarrow$ *G* is present; and (ii) *F* is absent $\Box\rightarrow$ *G* is absent (List and Menzies 2009: 483)

where the conditionals in (i) and (ii) are counterfactuals, understood in terms of the

[13] It is important to note, though, that Woodward's sees his interventionist account as aiming to "give an account of the content or meaning of various locutions, such as X causes Y " (2003, p. 38).

[14] Note too that it is not a reductionist account; there are too many causal features present in it for it to be a candidate for a reductive account.

standard Lewisian semantics for counterfactuals in which their truth conditions are specified in terms of a similarity relation between possible worlds. ($F \mathrel{\square\!\!\rightarrow} G$ is true in world w if and only if G is true in all the closest F-worlds to w.)

An example provided by List and Menzies helps to see how the account works to show that higher-level properties are causally efficacious. Suppose that there is an illness for which there is a drug which, given in an appropriate dose, causes recovery. The effect is a binary variable whose values are recovery/non-recovery, with the cause being a many-valued variable with possible values of, say, 0mg, 50mg, 150mg, and 200mg. The 'law' connecting this drug with this illness states that doses of at least 150mg are required for, and sufficient for, recovery; anything less results in death. Suppose now that a patient is given a 150 mg dose of this drug, and recovers. What is the cause of the patient's survival – that is, what can be substituted for F in the truth-condition for difference-making? According to List and Menzies, it cannot be 'Giving a dose of exactly 150mg' since this fails condition (ii) of the above condition (given that a patient given a dose of 200mg will also recover). 'Giving a dose of at least 150mg', however, satisfies both (i) and (ii) of the bi-conditional; so it is the difference-maker. ('Giving a dose of above 50mg' does not satisfy (i) either, so that is ruled out, according to List and Menzies.)

This is a very simple account of difference-making, though the account is given much further elaboration by Woodward, who introduces complexities to deal with cases exemplifying various problems that plague accounts of causation: over-determination, and especially what has been called 'late-pre-emption'. However, the cases we are considering, higher-level properties realized by lower-level ones, do not require the added complexities; the features essential for our discussion are contained in the bare version.

Turn now to the question of the causal relevance of mental properties. Suppose that in the actual world we have a mental/physical event causing a behavioural event. We can ask, in virtue of which property of that event, the mental one M, or the physical event P, does the behavioural effect B occur? The claim made by List and Menzies is that it is the M property, not the P one, that is causally efficacious vis à vis the B one; it is the difference-maker. The reason given is that, given multiple realizability of the M property, there will be closest nearby worlds in which M occurs and B occurs, but P does not occur (M being realized by a distinct physical property, say, P′). So M does, whereas P does not, meet condition (ii).

The difference-making account has the consequence of purchasing the causal relevance of the higher-level, mental property. Somewhat surprisingly, though, it looks like it *also* has the consequence, via *Exclusion*, of ensuring that the lower-level, microphysical properties that subvene on and realize the higher-level ones are *not* causally relevant properties vis à vis the B effect. And this is undesirable, since it is incompatible with the physicalist commitment of non-reductive physicalism.

However, the appearance of incompatibility is deceiving. The difference-making account rightly connects causal efficacy of properties to the *types* of effects instances of them have. But it is only if one assumes that mental and physical properties must be potential difference makers for the *same* type of effect that the difference making account has the consequence of ruling one, rather than the other, of these properties out as causally relevant to a single effect. In general, it is only on the assumption that the *effects* of events are events of only one type that *Exclusion* rules out one or the other of the lower-level or higher-level property as causally relevant. However, just as causes, being events, are of many types, only some of which are causally relevant to their effects, their effects themselves are of many types too. And a property's causal

relevance to an effect, being concerned with explanatory potential, will be sensitive to only some of that effect's properties as well as the causal/explanatory context. The moral is that properties are not causally relevant *tout court*; they are causally relevant in certain kinds of circumstances for certain types of effects and not others. Thermal conductivity is causally relevant for heating effects, not for electrical effects, irrespective of the fact that there is a law that correlates thermal conductivity with electrical conductivity (cf. Macdonald and Macdonald 1995).

The relevant reading of *Exclusion* in the psychophysical case, then, is that mental and physical properties of mental/physical events cannot both be completely and independently causally sufficient for an effect of a single *type*. However, behavioural effects of mental/physical events are of more than one type, since they are both of purely physical movement types and of action types. Mental properties of mental/physical events can therefore be causally relevant to their action-type effects, while their physical properties are causally relevant to their movement-type effects.[15]

The point is even clearer in the case of colour properties on the primitivist view. Colour properties are difference-makers for perceptual experiences as of them, and this does not compromise the causal relevance of their microphysical realizers. The reason, which examination of the psychophysical case reveals, is that colour properties can be co-instantiated with their microphysical realizers, despite multiple realization; and so too can their effects, perceptual experiences as of determinate colours. Effectively, as in the psychophysical case, instances of colour properties, as well as instances of their

---

[15] There is much more to be said about the psychophysical case but this must suffice for present purposes. For more on causal relevance and on how the co-instantiation thesis might work for properties that are not conceptually as well as metaphysically related (such as determinable/determinate ones), see Macdonald and Macdonald 2008 and Macdonald and Macdonald forthcoming. The considerations that are there brought to bear on the case of mental and physical properties work equally for colour properties and their microphysical realizers.

effects can be of (at least) two types. Colour properties are causally relevant to the perceptual effects they have in normal subjects in normal circumstances, which are co-instantiated with their neurophysiological effects in normal subjects' bodies. They are so because they meet conditions (i) and (ii) for being difference-makers. The microphysical realizers of colour properties are also causally relevant to the neurophysiological effects they have in the bodies of normal subjects. They are so because they meet conditions (i) and (ii) for being difference-makers.

Conclusion

Primitivism can meet the plausible constraints on a theory of colours, at least some of which are based on intuitions that Blackburn and I share about what a theory of colours should aim to explain. I have argued that two of the most prominent objections to it can also be met. This and its intuitive appeal should favour it, not only over response-dependence theories, but also over projectivist ones.

References

Armstrong, D.M. 1973. *Belief, Truth, and Knowledge*. New York: Cambridge University Press.

Bird, A. 1998. 'Dispositions and Antidotes'. *Philosophical Quarterly* 48: 227-234.

Blackburn, S. 1993. 'Hume on the Mezzanine Level'. *Hume Studies* 19: 273-288.

Blackburn, S. 1993/2010. 'Circles, Finks, Smells, and Biconditionals'. *Philosophical Perspectives* 7: 259-279. Reprinted in S. Blackburn, *Practical Tortoise Raising and Other Essays*. New York: Oxford University Press, 220-244.

Block, N. 1990. 'Inverted Earth'. *Philosophical Perspective*s 4: 53-79.

Byrne, A., and Hilbert, D. 2007. 'Color Primitivism'. *Erkenntnis* 66: 73-105.

Campbell, J. 1993. 'A Simple View of Colour'. In J. Haldane and C. Wright, eds., Reality, Representation, Projection. Oxford: Oxford University Press, 257-268.

Campbell, J. 2005. 'Transparency vs. Revelation in Color Perception'. *Philosophical Topics* 33 (1): 105-115.

Chalmers, D. 2006. 'Perception and the Fall from Eden'. In T. Gendler and J. Hawthorne, eds., *Perceptual Experience*. Oxford: Oxford University Press.

Dardis, A.1993. 'Sunburn: Independence Conditions on Causal Relevance'. *Philosophy and Phenomenological Researc*h 53: 577-598.

Gert, J. 2008. 'What Colors Could Not Be: An Argument for Color Primitivism'. *Journal of Philosophy* 105: 128-155.

Hitchcock, C. 2001. 'The Intransitivity of Causation Revealed in Equations and Graphs'. *Journal of Philosophy* 98: 273 – 299.

Holton, R. 1991. 'Intentions, Response-Dependence, and Immunity From Error'. In P. Menzies (ed.), *Response Dependent Concepts*. ANU Working Papers in Philosophy 1.

Jackson, F. 1996. 'The Primary Quality View of Color'. *Philosophical Perspectives* 10: 199 - 219.

Jackson, F. 1998. 'Colour, Disjunctions, Programming'. *Analysis* 58: 86-88

Jackson, F. and Pettit, P. 2002. 'Response-Dependence Without Tears'. *Philosophical Issues* 12: 97-117.

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

Johnston, M. 1998. 'Are Manifest Qualities Response-Dependent?' *Monist* 81: 3-43.

Kim, J. 1990. 'Explanatory Exclusion and the Problem of Mental Causation'. In E. Villanueva, ed., Information, Semantics, and Epistemolgy. Oxford: Basil Blackwell, 37-56.

Kim, J. 1998. *Mind in a Physical World*. Cambridge, MA: MIT Press.

List, C. and Menzies, P. 2009. 'Non-Reductive Physicalism and the Limits of the Exclusion Principle'. *Journal of Philosophy* 106: 475-502.

Locke, J. 1690/1975. *An Essay Concerning Human Understanding*. Ed. P.H. Nidditch (4th Edition). Oxford: Oxford University Press.

Mackie, J. 1973. *Truth, Probability, and Paradox*. Oxford: Clarendon Press.

Maund, J.B. 1995. *Colours: Their Nature and Representation*. Cambridge: Cambridge University Press.

Maund, J.B. 2008. 'Color. In Edward N. Zalta, ed., *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, URL = <http://plato.stanford.edu/archives/fall2008/entries/color/>.

McKitrick, J. 2003. 'A Case for Extrinsic Dispositions'. *Australasian Journal of Philosophy* 81: 155-174.

McLaughlin, B. 2003. 'Colour, Consciousness, and Colour Consciousness'. In Q. Smith and A. Jokic, eds., *Consciousness: New Philosophical Perspectives*. New York: Oxford University Press, 97-154.

Mellor, H. 1974. 'In Defense of Dispositions'. *Philosophical Review* 83: 157-181.

Menzies, P. 2008. 'The Exclusion Problem, the Determination Relation, and Contrastive Causation'. In J. Hohwy and J. Kallestrup, eds., *Being Reduced: New Essays on Reduction, Explanation, and Causation*. Oxford: Oxford University Press, 2008, 196-217.

Nolan, D. 2005. *David Lewis*. Chesham: Acumen Publishing.

O'Callaghan, C. 2007. *Sounds*. Oxford: Oxford University Press.

Pautz, A. 2006. 'Can the Physicalist Explain Colour Structure in Terms of Colour Experience?' *Australasian Journal of Philosophy* 84: 535-564.

Pearl, J. 2000. *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press.

Pettit, Philip. 1991. "Realism and Response-Dependence". *Mind*: 587-626.

Pettit, Phillip. 1993. *The Common Mind: An Essay on Psychology, Society and Politics.* New York: Oxford University Press.

Pettit, P. 1998. 'Noumenalism and Response-Dependence'. *Monist* 81: 112-132.

Prior, E., Pargetter, R., and Jackson, F. 1982. 'Three Theses About Dispositions'. *American Philosophical Quarterly* 19: 251-257.

Watkins, M. 2010. 'A Posteriori Primitivism'. *Philosophical Studies* 150: 123-137.

Wedgewood, R. 1997. 'The Essence of Response-Dependence'. *European Review of Philosophy* 3: 31-54.

Woodward, J. 2003. *Making Things Happen*. Oxford: Oxford University Press.

Yablo, S. 1995. 'Singling Out Properties'. Philosophical Perspectives 9: 477-502.

Yablo, S. 1999. 'Intrinsicness'. *Philosophical Topics* 26: 590-627.