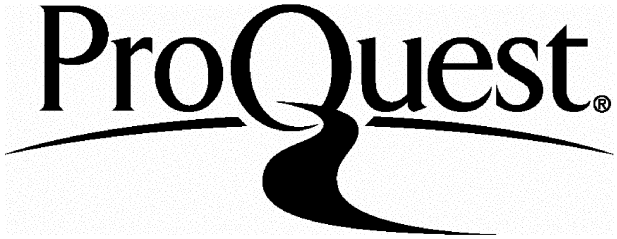
Colour, Dispositions and The Argument from Physics

Keith Allen

University College, London

MPhil 2002



Abstract

This thesis is an investigation into eliminativism about colour, the counter-intuitive thesis that the objects that surround us are not, despite appearances, coloured. A popular argument for this view – "The Argument from Physics" – is that because colours are not mentioned in a physical description of the world, a description which is assumed to be of ontological significance, then colours do not exist. One response to this argument is physicalism. Physicalists reject the eliminativist's claim that colours have no place within a physical description of the world, identifying colours either with disjunctions of physically describable properties or with physically specified dispositions. Because neither version of physicalism is pitched at the right explanatory level, however, neither is able to satisfactorily account for the similarity judgements about colour we make on the basis of visual experience.

Dispositionalism represents a different way of rejecting the eliminativist's claim that colours have no place within a physical description of the world. According to the dispositionalist, colours are logical constructions out of properties mentioned in the physical description of the world and the experiences these properties dispose their bearers to produce. As a response to the Argument from Physics, dispositionalism presupposes a metaphysical dispositional-categorical distinction. The only evident dispositional-categorical distinction, however, is conceptual.

The real problem with the Argument from Physics is the deference to physical theory that it presupposes. The only understanding of 'objectivity' strong enough to sustain the conclusion that describability in terms of physical vocabulary is of ontological significance, is *too* strong to support the eliminativist's conclusion. Rejecting eliminativism, we should conclude that colours are simple non-physical properties whose nature is entirely transparent to us.

Contents

Analytical Table of Contents

CHAPTER ONE: *Eliminativism about colour is the counter-intuitive thesis that the objects that surround us are not, despite appearances, coloured. A popular argument for this view – "The Argument from Physics" – is that because colours are not mentioned in a physical description of the world, a description which is assumed to be of ontological significance, then colours do not exist. Without engaging directly with this argument, one might suspect that eliminativism involves an incoherence. Even if this is so, however, it proves only that the argument is unsound; it does not enable us to identify which premiss is at fault and thereby determine what colours are.*

CHAPTER TWO: *Physicalists respond to the Argument from Physics by rejecting the eliminativist's claim that colours have no place within a physical description of the world. The primary quality view of colour identifies colours with disjunctions of physically describable properties. Physicalist dispositionalism, on the other hand, identifies colours with dispositions to either produce certain types of experience characterised neurophysiologically or to affect the light. Because neither version of physicalism is pitched at the right explanatory level, however, neither is able to satisfactorily account for the similarity judgements about colour we make on the basis of visual experience.*

CHAPTER THREE: *Dispositionalism, when understood as a thesis which presupposes a metaphysical distinction between properties that are dispositional and properties that are categorical, represents a different way of rejecting the eliminativist's claim that colours have no place within a physical description of the world. On this view, colours are logical constructions out of properties mentioned in the physical description of the world and the experiences these properties dispose their bearers to produce, that, though metaphysically derivative, still exist. Unfortunately, the dispositional-categorical distinction is not a metaphysical distinction; the only evident distinction in the area is a conceptual one.*

CHAPTER FOUR: *If colours are not (adequately) describable using purely physical vocabulary, then to avoid the conclusion of the Argument from Physics the deference to physical theory presupposed by the eliminativist must be rejected. Assuming that we are even able to give the notion of "the physical" any significant content, supposing that the physical exhausts what exists is deeply counter-intuitive. On one understanding of "objective", it may well be true that the physical world is the objective world; this understanding, however, is far too strong and does not support the eliminativist's conclusion.*

CHAPTER FIVE: *Non-reductive realism affords a more promising response to the Argument from Physics, rejecting both the ontological deference to physical theory presupposed by the eliminativist and the metaphysical deference presupposed by the dispositionalist. Identifying colours with simple properties whose nature is entirely transparent to us, the non-reductive realist (unlike the physicalist) is able to account for the similarity judgements that we make between colours on the basis of visual experience.*

CHAPTER ONE: Eliminativism

Eliminativism about colour is the view that the material objects that constitute the external world are not, despite appearances, coloured. This position dates back at least as far as Democritus, and even in something closely resembling its modern form can be found in writers from Galileo onwards. Thus Galileo writes:

> I think that tastes, odours, colours, and so on are no more than mere names so far as the object in which we place them is concerned, and that they reside only in the consciousness. Hence if the living creature were removed, all these qualities would be wiped away and annihilated (1957: 274).

More recently, the view that the world 'as it is in itself' – independent of our sensory interactions with it – lacks colour has become commonplace. It is, for example, the view of many, if not most, scientists working in the area. Cosmides and Tooby are fairly representative when they say:

> [Common sense] tells us, for example, that colour is out there in the world, an independent property of the objects we live among. But scientific investigations have led us, logical step by logical step, to escape our fantastically insistent, inelastic intuitions (1995: xi).

The view is popular even amongst philosophers, who, suitably impressed by the edicts of science, conclude that appearances to the contrary physical objects aren't really coloured. Jackson (1977), Hardin (1993) and Boghossian and Velleman (1989), for example, in different ways, all subscribe to this view.

1.1 Varieties of Eliminativism:

Perhaps the most popular argument for eliminativism starts from the observation that fundamental physical theory does not attribute colour properties ('colour-as-we-see-it') to material objects, but instead explains an object's colour-wise interactions with other objects – in particular sentient beings - entirely in terms of its primary, physically describable qualities.[1] From this, and a deference to physical theory,[2] the eliminativist concludes that, despite appearances, material

[1] The validity, if not the soundness, of this argument is agreed upon even by most colour realists. See, for example, Armstrong (1968: 289) and Jackson (2000).
[2] For example via "the philosophical principle of economy of postulation", Mackie (1976: 20).

objects are not coloured; colour is therefore in some sense 'subjective' or 'mind-dependent'. The argument may be explicitly stated as follows:

The Argument from Physics

(1) A physical description of the world exhausts what exists objectively.

(2) Colours have no place within a physical description of the world.

(3) Therefore, colours do not exist objectively.

Although agreeing upon the argument for their position, eliminativists disagree over what it entails. Traditionally popular is the view that if material objects are not coloured then there must be something else that *is* coloured, the properties of which we erroneously attribute to the objects in the external world. If it is required that this something be non-physical, we arrive at the projectivist sense-datum theory endorsed, for example, by Jackson (1977).

Closely related to Jackson's account is the projectivism of Boghossian and Velleman. According to Boghossian and Velleman, colour perception involves the projection of portions of the subjective visual field onto material objects, with the effect that these objects are represented, in the intentional content of experience, as the true bearers of that colour. The main difference between this view and that of Jackson is that Boghossian and Velleman's account is "potentially neutral on the metaphysics of mind" (1989: 95), depending on how the term "visual field" is understood.[3] In particular, they do not tie their eliminativism to a sense-datum theory of perception, allowing that "the visual field" may be neurophysiologically characterised, if so desired.

Precisely this suggestion was developed a few years earlier by Hardin, according to whom colour experience (indeed mentality *tout court*) is reducible to neural processes and is simply a product of our physiological makeup, explicable exclusively in neural terms (e.g. Hardin 1993: esp.109-112). For Hardin, all talk of the visual field is to be replaced by descriptions of neural processes which we have reason to expect will be "richer, more complete, and, in principle, more penetrable by the intellect" - although, as he himself admits, "at the present

[3] It should be stressed that Boghossian and Velleman's account is not neutral with respect to the *existence* of the visual field, only its metaphysical status. "Visual field" thus roughly corresponds

rudimentary state of our knowledge of the visual system, most of this is promise, program, and principle" (1993: 111).

The eliminativist, however, is not committed to the view that if physical objects are not coloured, then something else is. Consistent with eliminativism are theories of perception, which in the absence of genuine object-colour do not require the existence of suitable surrogates. The intentionalist, for example, who does not require that there be any entity, physical or otherwise, the awareness of which is essential to the possibility of visual experience, can maintain that the colour properties we project onto material objects have no bearers at all; they are nothing but simple qualities that exist in the intentional content of our experience.[4] Similarly the (disjunctivist) direct realist, who allows for the fact that our perceptual experience may be systematically misrepresentative in certain respects, can accept also eliminativism. Snowdon (1981), for example, who takes as his disjuncts veridical and illusory experience on the one hand and hallucination on the other, can allow that we are systematically mistaken with respect to the colour properties of objects whilst still being 'directly aware' of their other properties: visual experiences that are illusory in one respect can still be veridical, and so qualify as genuine perceptual experiences, in another.[5]

As such, eliminativism about colour, though counter-intuitive, is a very flexible thesis that can be accepted, in one form or another, by reductionists, anti-reductionists, sense-datum theorists, intentionalists and direct realists: if it is to be rejected, rejecting it on the basis of extraneous commitments in the philosophy of perception does not look promising. For those unwilling to accept the conclusion that material objects are not coloured, a different approach is needed. On the basis of a commitment to externalism about mental content, it might be argued that

---

to G.E. Moore's definition of "sense-data" as that which is the immediate object of perception, whatever that may be (see Mackie 1976: 48, n11).

[4] Price attributes this sort of position to Prichard (1947: 345). It is occasionally also attributed to Mackie (1976): this attribution receives qualified approval from Tye (2000: 166, n26) and is suggested, but rejected by Shoemaker (1994: 232, n4). Although an accurate model for his ethical theory, however, according to which values do not exist but the phenomenology is such as to make us believe that they do (Mackie 1977), this is not Mackie's official account of colour. His official position is that experience represents colour as an intrinsic, non-dispositional property of objects but in so doing it is mistaken; in fact colours are dispositions to produce certain kinds of experience in suitably placed suitably receptive observers.

[5] If, like Hinton (1967), *all* illusory experience is placed in the non-perceptual disjunct, however, we can concede that visual experience is systematically misrepresentative only if we concede that *no* visual experience is ever genuinely perceptual. This concession, however, has dire

eliminativism is conceptually incoherent. Alternatively, accepting the deference to physical theory presupposed by the eliminativist the assumption that colours lie beyond the scope of physical descriptions of the world may be disputed, either by arguing that colours are themselves physical properties or else by arguing that they are logical constructions thereof. If neither of these approaches are successful, the deference to physical theory presupposed by the eliminativist can itself be rejected. This thesis examines these responses.

1.2 The Argument from Incoherence:

The charge that eliminativism is conceptually incoherent is developed in detail by Stroud (2000: Chapter 7). Stroud's argument starts from the observation that there is clearly some connection between the colour properties that we attribute to objects in thought and those that are presented in visual experience: for instance, that there is some connection between the term 'yellow' as it is used in the descriptive sentences 'Jones sees that there is a yellow lemon on the table' and 'Jones believes that there is a yellow lemon on the table'. Assuming that perceptions of colour are not mere sensations with a certain distinctive character, in which case it is not clear that *any* intelligible connection can be forged,[6] the eliminativist needs to posit a direct connection between the colour properties that we attribute to objects in thought and those that are presented in visual experience; the word 'yellow', that is, must refer to the same property in both sentences.[7] But now, to conclude that physical objects are not coloured, the eliminativist must claim that neither our colour perceptions nor our colour thoughts, as thoughts and perceptions about physical objects, could ever be true. And this, Stroud claims, is problematic.

---

epistemological consequences and fails to respect the intuition that misperception in one respect does not preclude perception in another.

[6] As Stroud argues in chapter 5 and I simply assume here.

[7] If he posits an indirect connection, his view collapses into that of the dispositionalist, according to whom the colour properties that we attribute in thought are dispositional properties whereas the properties that we perceive "might be nothing more than seen instances of colour properties, or perhaps seen properties of things that are seen (or 'directly seen'), such as 'sense-data', or two-dimensional visual patches, or areas of one's 'visual field'" (2000: 119). The dispositionalist is unable to assume a direct connection between the contents of thought and perception because of the regress that would arise in trying to specify the content of perceptions of dispositional properties (2000: 141-4).

In order to give an 'unmasking explanation' of our colour beliefs – to explain away those beliefs without requiring that they be true, as we might explain away beliefs in ghosts or God – it is necessary to acknowledge that people have a great number of perceptions of, and beliefs about, colour. To attribute to others colour thoughts, however, the eliminativist must himself *at least* be able to grasp such thoughts: one cannot attribute to others the thought that lemons are yellow unless one knows what it would take to grasp that very thought - otherwise, how could one distinguish that thought from the thought that lemons are, for example, green?[8]

Understanding a given propositional thought, in turn, requires one "to know what would be so if it were true. And that is to know under what conditions it would be correct to judge or assert it to be true." (2000: 154) In other words, to attribute to others colour thoughts and perceptions, one must be able to specify the circumstances in which those thoughts would be true and those perceptions veridical. And this raises the following problem. How can the eliminativist specify the conditions under which it would be correct to judge that a lemon is yellow without appealing to at least some of the beliefs and perceptions that he is trying to reject, and thus rendering his position incoherent? Presumably, the circumstances in which it would be correct to judge that a given lemon is yellow are those in which we would have a yellow-experience as a result of seeing the relevant colour property of that lemon. But without making reference to his own colour thoughts and perceptions, the eliminativist will be able to give no content to the notion of a yellow-experience.

His only option, if he is to secure the existence of the *explananda*, is to hold that thoughts and perceptions are of properties, but not properties of physical objects: so, for example, they might be properties of objects that we 'directly' see, such as areas of the visual field, or, in the intentionalist's case, properties with no bearers at all. The eliminativist can then identify the thoughts and beliefs in question without himself having to believe that physical objects are coloured, and so without holding the contradictory beliefs that physical objects are both coloured and not coloured: ordinary thought, which is naively realistic about the objective existence of colour, attributes those properties that are perceived in

[8] The capacity for propositional thought is also a necessary precondition for the attribution, to

experiences of colour to physical objects, but in so doing it is systematically mistaken. In fact, our colour thoughts and perceptions are thoughts and perceptions about properties of the visual field or the intentional contents of experience.

But this too faces problems. Even if we allow that the eliminativist can identify the relevant thoughts and perceptions in his own case, there will be difficulties in moving from the first to the third person. The eliminativist will have to correlate *his* perception of regions of the visual field with non-colour facts about the world (he cannot correlate the perception of a yellow patch with the yellowness of lemons because *ex hypothesi* lemons are not yellow) and then generalise across all perceivers. But even if these correlations exist – which seems unlikely: "There is no distinctive feel, taste, or smell of those billiard balls that produce perceptions of yellow" (2000: 163) - if the properties of which the eliminativist is aware are private properties then this inference will be controversial, an idea traded upon by the inverted spectrum and absent qualia hypotheses.

But it is not even clear that the eliminativist will be able to identify the relevant perceptions in the first place, on Wittgensteinian grounds. Namely, what property is it that the subjectivist recognizes himself as perceiving? Is it that the region of his visual field is yellow, that it is a certain shape, that it is seen at certain time or in a certain position of his visual field? Ostension doesn't help either, as it leaves it indeterminate whether he is referring to the yellow patch, the event of seeing the yellow patch, the patch plus the surrounding area, or so on. In short:

> if we cannot attribute perceptions of and beliefs about the colours of things to anyone without ourselves having beliefs about the colours of objects, then the psychological facts that the project of unmasking the colours of things needs to explain cannot be acknowledged without our also accepting some non-psychological truths about the colours of things that the project means to deny. Accepting the relevant *explananda* violates a necessary condition of the project's success. Fulfilling that condition would render the relevant *explananda* unavailable. Either way, the project cannot succeed (2000: 169).

Stroud's argument against the eliminativist rests upon an externalist view of mental content. The idea is that unless the world were coloured, colour thoughts and perceptions would be impossible; the world, in a sense, 'reaches out' and determines the nature of our mental content. Content externalism, as an account

---

others, of colour perceptions (2000: 152-6).

of concept acquisition, is not uncontroversial: to discuss it fully, however, would take me too far afield. I shall therefore leave it open whether the eliminativist can plausibly resist Stroud's objection by taking a specific stance on the debate over the nature of mental content (perhaps, for instance, by arguing that our concepts are innately given). Even assuming that Stroud's reasoning is sound, however, the question of where this leave us remains.

According to Stroud, we are left in exactly the same epistemic position as we found ourselves before embarking on the metaphysical project: surrounded by physical objects that *appear* to be coloured. In particular, Stroud does not believe one should conclude from the eliminativist's failure to establish the metaphysical conclusion that physical objects really are coloured. To do so would:

> seem to require just what giving a negative answer required or presupposed – that we can get ourselves into a position to ask the question in the right way in the first place. We would have to be able to consider all human perceptions and beliefs concerning the colours of things, on the one hand, and the world as it is independently of us, on the other, and still manage to ask a still-open question about the relation between them (2000: 192-3).

Stroud's view can thus be summed up by the slogan 'objects are coloured but not really so'.[9]

This pessimism, however, is unwarranted. Against the eliminativist, Stroud argues that no specification of the conditions under which our colour judgements would be true is possible and consequently that the existence of the *explananda* that the eliminativist is trying to unmask cannot be acknowledged. But the strategy of the eliminativist is importantly disanalogous to that of the colour realist: they do not start from the same initial position. The colour realist has no qualms about the attribution of colour thoughts and perceptions. Before she can attribute these thoughts and perceptions to others she must be able to understand such thoughts herself, but given that this consists only in knowing under what conditions it would be correct to judge or assert it to be true that a given object is coloured, this condition is not difficult to fulfil. Specifically, it will be correct to judge that a lemon is yellow if it causes a yellow-experience in an appropriate way, where 'in an appropriate way' is a place holder for the colour realist's preferred account of what being coloured consists in. So, for example, on the

[9] This does not mean that Stroud is unable to distinguish between veridical and non-veridical perceptions of colour. His objection is to uses of 'really' that carry metaphysical force.

primary quality view of colour, it will be correct to judge that a physical object is coloured if the colour experiences that it produces in perceivers are caused by its primary qualities. Similarly, according to physicalist dispositionalism, the judgement will be correct if the experiences an object causes are a result of its disposition to affect the light in certain ways, or produce certain types of experience, neurophysiologically characterised.

The task facing the colour realist is far less demanding than that facing the colour irrealist. The colour realist starts with our colour thoughts and experiences and works backwards by trying to identify a condition (that to her is metaphysically acceptable) that suffices to render (at least the majority) of these beliefs and perceptions true; that is, to identify a common feature amongst the causes of these beliefs and perceptions. Essential to this strategy is the assumption that (most of) our thoughts and experiences are veridical. It is precisely this assumption that Stroud argues the eliminativist also requires. The difference is that it is an assumption that only the realist can consistently make.

Even if it turns out that Stroud's argument against the eliminativist is successful, then the *most* that it can show is that the Argument from Physics is unsound; nothing follows concerning the nature of the colour properties whose existence is thereby assured. Alternatively, if it turns out that Stroud's argument is unsuccessful then eliminativism remains a conceptual possibility. Either way, direct evaluation of the Argument from Physics is required.

CHAPTER TWO: Physicalism

Physicalism represents the most straightforward response to the Argument from Physics: by identifying colours with properties that are describable in terms of a physical vocabulary, the physicalist seeks to reject premiss (2) of the Argument from Physics outright.[1] This simplicity, however, is off-set by the need for complexity in the physicalist's account, in light of the empirical facts of colour perception.

Hardin (1993: 2), following Nassau, lists fifteen types of physical property that 'standardly' cause an object's perceived colour (these include molecular vibration, incandescence and the scattering of light). Take as an example the class of things that are perceived to be yellow. There is no single (micro-) physical property that yellow lights, the sun, lemons and sunflower oil all share (and this is to say nothing about 'subjective colours' – for example yellow after-images – which appear to have no physical bearers in the first place). Even restricting the discussion to *surface* colours, variation amongst objects that are perceived to be yellow is significant enough to suggest that there is no unique physical property that they all instantiate. Plastics, glass, rocks, paper (etc) can all be seen to be yellow even though they have radically different physical compositions. Moreover, because of the phenomenon of metamerism, there is not even any straight-forward connection between perceived colour and the character of the incident light that yellow objects are disposed to reflect, and so we cannot say that yellow objects all share a property that causes them to reflect light of the same composition. Objects that reflect radically different proportions of the incident light that strikes them, for instance, can nevertheless all appear to be the same colour providing that the light entering the eye of the perceiver is such that it is of the right composition to issue in the relevant opponent processing.[2] All that these

[1] More detailed consideration of what is meant by 'physics' and 'physical vocabulary' is postponed until Chapter 4.

[2] There is a corresponding problem for coloured lights: there are an indefinite number of light combinations that when mixed will produce in a standard observer an experience of yellow. Most importantly, these lights need not be drawn from the 'yellow' part of the spectrum: a mixture of

disparate objects appear to have in common is that they reflect light that produces distinctive types of experience in suitably placed observers.

Physicalists tend to react to these facts in one of two ways. They either identify colours with the disjunction of the configurations of primary qualities that are individually sufficient to bring about colour experiences of a given type or else they identify them with physically specified dispositional properties. Either way, however, they face problems.

2.1 The Primary Quality View of Colour:

The primary quality, or 'Australian', view of colour identifies colours with disjunctions of physical properties.[3] These disjunctive properties are not, given the empirical facts of colour perception, "perfectly natural properties" (Lewis 1983: 192) - they are not amongst those properties that are fundamental to physical enquiry.[4] But they are properties nevertheless; namely intrinsic (non-relational) properties that are disjunctions of, and therefore grounded in, an object's perfectly natural properties.

As such, the primary quality view respects two very important intuitions about colour. The first, *pace* the eliminativist, is that there is no *ontological* difference between properties like colour and properties like shape: colours, like shapes, exist. The second, assuming that naturalness is a property of properties that comes in degrees (Lewis 1983: 193, Hirsch 1993: 74-9), is that there is no qualitative *metaphysical* difference between properties like colour and properties like shape, either; both are intrinsic properties of objects. The second of these intuitions is more controversial than the first (we shall see below that it is rejected by

---

'green' light that is 550 nanometers ('nms') and 'red' light that is 650 nms, for example, exactly matches a 'yellow' light of 575nms (see Hardin 1993: 63, n6).

[3] See, for example, Armstrong (1968, 1987), Smart (1975) and Jackson (1996, 2000 and with Pargetter 1987). A similar view is suggested by Kripke's remark, "the reference of 'yellowness' is fixed by the description 'that (manifest) property of objects which causes them, under normal circumstances, to be seen as yellow (i.e., to be sensed by certain visual impressions)" (1980: 140, n71). Given that there is no physical property common to all and only things seen to be yellow, Kripke must either: (i) reject the assumption that the property of objects which causes them to be seen as yellow is a physical property, (ii) conclude that many of the objects we commonly assume to be yellow are not in fact yellow, or (iii) take the micro-structural property which determines the reference of 'yellow' to be disjunctive. (ii) is counter-intuitive; (i), though much more attractive, is apparently ruled out by Kripke's earlier remark, "Yellowness certainly expresses a manifest *physical* property of an object" (1980: 128 n66, emphasis added); this leaves (iii) - colours are disjunctive properties of objects.

dispositionalists), but the idea that colours are, at the very least, experienced in naïve perceptual consciousness as intrinsic properties of objects has a distinguished history and is not obviously misguided.[5]

The primary quality view of colour is nevertheless deeply problematic. For a start, not everything that appears to instantiate colour properties has primary qualities with which to identify these colours. After-images, for example, can appear coloured even though they are plausibly not physical objects, and therefore have no physical properties with which their colour may be identified. A common physicalist response to this problem is to put after-image perception down to hallucination (a red, square after-image is no more red than it is square); it is no objection to a physical theory of colour that after-images are non-physical entities that nevertheless appear coloured if the appearance is a *mere* appearance. There is, however, a problem with this response. Typically, coloured after-images (a result of 'successive contrast') are taken to be only one of the types of 'subjective colour' that we can experience, including amongst others coloured shadows, the colours caused by spinning (black and white) Benham discs in bright tungsten light, the 'filling in' that occurs in the area of the visual field corresponding to the blind spot and simultaneous contrast effects, in which the perceived colour of one patch of colour changes when viewed against backgrounds of different colours (Hardin 1993: 91). Hardin remarks that the reason for treating these types of colour experience collectively is that, "The colours experienced in after-images, coloured shadows, and simultaneous contrast [etc.], are explicable in terms of the operation of nervous systems and cannot plausibly be supposed to exist apart from them" (1997: 289).[6] If the principle of 'chromatic democracy' that this observation suggests is accepted, however, the physicalist is committed to the view that these other phenomena involve illusion as well. And, at least in the case of simultaneous contrast, this commitment causes difficulties.

---

[4] As Quine remarks, "the distinctions that matter for basic physical theory are mostly independent of colour contrasts" (1969: 49).

[5] See, for example, Locke (1975: 2.8.25) and following him Mackie (1976: 11); this intuition is exploited by Berkeley (1998: 1.10) in his argument against Locke's primary-secondary quality distinction.

[6] There is a sense in which *all* colour experience is explicable in terms of the functioning of the visual system (to have a colour experience just is for your visual system to be functioning). Hardin's point, however, is that these colour experiences are "inextricably bound up" with the function of the *human* brain; there may be ways of experiencing colour that do not yield experiences of these kinds.

Simultaneous contrast is a common phenomenon that affects colour perception of all types. Assuming that it is the result of illusion would mean that a large proportion of colour perception is rendered, to some extent, illusory. Perception of the 'pure contrast colours' (black, white, grey and brown), would be particularly badly effected as these can *only* be seen relative to a certain backgrounds; their perception *depends* on simultaneous contrast. Black objects, for example, are only seen to be black if they are perceived relative to a background of a lighter colour. Conversely, to look white, it is necessary (although clearly not sufficient) that an object be seen relative to a darker background. If our perception of these colours almost always involves simultaneous contrast – a species of illusion - however, then hardly any pure contrast colour experience will ever be veridical; we will have to conclude that almost no object is really white, black, grey or brown. And selective colour irrealism is, if anything, only slightly more acceptable than the eliminativist's more general irrealism; if we reject one, then we should also be disdainful of the other.

At this point, one might, following Tye (2000: 150-8) try a damage limitation exercise, conceding that whilst simultaneous contrast does involve perceptual illusion, colour perception does not differ fundamentally in this respect from shape perception, which is also subject to background effects in the form of the Hering and Orbison illusions (colour illusion is just more frequent than shape illusion). Moreover, selective irrealism about colour does not follow from the fact that perception of the pure contrast colours almost invariably involves colour contrast, if we distinguish between object colour and perceived (pure contrast) colour. So, for example, if we distinguish the actual colour black from the pure contrast colour BLACK, such that something is BLACK iff it is black and darker than its surroundings (BLACK is thus a compound, relational property), then although most of the time objects appear will appear BLACK, this BLACKNESS is parasitic on the actual colour, black (blackness will need to be characterised in independent terms, but overlook this at present). Crucially, visual experience will not be *necessarily* misrepresentative with respect to black, white, grey and brown, as occasionally it will be *these* properties, as opposed to their pure contrast counterparts, that we perceive: furniture in an unlit room, for example, will be

black but not BLACK; an enveloping light brown fog will be brown but not BROWN; and so on.

In the context of Tye's own perceptual theory this response is unusual to say the least, entailing that our colour experience is largely illusory: part of Tye's argument for a representationalist theory of perception is that the supposition that visual experience is systematically misleading in the way required by sense-datum theories of perception is "just not credible" (2000: 46). Admittedly the kind of error presupposed by sense-datum theories of perception differs from that required by Tye to explain our perception of the contrast colours: the first case involves the *necessary* confusion of one *object* for another, whilst in the second case the error lies in the *almost inevitable* confusion one *property* for another. But if, as Tye appears to suggest (2000: 53), the incredibility claim is merely an inference to the best explanation (i.e. it much more straightforward to suppose that visual experience is not systematically misrepresentative), there seems to be no reason to think that the argument does not generalise: if it is more plausible to suppose that visual experience is not systematically misleading in the one case, then the same should go for the other.

More generally, the attribution of wide scale perceptual illusion in respect of colour is methodologically problematic. On the basis of visual experience, we assign things to the extensions of the different colour terms. To rule that some of these experiences are not veridical, and therefore than some of these things are not genuine members of the extensions to which they are assigned, there have to be grounds for distinguishing between the genuine and the non-genuine. Whatever grounds we chose, however, will be controversial. This is because colour perception is importantly disanalogous to shape perception: if we undergo an illusion with respect to the primary qualities, we have a method for determining whether or not our experience actually is illusory - we can measure the angles and the length of the sides of the shape in question. The same is not true of colours. In the case of colour there is no higher court of appeal than visual experience.

Just because any grounds we chose are likely to be controversial does not preclude the existence of *some* distinction between illusory and veridical perception of colour properties, but it does suggest that unless the grounds for this distinction are well founded we should try to minimise the discrepancy between

the pre- and post-reflective extensions of our colour terms as far as possible. In assuming that the perception of simultaneous contrast involves perceptual illusion, the colour realist places a lot of weight on his basic assumptions; in Tye's case, these assumptions are physicalist, and therefore (as we shall see below) questionable. But whatever they are, it is better not treat simultaneous contrast as a species of illusion if at all possible: no basic assumptions will be *that* certain.

But if the physicalist does not treat simultaneous contrast as a species of illusion, does this mean that he cannot also treat after-images as illusions, and therefore that the initial objection holds? No: even though both effects are inextricably bound up with the human nervous system, they are importantly disanalogous. After–images exhibit a number of distinctive traits: they are, for example, presented at no depth in the visual field, do not cast shadows and move with the eye. As such, they are unlike simultaneous contrast effects, in which the colour is still represented as an external quality; after-images, it is not unreasonable to suggest, are not even *experienced* as physical objects. This makes it much easier to put after-image experience down to illusion, whilst at the same maintaining that simultaneous contrast is not illusory, because after-image experience wears its non-veridicality on its sleeve.

An account of simultaneous contrast that does not threaten to lend support to the eliminativist's denial of the existence of colour or else render effectively all colour experience illusory is still required. But this turns out to be relatively unproblematic if the holding of the relevant background facts is written into the account of normal conditions for the perception of the contrast colours, and it is not insisted that when these conditions are not met our experience is in some sense illusory. Just as some people 'don't seem themselves' early in the morning, so you get a much better idea of what the contrast colours look like if you see *them* under the right conditions, too. But this is not to say that even in these less than perfect conditions, you do not see the colour in question, just as you still see the person in the morning. Properties seen under non-standard conditions are the same properties as are seen under standard conditions; there is no illusion.

Even if the physicalist can write-off after-image perception as illusory without the rot spreading, however, there are other things that appear to instantiate colour properties but which seem to lack primary qualities with which to identify these

colours, for which this kind of treatment is less appropriate. Rainbows are a case in point. Rainbows, unlike after-images, are presented perceptually as spatially located and are not seen to change this location merely by moving the eye. Moreover, unlike after-images, they are publicly observable phenomena that can be photographed (and at least under normal conditions, the camera never lies – it photographs only what is there to be photographed). Writing rainbows off as mere illusions is therefore far more implausible than writing-off after-images similarly.

The physicalist's best response is to try to reconcile his physicalism with the existence of rainbows, identifying rainbows with the collection of water droplets that are responsible for the necessary light diffraction and locating the colour accordingly. But even this identification is problematic. Specifically, collections of water droplets lack many of the properties borne by rainbows, which by an application of Leibniz's Law would show them to be non-identical. So, for example, a rainbow can appear exactly the same even though the collection of droplets with which it is identified are constantly changing. This example is contentious, as the discussion of the corresponding problem in personal identity indicates, but is not obviously beside the point. In discussions of personal identity, for example, part of the reason for adopting a criterion of psychological continuity is that every seven years or so, the molecules that make up a human body are entirely replaced, which makes identifying people with the sum of their micro-physical parts rather difficult; on this criterion, I would have been four different persons by now. And at least *prima facie*, this is reason to reject physicalist theories of personal identity.

More seriously, however, the rainbow and the collection of water droplets with which it is identified need not even be perceived to instantiate the same spatial location. Rainbows are nearly always perceived to be on the horizon, yet the collection of water droplets with which they are identified need not be. To write this off merely as misperception would appear to be deeply uncharitable to visual experience; but unless something like this manoeuvre is attempted, the primary quality account of colour will not generalise to cover the colours of phenomenal objects like rainbows.[7]

[7] Allowing for the existence of phenomenal objects like rainbows provides a further reason (to those presented below) for relaxing the physicalist criterion for objective existence expressed in premiss (1) of the Argument from Physics. Once it is conceded that objects need not be physical

Even restricted to a thesis about physical object colour, however, the primary quality view faces problems. To accommodate the empirical fact that the set of (physical) yellow things exhibits spectacular diversity, the primary quality theorist has to concede that the property that all these things share will have to be (probably) highly disjunctive. But many would rather conclude either that colours do not exist or that the primary quality view of colour is simply wrong, than that colours are disjunctive properties.[8]

One common source of hostility to disjunctive properties is the belief that they are causally inefficacious. Assume, for the moment, a counterfactual analysis of causation, according to which property p of object c causes event e iff if c had not been p, then e would not have occurred. Now, oversimplifying, assume that there are only two types of physical state that will make an object look yellow (s1 and s2). The problem is this: c's being s1 cannot be the cause of e (where e is an experience of yellow), because it is not necessarily true that if c had not been s1, then e would not have occurred; c's being s2 would also have caused e. *Pari passu*, c's being s2 cannot be the cause of e, because e could have occurred in the absence of s2 (namely if c were s1). However, if neither s1 nor s2 can cause e, then the disjunction of these two properties certainly cannot.

The first thing to notice about this argument is that, if valid, it shows only that disjunctive properties are causally inefficacious, not that they do not exist. Causal efficacy may well be too strong a criterion for existence: even though we do not have any reason to posit the existence of causally inefficacious properties or objects, it does not thereby follow that they do not exist; for all we know they might (only assuming something like a Principle of Sufficient Reason does this inference look at all appealing). Nevertheless, for the physicalist a lack of causal efficacy is bad enough: if colours are disjunctive properties and disjunctive properties are causally inefficacious, then assuming that perception is a causal process, what we see when we look at the world around us will not be colours. And this renders the physicalist theory otiose.

---

objects and colours need not be physical properties, accounting for the existence and colour of rainbows becomes relatively unproblematic.

[8] Armstrong, for example, once opted for the first of these alternatives: if it "is simply a matter of *irreducibly* diverse causes in the physical surfaces bringing about identical colour-appearances for human observers...I think we would then have to conclude that colour is a pseudo-quality" (1968: 289). Subsequently, however, his hostility to disjunctive properties waned (1987: 40).

Nevertheless, this objection is far from conclusive: as it stands, there is no more reason to reject the causal efficacy of disjunctive properties than there is to reject (simple) counterfactual analyses of causation which, on independent grounds, are highly questionable anyway. Aside from attempting to amend the counterfactual account to accommodate this sort of problem, we could just reject it outright and replace it with a theory of causation on which this problem does not arise in the first place: on a simple sufficiency account of causation, for example, according to which c's being p causes e iff c's being p is *ceteris paribus* sufficient for e, either c's being s1 or c's being s2 would cause e because both are sufficient for it. Either way, it does not obviously follow that disjunctive properties are the problem.[9]

A more genuine worry with highly disjunctive properties consists in filling out the disjunction without adding a 'whatever it takes' clause. For a property to be of epistemological significance – to non-trivially explain or facilitate the predication an object's interactions with other objects – there have to be clear criteria of instantiation which are not too easy to satisfy. Explanations impart information; it must therefore be informative to say that one thing is the cause of another. However, if the likelihood of the effect occurring, given the cause, is near certainty, this condition will not be met; there has to at least be the chance that the one event (or fact...) will not be followed by the other. The problem in the case of colour is that there appears to be no obvious limit to the types of physical object, exhibiting novel physical structures, that can be seen to be yellow (indeed even those that are actually yellow are very varied). There are, that is, apparently no clear criteria of instantiation for colour properties.

No doubt it is, at least in principle, possible to specify the full range of physical states that objects perceived to be yellow could instantiate. But then presumably, the same is true of, for example, comic situations: there will be some highly complex disjunction which describes the physical states of the things that partake in every conceivable comic situation. But just as it isn't very illuminating to say that seeing someone slipping on a banana skin is funny because the banana

[9] A different metaphysical problem arises, if like Armstrong (1978), we assume both that all genuine properties are universals and that there can be no disjunctive universals. Again, however, and as Armstrong himself later concedes (1987: 40), this does not force the rejection of the primary quality view of colour as the option remains of simply rejecting one or both of the original assumptions.

skin has physical properties a, b and c and the man has properties d, e and f *or* the banana skin has properties g, h and i and the man has properties j, k and l *or*...so, the identification of colours with disjunctions of primary qualities is similarly imperspicous.

Depending on how you understand the term, these primary quality accounts of colour and comedy may still constitute explanations: on Lewis's view, for example, whereby "to explain an event is to provide some information about its causal history" (1986: 217), an account of an experience of yellow in terms of the primary qualities of the objects in the subject's environment *is* an explanation.[10] But we must not be misled by terminology. If we assume a weak criterion for explanation, like that proposed by Lewis, we then require a further distinction between those explanations that are good and those that are not: the information must be relevant, comprehensible, and easy to manipulate. More demanding accounts of explanation will reserve the term 'explanation' for what are, on Lewis' account, good explanations. Either way, the explanations of colour perception and comedy in terms of basic physics are not what we're looking for. The primary quality view of colour is pitched at the wrong explanatory level.

Because the primary quality view of colour is pitched at the wrong explanatory level, it violates a crucial adequacy constraint on any putative theory of colour. Visual experience grounds the judgements that we make about the colours. It is purely on this basis that we judge of the structural relations that hold between the colours: for instance, that orange is more similar to red than red is to blue; or that red, green, yellow and blue are unitary hues, determinates of which can appear pure and unmixed, whilst purple, orange, pink and yellow-green are binary hues, determinates of which always appear to be a mixture of two unitary hues. An adequacy constraint on any putative theory of colour is that it explains our judgements concerning these relations. By identifying colours with disjunctions of properties that fail to display any substantial metaphysical unity, the primary quality view of colour violates this constraint: why we should judge that red is more similar to orange than it is to blue appears totally inexplicable if there is no

[10] For the explanation relation, as opposed to epistemically qualified relations such as '*known* to explain', the "explanatory equivalence principle" that Lewis' account thus respects does indeed look attractive: "If S and T are (*a priori* necessarily) equivalent, respectively, to S' and T', and S is the explanans in a correct explanation in which T is the explanandum, then S' is the explanans in a correct explanation in which T' is the explanandum" (Hirsch 1993: 80).

basis for this judgement in similarity relations that hold amongst the properties perceived (unless, of course, these judgements are incorrect, in which case visual experience would once again uncharitably be accused of systematic misrepresentation).[11] Colours, then, are best not identified with primary qualities of objects.

2.2 Physicalist Dispositionalism:

Physicalist dispositionalism represents a different way of attempting to reject premiss (2) of the Argument from Physics consistent the empirical facts of colour perception. Physicalist versions of dispositionalism typically come in one of two forms.[12] First, are those that characterise the visual experiences that coloured objects tend to produce in scientific terms - for example, in terms of the visual experiences *characterised psychophysically* that they are disposed to produce.[13]

The human eye contains three types of cone, one sensitive to short-wave light (S), one to medium wave light (M) and one to long wave light (L). These cones are connected to three neural channels, each responsible for different types of colour-experience: one achromatic, one red-green and one yellow-blue. Depending on the values assigned to S, M and L, these channels code different colours and so issue in different colour experiences. For example, the sum of L+M determines the coding of the achromatic channel: where '0' represents the base level of neural activity, if L+M>0, then the achromatic channel codes white, if L+M<0 then the achromatic channel codes black, and if L+M=0 then the achromatic channel codes "brain grey". Similarly, L-M determines the red-green signal: if L-M>0, then it is red and if L-M<0, then it is green. Finally, (L+M)-S determines the yellow-blue signal: if (L+M)-S>0, then it is yellow and if (L+M)-S<0, then it is blue. On Smith's proposal, green objects, for example, are those objects that are disposed to produce opponent processing characteristic of green experiences, in which the red-green channel codes green and the yellow-blue channel codes zero.

---

[11] This sort of objection can be found in, amongst others Hardin (1993) and Boghossian and Velleman (1991).

[12] Consideration of what is meant by identifying colours with *dispositions* is postponed until Chapter 3.

The second type of physicalist dispositionalism identifies colours, not with dispositions to produce experiences suitably spelt out, but with dispositions to affect light, again described in physical vocabulary.[14] Coloured objects reflect different amounts of light at most spectral wavelengths (this is part of the reason why Armstrong's (1968) identification of colours with light emissions of different wavelengths is too simplistic: yellow objects, for example, do not just reflect light from the 'yellow' part of the visible (*to humans*) spectrum, which falls roughly between 570 and 590 nanometres). For each wavelength at which an object reflects light, it is possible to specify the percentage of light reflected. These percentages can then be used to generate a surface reflectance profile (SRP), which in turn determines the colour that any object will appear to have.

Achromatic objects tend to reflect roughly the same percentage of light at each wavelength, with ideal white objects reflecting 100% of the light that strikes it, ideal black objects reflecting none of the light that strikes it and grey objects reflecting some intermediate percentage (roughly 20%). Chromatic objects, on the other hand, reflect different percentages of light at different wavelengths. Green objects, for example, tend to reflect more light in the middle-wavelength part of the visible spectrum than in the long-wavelength part, and roughly the same amount of light in the short-wavelength part of the spectrum as in the long- and middle-wavelength parts put together.

Any two objects with identical SRPs appear identical to normal observers in standard conditions. Unfortunately, so do many objects that differ radically in their SRP. This phenomenon (metamerism) constitutes a *prima facie* problem for the light-affecting physicalist dispositionalist as it blocks the straightforward identification of colours with SRPs: if two objects can appear to be the same colour in standard conditions, and yet differ radically with respect to their SRP, then perceived colour cannot be identified with SRP.

The physicalist can avoid this problem, however, by identifying colour, not with particular SRPs, but rather with *types* of SRP that issue in a similar net

---

[13] See, for example, Smith (1987). See also Harman: "objective color is plausibly identified with a tendency to produce a certain reaction in normal observers, where the relevant reaction is identified in part with the ["*biological*"] mechanisms of color perception" (1996: 259).

[14] See, for example, Hilbert (1987), Byrne and Hilbert (1997b), Hilbert and Kalderon (2000) and Tye (2000). Similar accounts are proposed by Westphal (1987) and Broackes (1992), although, these authors do not require that the relevant dispositions be specified in physical vocabulary. See Chapter 5.1 below for more details.

effect. So, for example, even though two apparently identical objects may have radically different SRPs in certain respects, with one reflecting a high percentage of light at a given wavelength, λ, and one reflecting almost no light at this wavelength, all that matters is their overall profile: in the case of a green object, for example, that it meets the rough specifications given above. The relevant type of SRP shared by green objects is thus one that, *as it happens*, issues in the opponent processing associated with green experiences: green objects reflect more medium-wave than long-wave light, so the red-green channel codes green, and approximately the same amount of short-wave light as the amount of medium- and long-wave light put together, so that the yellow-blue channel codes nothing, thus yielding an experience of green.

Physicalist dispositionalism of either variety has to respect the adequacy constraint on theories of colour identified in the last section: they have to be able to explain our judgements concerning the internal relations that hold between the perceived colours. This turns out to be problematic, however, particularly for physicalist dispositionalism of the latter variety. Like the primary quality theorist, the light-affecting physicalist dispositionalist needs to locate some similarity between the colours, conceived of as he conceives of them, that accounts for the relations that we judge to hold between the colours purely on the basis of visual experience. The difficulty lies in determining what this similarity is.

Byrne and Hilbert's response (1997b) to this problem is to distinguish between 'relative' similarity (similarity relative to a family of properties) and 'natural' similarity (similarity *simpliciter*), claiming that what we know on the basis of visual experience is just the *relative* similarity claim that orange is more similar to red than red is to blue. Identifying colours with types of surface reflectance profile, of which objects can instantiate more than one at any given time (orange objects, for example, are simultaneously represented as SRP_{ORANGE}, $SRP_{RED\text{-}OR\text{-}ORANGE\text{-}OR\text{-}PURPLE}$, i.e. reddish, and $SRP_{COLOURED}$), Byrne and Hilbert explain the greater similarity between red and orange than between red and blue on the grounds that red and orange objects instantiate more of the same SRPs than blue objects (both instantiate the second and third SRP, whilst blue objects instantiate only the third).

However, even assuming that it is the relative similarity claim, as opposed to the natural one, that we want the physicalist to respect, this response is inadequate

because it does not generalise across all the internal relations that we judge to hold between the colours. The similarity that orange and red bear to each other, for example, and that neither bears to blue, is explained by Byrne and Hilbert by the fact that orange and red objects instantiate the property $SRP_{RED\text{-}OR\text{-}ORANGE\text{-}OR\text{-}PURPLE}$, whilst blue objects do not. But suppose that red, in turn, is more similar to orange than it is to purple. Byrne and Hilbert cannot explain this by generalising their theory and claiming that red and orange objects instantiate the common property $SRP_{RED\text{-}OR\text{-}ORANGE}$, whereas purple ones do not, therefore red and orange objects instantiate more shared properties than red and purple objects, because red and purple objects also instantiate a common property, $SRP_{RED\text{-}OR\text{-}PURPLE}$; red and orange things instantiate exactly the same number of representable colour properties as red and purple things.

Clearly, this only a problem if red really is more similar to orange than purple; but unfortunately for Byrne and Hilbert it *is*. The human colour space is asymmetrical. One way of representing this fact is in terms of the Munsell colour system. The Munsell colour system divides the colour space up into 100 equal sized perceptual hue steps. The number of steps between red and yellow is 23, whilst that between red and blue is 31 (Kuehni 1997: 61). Assuming that the binary colours orange and purple lie half way between the two sets of unique hues, it follows that red is more similar to orange than it is to purple, given that the latter is 'further away' from red than the former. Byrne and Hilbert, then, are unable to account for the structural relations that we judge to hold between colour properties on the basis of visual experience.

Hilbert and Kalderon (2000) propose a more sophisticated account of light-affecting physicalist dispositionalism, introducing a role for the observer in determining the identity of colour properties, in order to accommodate our judgements concerning the relations that hold between the colours and thereby avoid this problem. According to Hilbert and Kalderon, our ability to sort colour samples into similarity classes depends on the similarity relations between the colours (modelled in the colour space) that objects instantiate. These similarity relations are in turn fixed by a biologically determined classificatory function of the visual system, i.e. the way that we classify colours fixes the relations that each colour bears to every other colour. Finally, the content of a colour experience is determined (partly) by "*all* the relations of comparative similarity in which red is

represented as standing vis-à-vis the other colours" (2000: 198); that is to say, the content of a colour experience is determined partly by its position in the colour space. The qualification 'partly' here is important because of the possibility that different systems of similarity class might all exhibit the same structure: all our biology determines is that the similarity classes exhibit the structure of the colour space, not that they exhibit it uniquely. This further factor comes from the relations that the subject bears to his environment, i.e. colours cannot bear relations to each other in a subject's colour space that they do not bear to each other in the world; colour properties must be instantiated in collections of things that jointly exhibit the structure of colour space.

The advantage of this account in the present context is that it guarantees *a priori* the discoverability (given sufficient reflection) of the similarity relations that hold between the different colours, on account of the fact that these similarities actually determine the content of our colour experiences. Moreover, it does not fall into the trap, noted by Boghossian and Velleman (1991: 128-30), of crediting us with knowledge of all the similarity relations constitutive of the colour space on the basis of a single experience: on this account, the content of an experience is not specified by explicitly *describing* the colour's location in colour space (which requires knowledge of the full extent of that space), but rather by its *location* in that space. Physicalism of this sort, then, looks to be capable of accounting for our judgements concerning the structural relations that hold between the colours.

However, flattening one bulge in the carpet only causes others appear elsewhere. For one thing, there is the question of how the biologically determined classificatory function of the visual system fixes the similarity relations that hold between the colours. Opponent process theory is only so successful in meeting this challenge. Whilst it does determine the opponency of the unique hues, red and green and yellow and blue - the red-green channel, for example, cannot simultaneously code red and green; it codes red when the signal it receives is 'positive' and green when the signal it receives is 'negative' – at least in its present form, it cannot be used to explain why our colour space is asymmetrical; we currently have no explanation of how our biology fixes the fact that red looks more like orange than it looks like purple. This much must remain programmatic.

A more fundamental problem concerns the possibility of spectral inversion. If the content (both phenomenal and representational) of our colour experience is determined by the similarity relations it bears to other experiences, then in a symmetrical colour space, all experience will be phenomenologically identical.[15] Assuming, as the intentionalist about perceptual experience does, that there can be no difference in phenomenological character without a corresponding difference in representational content, this means that if our colour space were symmetrical, all our experience would be exactly the same; in effect, that we would have no experience at all.

But this is a very strong claim indeed. For a start, it is open to empirical disconfirmation. More importantly, the claim looks much less plausible for simple achromatic colour spaces, in which there is only one dimension along which colours vary (brightness - from white to black through grey), than it does for our three dimensional space colour space.[16] To this Hilbert and Kalderon object that imagining the world in black and white is not the same as imagining what visual experience would be like for a creature whose colour space consisted entirely of these two qualities (2000: 206); such a space would be radically different from our own, and it is therefore far from obvious that the target state of affairs can even so much as be conceived. But even accepting this claim, it at the very least exacerbates the biological determination problem and increases the explanatory debt that Hilbert and Kalderon owe. Specifically, they now need to explain why our biology could not *possibly* be such as to realise a symmetric colour space. In the absence of such an explanation, this begins to look more like a reductio than a welcome consequence of Hilbert and Kalderon's dispositional physicalism.

Hilbert and Kalderon attempt to meet the adequacy constraint on theories of colour by explicitly introducing the observer into their account. This strategy is more clearly in evidence in the physicalist dispositionalism of Smith (1987), who identifies colours with dispositions to produce neurophysiologically characterised experiences of a certain sort. Experience-introducing versions of physicalist dispositionalism are able to account for our judgements about the structural

[15] The situation is "strictly analogous" to a geometrical sphere in which "there is no way to individuate the points on the surface of a sphere purely in terms of their relations to one another" (2000: 208).

[16] A similar point holds for the less complicated quality spaces associated with the other sensory modalities.

relations that hold between the colours in virtue of structural relations that hold between our experiences of the colours: we judge that red is more like orange than blue in virtue of features of the *experiences* of red, orange and blue. In failing to require that any similarity amongst colour properties themselves explain the similarity judgements that we tend to make on the basis of visual experience, however, experience-introducing physicalist dispositionalism fails to meet the adequacy constraint in a satisfactory manner. Most people will agree that similarity relations hold amongst our colour experiences: that an experience of red is more like an experience of orange than an experience of blue. It is natural to suppose, however, that the similarity relations that hold between experiences are parasitic upon similarity relations that hold between the properties that these experiences are experiences of: experiences of red are similar to experiences of orange because *red* is similar to *orange*. The experience-introducing physicalist dispositionalist faces a dilemma. If the similarity relations that hold between colour experiences *are* underwritten by relations that hold between the colours, then an account of what these relations are is still owed. If, however, they are not, then physicalist dispositionalism will end up being deeply uncharitable to visual experience. Either way, the account is unsatisfactory.

Although the physicalist's attempts to reject premiss (2) of the Argument from Physics have been unsuccessful, this premiss may yet be rejected. Dispositionalism, in its more traditional form, represents a different way in which to reject it.

CHAPTER THREE: Two Theses about Dispositions

Central to any dispositional account is a biconditional of the general form:

(D) x is F iff were x to φ, then it would ψ.

Dispositional accounts differ depending on how (D) is interpreted, but fall essentially into two groups. On the one hand, there are those that interpret (D) as telling us something primarily about the *concept* F-ness (or, equivalently, assuming a straight forward link between language and thought, a semantic claim about the predicate 'is F'[1]). On the other, there are those that interpret (D) as a metaphysical claim about *properties*.

If it is to afford any response to the Argument from Physics, dispositionalism must be understood as a metaphysical thesis.[2] Though necessary, however, this condition is not sufficient: understood as a claim about properties, (D) also admits of two distinct interpretations. The weaker interpretation is that (D) is a truth merely *about* the property F: that, as a matter of fact, everything that instantiates this property ψ's when φ. The stronger interpretation regards (D) as an exhaustive constitution claim that tells us what F-ness actually *is*. Only when it is understood thus does dispositionalism have the potential for grounding a response to the eliminativist: identifying colours with dispositions that are logical constructions out of physically acceptable properties and the types of experience that the bearers of these properties are thereby disposed to produce, metaphysical dispositionalists, like physicalists, reject premiss (2) of the Argument from Physics. Colours, on this view, are essentially nothing over and above those properties describable using physical vocabulary: unlike the physicalist dispositionalist, the more traditional metaphysical dispositionalist does not require that the behaviour that coloured objects exhibit - and which is partly constitutive

[1] Henceforth qualifications of this kind will be suppressed. What might more perspicuously, but less elegantly, be labelled 'semantico-conceptual dispositionalism' will be sometimes be discussed in semantic terms and sometimes in conceptual terms depending on the context. The difference is not important.

[2] This applies equally to physicalist dispositionalism. Physicalist dispositionalism must, moreover, be understood as a strong metaphysical thesis. See, for example, Smith: "On the preferred account,

of colour properties – be itself describable using physical vocabulary. Metaphysical dispositionalism still represents a rejection of premiss (2), however, in virtue of the fact that the properties causally responsible for this behaviour are physical properties. In contrast to the stronger interpretation, the weaker interpretation entails nothing about what colours actually are: it is consistent with this view that colours are not in themselves physically describable properties, and therefore that the Argument from Physics is sound.

This weaker interpretation of (D) will receive fuller discussion in Chapter 5.1; here, it is metaphysical dispositionalism understood as a response to the Argument from Physics that is primarily of interest. To focus the discussion, however, it is important to first get a clearer grasp of dispositionalism understood as a *conceptual* thesis, in order that the truth or falsity of the conceptual claim may not be confused with the truth or falsity of the more fundamental metaphysical claim. Only given a good understanding of the conceptual thesis is it possible to adequately evaluate the metaphysical thesis.

3.1 Conceptual Dispositionalism:

Understood as a conceptual thesis, dispositionalism aims to draw a distinction between different types of concept based on peculiarities in our understanding of the dispositional biconditionals associated with each. The difference is typically identified in terms of *a priority*, such that, for example, (D) is knowable *a priori* for some concepts, whereas for others its truth can only be determined *a posteriori*: for example, that whilst the behaviour associated with objects that instantiate the property to which the concept 'square' refers can be determined only through experience, the behaviour characteristic of objects that instantiate the property to which the concept 'fragility' refers can be known *a priori*.[3]

As a thesis specifically about colour, conceptual dispositionalism is compelling: with respect to colour judgements it is common to assume that there

---

what blueness *is* (ontologically speaking) is a disposition normally to produce in us dominantly YB-negative responses" (1987: 249).

[3] The dispositional-categorical distinction is drawn in this way by Mumford: "Disposition ascriptions *are ascriptions of properties that occupy a particular functional role as a matter of conceptual necessity and have particular shape or structure characterizations only a posteriori.* Categorical ascriptions *are ascriptions of shapes and structures which have particular functional roles only a posteriori*" (1998: 77). See also Wright, who distinguishes between shape and colour judgements on the basis of whether they meet the "*A priority condition*" (1992: 117).

is no higher court of appeal than visual experience;[4] that there is no test we could perform that could overturn the judgment of a suitably placed, physiologically 'normal' observer in 'standard conditions'. In saying this, the conceptual dispositionalist owes an account of normalcy conditions for the manifestation of an object's characteristic behaviour. It has been objected (notably by Hardin 1993: 67-82) that this presents a problem, given that it is highly implausible to suppose that it will be possible to give any general account of normalcy conditions for the application of our colour concepts: amongst statistically normal observers (broadly construed), for example, there is significant variation in the location in the spectrum of the unique hues; similarly, whilst normal viewing conditions for most objects involves being viewed in north daylight, stars and bioluminescent fish, for example, can only be seen when there is *no* daylight at all.

The problem that this presents for the conceptual dispositionalist, however, should not be overstated: colour concepts are hardly unique in this respect. Ascriptions of squareness, for example, depend crucially on the situation for their applicability: if there are four quite differently shaped pegs in front of you and I ask you to pick the square one, you should have no problem in doing this even if the peg to which I am referring is only approximately square – if, for example, it has rounded corners, jagged edges etc. However, an engineer (say), will demand a much higher standard of precision if he asks for a square peg as part of his work. Similarly, whether you judge that a wine glass is fragile, for instance, will depend upon your reasons for making that judgement. If you want to know whether you can put the glass in the dishwasher, and the wine glass is made from thicker than average glass, then the judgement that the glass is *not* fragile will be to the point. If, however, you want to know whether the wine glass will withstand being dropped from a first floor window, then the correct judgement to make is that it *is* fragile.[5] It seems, therefore, that there is good reason to require that the

[4] In Wright's terminology (1992: 110), whilst shape judgements are "extension-reflecting", insofar as the extension of the truth predicate amongst judgements concerning the primary qualities is determined independently of the deliverances of best opinion, colour judgements are "extension-determining"; the extension of the truth predicate is, at least in part, determined by what the verdict of best opinion would be.

[5] As Travis remarks: "Something may be describable in a given way for some purposes, but not for others. There is such a thing as describability for a purpose" (1996: 461).

specification of the conditions under which *any* concept should be applied is best relativised to specific circumstances.[6]

To guarantee that our judgements involving dispositional concepts ever qualify as true, conceptual dispositionalism needs to presuppose certain metaphysical commitments. Commonly, these consist in the assumption that the properties to which our dispositional concepts refer *are* (metaphysically speaking) dispositions. Bennett, for example, reads into Locke's claim that secondary qualities are "nothing in the Objects themselves, but Powers to produce various Sensations in us" (Locke 1975: 2.8.10) two distinct theses roughly corresponding to the conceptual and the metaphysical theses identified above. The first - the Analytic Thesis – is that "any statement attributing a secondary quality to a thing is equivalent [in meaning] to a counterfactual conditional" like that given in (D) (Bennett 1971: 94). The second - the Causal Thesis - is that "in a perfected and completed science, all our secondary quality perceptions would be causally explained in terms of the primary qualities of the things we perceive" (1971: 102). Although, of these, Bennett's primary concern is with the Analytic Thesis - which he regards, in contrast to the Causal Thesis, as "a philosophical thesis whose support involves no recherché scientific information" (1971: 105) - Bennett nevertheless accepts the Causal Thesis: even in the absence of the relevant recherché scientific information, he suggests that once we accept the Analytic thesis it does not take very "strenuous argument" to arrive at the Causal Thesis.

The conceptual dispositionalist need not, however, be so committal as to accept the stronger metaphysical interpretation of (D). At a minimum, it is sufficient for our judgements involving dispositional concepts to qualify as true that the properties to which our concepts correspond indeed be such as to cause their bearers to exhibit certain sorts of behaviour in certain situations: the conceptual dispositionalist, that is, can allow for the truth of dispositional concept judgements by accepting merely the weaker metaphysical interpretation of (D). Johnston, for example, a dispositionalist with respect to colour concepts is (at least officially) neutral with respect to metaphysics, saying our choice of concepts

[6] There is also good reason to require that the description of these circumstances be rigidified to the actual world, as it allows the dispositionalist to respect the intuition that even in a world in which there are no sentient beings, lemons are still yellow. It should be noted, however, that not all dispositionalists share this intuition: see for example, McGinn (1983) and Jackson and Pargetter (1987).

is "not dictated by the natures of the things under discussion. For those natures admit of many types of true descriptions" (1992: 168).[7] Being more forthright, the conceptual dispositionalist may explicitly *deny* that colours are, metaphysically speaking, dispositions: even though our colour *concepts* are determined by the types of experience that the objects to which they apply produce, the *properties* to which these concepts refer need not themselves be exhaustively constituted by this behaviour. Rather, it might be claimed that colours are relatively abiding properties that inhere in objects independent of the possibility of being perceived and ground those objects' dispositions to produce in suitably placed, suitably receptive observers, certain kinds of experience.[8]

As well as disagreeing over the metaphysical status of the properties to which dispositional concepts refer, conceptual dispositionalists also differ as to what they take (D) to tell us about our concepts or terms. We have already seen that Bennett, for example, claims that (D) serves to reductively define the predicate 'F' (at least where 'F' is a secondary quality predicate): that "any statement attributing a secondary quality to a thing is equivalent [in meaning] to a counterfactual conditional" like that given in (D). This proposal, however, faces a grave difficulty. The purpose of reductive definition is to define a word (or equivalently: analyse a concept) in terms of others, in such a way that would enable someone who did not understand that term to acquire it, given knowledge of the other words used in the *definiens*, on the basis of that definition. A paradigmatic example of this strategy is the definition of 'bachelor' in terms of 'unmarried man'. The problem for Bennett lies in explicitly formulating (D) in such a way as to avoid using any terms in the *definiens*, the meaning of which can only be explicated with reference to the meaning of the *definiendum*.

One way in which it might be thought that a difficulty in this respect arises is in response to a quite general problem that affects any dispositional analysis, either conceptual or metaphysical: finkishness (Martin 1994). Finkishly disposed objects cease to be disposed to produce their characteristic response in exactly those conditions under which they would otherwise manifest that very disposition: a glass that would cease to be fragile when, and *because*, it is

[7] Johnston's views may subsequently have changed. Cuneo (2001), for example, adduces evidence that Johnston accepts dispositionalism in both its conceptual and metaphysical versions.
[8] This view will be defended in Chapter 5.

dropped, for example, is finkishly fragile; similarly, a shy green chameleon, which blushes bright red when it intuits that it is about to be observed is finkishly green (Johnston 1992: 145).[9] Finkishness is a problem for the semantic dispositionalist, according to whom there is an important semantic relation (according to Bennett: analytic equivalence) between 'x is F' and 'x would manifest certain behaviour in certain circumstances', because if x is finkishly F then either: (a) 'x is F' is false, which violates our intuition that something finkishly F is still F; or (b) the semantic relation posited between 'x is F' and 'x would manifest certain behaviour in certain circumstances' does not hold. Finkishness similarly poses a problem for the metaphysical dispositionalist, according to whom being F is nothing over and above exhibiting certain behaviour in certain circumstances, because if x is finkishly F then either: (a) finkishly F things are not F (they never manifest the appropriate behaviour); or (b) being F cannot consist in exhibiting the behaviour in question.

To avoid problems of this kind, Lewis (1997) and following him Mellor (2000), amend the general form of (D) to,

(D') x is F iff were x to φ without ceasing to be F, then it would ψ.

Calling these conditionals "reduction sentences", Mellor notes that whilst they nevertheless avoid the problem of finkishness, they have seemingly "unacceptable semantic side-effects, since to understand [e.g.] "if x were stressed without ceasing to be fragile..." we must already know what 'fragile' means". He proceeds to claim that "This however is no real objection to them, since it does not in fact stop us using them to remedy the ignorance of those who do not know what to call 'fragile' by saying that, by definition, all and only those things that remain or become fragile when (relatively suddenly and lightly) stressed will then break" (2000: 763). For someone with Bennett's aspirations, however, such side-effects *would* be fatal: (D') could not be used to reductively define 'is F' because exactly the same undefined predicate appears in both the *definiens* and the *definiendum*. For Bennett's account to even get off the ground, some antidote to

[9] Kripke's killer yellow things, whilst not themselves finkishly disposed, require the instantiation of a finkish disposition in the observing subject: the subject fails to manifest his disposition to undergo a certain sort of experience when presented with a killer yellow object precisely *because* of the particular shade of yellow he is presented with. Finkishly disposed things fail to display their characteristic behaviour because of a change in *them* brought about by a change in their environmental circumstances.

finkishness that does not require circularity in the dispositional biconditional is required.

One response to finkishness which would meet this constraint is Blackburn's. To avoid the problems posed by finkishness, Blackburn inserts a 'is naturally such that' – i.e. "in accordance with nature and its laws" - clause into the dispositional biconditional (1993: Section 2). On this account, x is fragile iff it is naturally such as to break when dropped, even if, as it happens, this natural disposition never manifests itself. Whilst this manoeuvre avoids circularity, however, there remains the question of exactly what is meant by "in accordance with nature and its laws". In particular, if *no* glass were ever such as to break when dropped because all glasses were finkishly fragile, then it would not seem obviously unreasonable to count as a law of nature 'dropped glasses do *not* break': at least according to the Humean, for example, this law would display all the necessity required of a law of nature. And if this were the case, then 'x is fragile iff it is naturally such as to break when dropped' would be false; finkishly fragile glasses would not be fragile after all.

A more straightforward response treats the occurrence of 'F' on the right-hand side of the biconditional as an instance of anaphora and paraphrases (D') accordingly: instead of requiring that the very term 'F' reappear on the right-hand side, thus rendering the biconditional circular, (D') is recast as,

> (D'') x is F iff were x to φ without ceasing to be such as to ψ when φ, then it would ψ.

Stated thus, there is no necessity that the dispositional biconditional contain any terms on the right-hand side the meaning of which is a function of the meaning of the terms on the left-hand side.

Bennett is not yet in the clear, however: whilst there may be no necessity in general that the right-hand side of (D'') contain any terms whose meaning is a function of any of the terms on the left-hand side, in the specific case of colour it is often argued that such circularity is unavoidable. That is, that there is no possible positive characterisation of the sort of experience that coloured objects are disposed to produce that does not use a term whose meaning is a function of the meaning of the *definiendum* term: that 'x is yellow iff x is disposed to look

yellow in normal circumstances' (or some variation thereon) is the best we can manage.[10]

For someone like Bennett with aspiritations of reductive definition, there are only two possible replies to this objection once it is conceded that no characterisation of the relevant experiences in non-colour vocabulary (which at the same time respects the *a priori* status of the dispositional conditional) is possible. Either, it can be claimed that 'looks yellow' as it appears on the right-hand side is semantically unstructured and so does not admit of semantic decomposition into its constituent parts (like, for example 'red herring', the meaning of which is a function of neither 'herring' nor 'red'). Or, it can be claimed that the semantic structure of 'looks yellow' is non-standard, and so its meaning is not problematically a function of its constituent parts: Peacocke (1983), for example, develops this sort of approach by arguing that 'looks yellow' should be understood, not as saying that yellow objects give the appearance of instantiating the property yellow, but rather as saying that they cause the instantiation of an intrinsic property of the visual field, yellow'; that 'x is yellow iff x is disposed to be represented in yellow' patches of the visual field in normal circumstances'.

Of these, the first approach is particularly unappealing: 'looks yellow' does not seem at all like 'red herring'. The second type of strategy is more promising - at least as it is developed by Peacocke, however, it presupposes the existence of intrinsic sensational properties (whether any exist is highly controversial; see, for example, Harman 1990). Either way, the more usual response for the semantic dispositionalist to make in face of the circularity objection is to renounce any intention of reductive definition of the predicate 'is F' in the first place. So, for instance, McDowell remarks, "it is quite unclear that we ought to have the sort of analytic or definitional aspirations that would make the circle problematic" (1985: 124, n6), referring the reader to McGinn (1983) for a more detailed defence. This reference to McGinn is surprising as, at least in the passage McDowell refers to, McGinn is proposing to avoid the circularity objection by understanding dispositionalism as a metaphysical thesis; nevertheless, the remark is suggestive.

[10] See, for example, Evans: "It seems decisive against any dispositional account of the meaning of such a term as 'red' that the only way to characterize the experience that red objects produce in us is as such" (1980: 272, n27).

At the very least, the dispositionalist can claim to be offering, if not definition, then commentary or conceptual elucidation. This, for example, is Wiggins's approach, who sees the dispositional account as bringing out the subjectivity implicit in our colour terms (1987: 189).

Circumscribing the circularity objection is not the semantic dispositionalist's only worry. Transposing from the semantic into the conceptual, for example, and assuming, in Stroud's terminology, a "direct connection" between colour properties as they feature in thought and colour properties as they are presented in visual experience, conceptual dispositionalism is a theory about both the cognitive *and* the perceptual representation of colour. As such, it is open to two kinds of objection. The first is phenomenological: colours do not look like dispositions, but rather intrinsic properties of objects.[11] The second consists in the challenge of specifying the contents of perceptions of dispositional properties without generating a vicious regress: in specifying the content of a perception of yellow we will have to use the term 'yellow', which itself will have to be unpacked in terms of 'a disposition to look yellow in normal circumstances' and so on.[12]

Whether either objection is decisive is questionable: the second objection looks to be stronger than the first, given that whether colours can look like dispositions will depend at least in part upon extraneous commitments in the philosophy of perception;[13] in response to either objection, however, it is open to the conceptual dispositionalist to reject the underlying assumption that there is a direct connection between colour properties as they feature in thought and colour properties as they are presented in experience. To discuss these issues in any more detail would take me too far afield: it is at least not obviously unreasonable, however, to suppose that in some form or other, conceptual dispositionalism about colour is a viable thesis.

For our purposes, conceptual dispositionalism is important primarily only insofar as it more clearly defines the logical space of the metaphysical

---

[11] See, for example, Boghossian and Velleman (1989: 85), Johnston (1992: 141), McGinn (1996: 300).

[12] See, for example, Boghossian and Velleman (1989: 86), McGinn (1996: 303) and Stroud (2000: 141).

[13] If, for example, perception is assumed to be essentially sensational, then the phenomenological objection may well be decisive. If, however, an intentional theory of perception which allows for the possibility of seeing *that* is presupposed, the issue is far less clear cut. These questions are discussed briefly by Stroud (2000: 140-1).

dispositionalist thesis. It is this thesis that has the potential for grounding a response to the Argument from Physics.

3.2 Metaphysical Dispositionalism:

According to the metaphysical dispositionalist, (D) is a constitution claim to the effect that being disposed to exhibit characteristic behaviour in certain situations is *exhaustively constitutive* of a given property: F-ness is a logical construction out of an F-object's physical properties and the behaviour that it is thereby disposed to exhibit and consists in nothing over and above - there is no abiding property F-ness that inheres in the object independent of the possibility of its manifesting its distinctive behaviour. Typically, metaphysical dispositionalism is wed to the assumption that some properties differ, in this respect, from others; that whilst there is nothing to objects bearing certain properties over and above their disposition to exhibit certain kinds of behaviour, this is not universally so - consequently, that (D), understood as an exhaustive constitution claim is true for some properties but false for others (this is in contrast to conceptual dispositionalism according to which the difference between the dispositional and the categorical lies in the manner in which (D) is known to be true, and not its truth or falsity).[14] The assumption of property dualism is not essential to the metaphysical dispositionalist's position, however: it is consistent, for example, to hold that no particular subset of properties is unique in being exhausted by their casual dispositions, and indeed the same is true of *all* properties; that properties are, in Blackburn's term dispositional "all the way down" (1990: 63).[15]

Whilst, as a conceptual thesis, dispositionalism is commonly associated with certain metaphysical commitments, there is much greater scope for developing dispositionalism as a metaphysical thesis independently of conceptual considerations. Mackie (1976), for instance, believes that whilst, metaphysically speaking, colours *are* dispositions to produce certain kinds of experience in certain circumstances, this is not how we experience them. Rather, we experience

[14] See, for example, Bennett (1971: 104) and Evans (1980: 276).

[15] See, for example, Mellor (1974) and Shoemaker (1980). It is also consistent, in a very loose sense, to identify dispositions with their categorical bases (see for example Armstrong 1968: 85-9 and Armstrong *et al* 1996); I say 'in a very loose sense' because in this eventuality, talk of dispositions begins to lose its point. This combination of views is probably best not labelled 'dispositionalism' at all.

colours as intrinsic, non-relational properties of objects. A more extreme position is also possible. According to Mackie, although visual experience is systematically misrepresentative, it is only misrepresentative with respect to nature of colour properties and not with respect to which properties our experience tracks. The more extreme position is that whilst visual experience tracks properties of a certain sort – dispositions to produce in us certain kinds of visual impression – these are not the properties that we commonly take ourselves to track: it is not just that we conceptualise the colour properties that we track erroneously; we do not even track the properties that we think we do. This position is much closer to Mackie's account of moral properties, and also bears certain affinities to Boghossian and Velleman's eliminativism: to allow that there is an interesting sense in which our colour-talk is still truth evaluable (1989: 98-101), Boghossian and Velleman allow that our experience tracks worldly properties, even though these properties are not the properties we ordinarily assume ourselves to be tracking (indeed, the properties we ordinarily assume ourselves to be tracking do not exist).

Mackie's error theory - according to which colours, though dispositions, are not experienced as such – is presented as a response to the phenomenological objection to concept dispositionalism. Hence the debate with McDowell: both accept that colours are dispositional properties, but disagree over how best to describe the phenomenology. In formulating dispositionalism as an explicitly metaphysical thesis to avoid the problems associated with conceptual dispositionalism, Mackie is not alone. McGinn, this time impressed by the circularity objection, for example, also believes that dispositionalism is best understood as a metaphysical thesis. Thus, he remarks, "I will allow myself to speak of *analysing* secondary qualities in terms of dispositions to produce sensory experiences, understanding by this the thesis that these experiential facts are *constitutive* of the presence of the quality in question" (1983: 8).[16] In a similar vein, Evans regards (D) as grounding the truth of ascriptions of dispositional properties, saying, "I differ from Bennett in not making the dispositional character

[16] Although McGinn does not explicitly claim that the experiences coloured objects are disposed to produce are *exhaustively* constitutive of those objects' colour (i.e. that there is no 'abiding residue'), this is certainly the implication. He is more explicit subsequently, saying that, "Colour properties, we might say, are logical constructions from categorical physical realizations and higher-order causal relations to perceptual experiences" (1996: 299).

of the secondary qualities a matter of the meaning of sentences ascribing secondary qualities, but relying instead upon the obscurer notion of *that in which their truth consists.*" (1980: 272, n27, emphasis added).

Understood as a property analysis, dispositionalism circumscribes worries about circularity. On this reading, colour properties are identified with the set of experiences that their bearers would produce in a given situation. How these experiences are referred to, however, makes no difference to the analysis *per se.* Clearly, there will be differences in terms of utility and comprehensibility between the different descriptions of the relevant experiences: for instance, it is not much use describing these experiences in phenomenal terms to someone who does not share the same phenomenology as yourself. But these differences are cosmetic. As long as the same set of experiences is identified by each description, the analyses are mere notational variants.[17]

Whether or not the desire to avoid these objections is sufficient motivation for understanding dispositionalism as a metaphysical thesis, however, it is not the only possible motivation. Perhaps the most persuasive reason for interpreting dispositionalism as a metaphysical thesis is its potential for blocking the eliminativist's conclusion. According to the metaphysical dispositionalist, colours are properties that are exhaustively constituted by their bearer's dispositions to produce certain kinds of experience, grounded by that object's physically acceptable properties. Metaphysical dispositionalism thus represents one way of denying premiss (2) of the Argument from Physics: in virtue of the fact that colours are nothing over and above those physical properties that dispose their bearers to produce certain kinds of experience in certain situations, colours *do* have a place within a physical description of the world. Resisting the *ontological* distinction between colours and physically acceptable properties presupposed by the eliminativist, the metaphysical dispositionalism puts in its place a merely *metaphysical* distinction: both colour and shape, exist, they are just of metaphysically different kinds – shape is more fundamental than colour.[18]

---

[17] Consequently, it is not an advantage of physicalist dispositionalism that it avoids worries concerning circularity. Circularity worries affect only dispositionalism understood as a conceptual thesis.

[18] As McGinn remarks, metaphysical dispositionalism about colour is "Logically...analogous to functionalist theories of mental properties" (1996: 299), which are themselves often seen as attempts to bring a supposedly problematic feature of the world within the compass of the physical.

To sustain this response to the eliminativist, the dispositionalist owes a more detailed account of the nature of the metaphysical distinction between shape and colour. One such account is given by Evans. This account of the nature of the dispositional-categorical distinction is based upon a distinction between primary and secondary quality concepts. According to Evans, to grasp primary quality concepts requires mastering a set of "interconnected principles": in the case of concepts for properties such as electric charge, these principles are explicit, whilst for concepts for "the properties constitutive of the idea of material substance as *space-occupying stuff*", i.e. shape, size etc., such principles are implicit (1980: 269). Primary quality concepts differ in this respect from secondary quality concepts, because unlike primary quality concepts, secondary quality concepts may be 'distilled directly out of experience'.

On the basis of this distinction between primary and secondary quality *concepts*, Evans proceeds to argue that there is a corresponding difference between the *properties* to which these concepts refer: namely, that whilst squareness is a relatively abiding property of objects, the instantiation of which explains an object's causal interactions, there are no abiding colour properties - "All it can amount to for something to be red is that it be such that, if looked at in the normal conditions, it will appear red" (1980: 272). Anyone who rejects this sort of metaphysical dispositionalism about colour in favour of a view that presupposes the intelligibility of conceiving of colours as abiding properties of objects that resemble our experiences of them, is challenged by Evans to explain "in what the difference between such an objective colour property, and the dispositional property, consists" (Evans 1980: 273; Bolton 1983: 360-4 attributes a similar argument to Locke).

The problem with this account of the dispositional-categorical distinction, however, is that the metaphysical claim does not follow from the conceptual claim.[19] Even conceding that there is no difference between dispositional and non-dispositional colour *concepts*, nothing follows about the properties to which these concepts refer: just because we find it difficult to conceive of a colour property that resembles our experience of it characterising an object in the dark, for instance, there is no reason to suppose that this tells us anything about the way the

world is as opposed to merely telling us something about our own conceptual capacities. Perhaps in this respect, like in so many others, conceivability is just not a reliable guide to possibility. To answer Evans's challenge directly, perhaps the difference between objective and dispositional colour properties is simply that the former, but not the latter, are relatively abiding properties of objects. No conclusive reason why this cannot be so has yet been given. An account of the dispositional-categorical distinction that is of more obvious metaphysical significance is required.

According to Prior, "The commonly accepted view" of the nature of this distinction, "is that dispositions differ from categorical properties because the former possess a special relationship to subjunctive conditionals not possessed by the latter. More specifically, one common view is that sentences which ascribe dispositions entail certain subjunctive conditionals while those which ascribe categorical properties do not" (1985: 59). Granting that the putative relationship sentences ascribing dispositional properties bear to subjunctive (counterfactual) conditionals *is* entailment, to sustain this version of the distinction it will have to be shown that sentences ascribing dispositional properties are unique in this respect.[20]

Prior's statement of the Common View requires disambiguation. One way of interpreting it is as the claim that whilst ascriptions of dispositional properties entail subjunctive conditionals, ascriptions of categorical properties do not, *tout*

---

[19] In keeping with this, as it appears in McDowell (1985: 113), this argument is used to support dispositionalism as a merely conceptual thesis.

[20] The assumption that sentences ascribing dispositional properties entail subjunctive conditionals is controversial. Defined in classical terms, entailment is a relation that holds between propositions, such that x entails y just in case there are no circumstances in which x is true and y is false. If, however, as, for example, Adams (1975) and Edgington (1986) believe, subjunctive conditionals have no truth-conditions, then there can be no entailment where one of the relata is a subjunctive conditional. Even if the conclusion is granted, however, it is not disastrous if it is taken to show only an inadequacy in the classical notion of entailment. Generalising the classical notion in terms of assertability conditions, we might instead claim, for example, that x 'entails' y only if we are prepared to accept y on the basis of accepting x. This would then yield a degree theory account of conditionals, according to which you should accept 'If A, B' if the credence that you give to A&B is nearly that which you give to A itself (or equivalently, that A&B is more likely than A&~B). At least *some* notion of entailment may therefore be assumed to hold between sentences ascribing dispositional properties and their corresponding conditionals, even if this relation is more properly characterised in terms of assertability conditions than classical entailment: that, for example, if you accept 'x is fragile' then you should be prepared to accept the corresponding conditional, 'if x were dropped, then it would break'.

court.[21] Understood thus, however, the Common View is surely false: on this reading 'x is fragile', for example, is supposed to differ from 'x is square' in bearing a special relationship to subjunctive conditionals of the form 'were x to φ, then it would ψ' – in the case of fragility to the conditional 'were x to be dropped, then it would break'.[22] But the truth of sentences ascribing squareness to a (rigid) object also seems to require the truth of corresponding conditionals: for example, 'were x to be placed at the top of shallow incline, then it would not slide down' and 'were x to be brought up to a round hole of equal area, then it would not fit through'.

A more promising interpretation of the Common View is that the difference between ascriptions of squareness and ascriptions of fragility lies in the *kinds* of subjunctive conditional that dispositional and categorical property ascriptions entail. So, it might be suggested, whilst fragility bears an intimate relation to counterfactual conditionals of just one general form – 'were x to be dropped, then it would break' – there is not just one type of conditional that ascriptions of squareness bear this relation to; the idea being that properties that only manifest themselves in one way are 'less real' than properties that can act on other objects in more ways than one.

In contrast to the dispositional-categorical distinction drawn by Evans, which is based on purely conceptual considerations,[23] this distinction is empirically grounded, turning on a difference in causal role between dispositional and categorical properties. As such, it is more explicitly metaphysical. Unfortunately, it does not yield the kind of distinction that the metaphysical dispositionalist originally envisaged.

For one thing, ascriptions of fragility are not as simple as this suggestion implies. Fragility is a property that objects can manifest without being dropped. The pane of glass that cracks when it is tapped, for example, is just as fragile as

[21] Quine (1960, 1969) seems to endorse this view, suggesting that the scientific discovery of micro-structural properties renders respectable otherwise disreputable dispositional properties. This is noted by Mellor (1974: 115).

[22] In discussing these issues, it is helpful to use as an example of a paradigmatic dispositional property fragility, rather than colour: if there *is* a significant distinction between dispositional and categorical properties, colours may not fall squarely on the dispositional side (see, for example, Jackson 1996). In this eventuality, citing colour as a paradigm example of a dispositional property would obscure what genuine distinction there is.

[23] Reid, from whom Evans takes his version of the distinction, for example, remarks, "The account I have given of this distinction is founded upon no hypothesis" (1983: 183).

the glass that breaks when it is dropped. If the difference between the dispositional and the categorical is that ascriptions of the former, but not the latter, bear an intimate relation to subjunctive conditionals of just one type, then it looks as though fragility will not be a dispositional property. But if fragility is a not dispositional property, then this cannot be the distinction that we intended to capture, as it will be agreed on all sides that if any property is dispositional then fragility is.

In response, it might be suggested that all this objection shows is that our current formulation of the counterfactual conditional associated with fragility is inadequate, and not that fragility is not a dispositional property: perhaps breaking when dropped and breaking when tapped are not really different manifestations of the same property at all. Fragility is a property that causes objects to crack, break *or* chip when they interact suitably with other objects. And intuitively, at least, there is a unity to these types of behavioural manifestation.

The problem for the dispositionalist, however, lies in spelling this out more explicitly. Particularly problematic will be cashing out 'suitable interaction': if a fridge 'suitably interacts' with the floor when dropped from a first floor window it will break into smaller pieces; yet we do not really want to say that fridges are fragile. There are even certain 'suitable interactions' that will leave diamonds cracked or chipped. Diamonds, however, are quite clearly not fragile. What is needed is some account of 'suitable interaction' that excludes the cases we want excluding but doesn't exclude those that we don't want excluding. This, however, will not be an easy task.[24]

But even if we grant that some account of 'suitable interaction' can be given, the current proposal still does not capture the dispositional-categorical distinction that it was intended to. The fact that dispositional properties only manifest themselves in one type of way may distinguish them, in some sense, from properties that enjoy a 'wider cosmological role', but it is far from clear that the distinction is *substantially* metaphysical and that those properties that enter into a more limited range of causal interactions are therefore any less real. All that would appear to follow is that shape is just a more complicated disposition than fragility, and so something along the lines of,

(D''') x is square iff x is disposed to enter into the causal interactions characteristic of square things,

is true - in Ryle's terminology, that primary qualities are "generic dispositions", whilst secondary qualities are "specific dispositions" (1949: 118).[25]

A more substantially metaphysical distinction between dispositions and their grounds can be drawn by arguing, as in fact many dispositionalists do, that dispositional properties are causally inefficacious *tout court*.[26] Though more obviously to the point, however, this approach, too, faces problems. Intuitively, at least, dropped wine glasses break when dropped *because* they are fragile; instantiating the property fragility causes them to exhibit behaviour of this kind. Similarly colours, it is ordinarily assumed, are causally efficacious. Following Broackes (1992: 194), there are three ways in which it might be considered that this is not the case: (i) colours do not explain the effects of material objects on sentient beings, (ii) colours do not explain the effects of material objects on the colour properties of other objects and (iii) colours do not explain the effects of material objects on the non-colour properties of other objects. There is good reason to suppose, however, that colours are causally efficacious in *at least* two of these ways.

As a counter-example to (iii), for example, Hacker (1987: 139) presents the case where one hut is cool and another hot because the former is painted white and the latter black: black objects absorb more infra-red light than white objects and so heat up faster. Broackes objects (1992: 194) that this does not constitute a genuine counter-example to (iii): the failure of black objects to reflect infra-red light is independent of their failure to also reflect visible light; not only can objects that reflect visible light fail to reflect infra-red light, but objects that fail to reflect visible light can nevertheless reflect infra-red light.

On Hacker's behalf (though it would no doubt be help that he would not be keen to accept), however, we might respond to Broackes as follows. The colour concepts with which we operate are a function of our physiological makeup; they

---

[24] Prior, who accepts a metaphysical distinction between the dispositional and the categorical, concedes as much (1985: 61).

[25] Worse still, if colours enter into more complicated causal interactions than is often assumed, then according to this dispositional-categorical distinction colours will not be dispositions, and so the dispositionalist response to the Argument from Physics would fail anyway.

[26] See, for example, Mackie (1976: 18), McGinn (1983: 14), McDowell (1985: 118) and Nagel (1986: 75).

are anthropocentric. So, for example, honeybees are capable of discriminating hues at the ultra-violet end of the spectrum that are entirely alien to human observers, whilst humans have concepts for hues at the other end of the spectrum that are entirely alien to bees. Now, imagine a species whose 'visible window' extended into the infra-red, i.e. who could discriminate spectral hues in the infra-red region of the spectrum.[27] This species may well count those objects that reflect infra-red light as belonging to one colour category (call it 'Infra-Red') and those that reflect visible light as belonging to another. Given the anthropocentricity of colour concepts, we would have no reason to privilege our set of concepts over any others; but now, it seems the following explanation is possible: the first hut is cool and the second hot because the first hut is painted Infra-Red and the second hut is not painted Infra-Red. If this is correct, then we can even explain the effects of material objects on the non-colour properties of other objects.

The dispositionalist may still not accept this argument: perhaps Infra-Red is not a *colour* but just a light-related property similar to colour - in the same way that if all and only knives were yellow, being yellow would not causally explain a knife's power to cut bread, maybe the causal basis for an object's non-colour-related effects on other objects differs from the causal basis for its experience inducing effects on sentient beings; two properties can, after all, be always co-instantiated without being identical. This response may not ultimately be tenable (there may be a suspicion, for example, that it is *ad hoc*). Even if it is, however, it does not help the dispositionalist who is attempting to argue that colours are causally inefficacious *tout court.*

Because two objects can look to be the same colour even though their spectral reflectance profiles differ and two lights can look to be the same colour even though their spectral compositions differ, the colour that an object appears to have is not determined by the colour of the light and the colour of the object. Rather, it is determined by the object's spectral reflectance profile and the light's spectral composition: to use Broackes' example, the fact that a tomato is placed in a green light is neither necessary nor sufficient for (and hence cannot explain) the shade it will look to be. Nevertheless, the following sort of causal explanation is still possible: the mug looks blue to John because it is blue and John is looking at it in

[27] It is believed that some fresh water fish (e.g. salmon), can see into the near infra-red (over

decent lighting and has good colour vision. Here it is irrelevant that blueness can be realised by different spectral reflectance profiles; that it is realised by *this* reflectance profile is all that matters. Colours, therefore, are not causally inefficacious in the first of the ways identified by Broackes. Less controversially still, colours are not explanatorily idle when it comes to explaining the effects of material objects on the colour properties of other objects: Grassman's Third Law, which explains the colour of a mixture of lights in terms of the colour of the combined lights, Chevreul's laws of colour contrast and Rood's laws which explain the additive mixing of white with coloured light, all provide examples (Broackes 1992: 194-202).

The dispositionalist will probably object at this point that 'high-level' explanations of this sort in terms of dispositional properties are pre-empted by 'lower-level' explanations that do not mention these properties at all. In Lewis's weak sense of explanation discussed above, this may be correct; as we saw there, however, this notion is not philosophically interesting. What we are interested in is *good* explanation. And in this respect low-level explanations of colour perception fail.

The situation with respect to colour perception is analogous to that of object interaction discussed by Putnam. Explanations of why rigid square pegs do not pass through round holes cut into rigid boards, for example, that refer to the fundamental constituents of pegs and boards are unilluminating because they fail to bring out the relevant features of situations of this sort. Similarly unilluminating are explanations of colour perception in terms of fundamental physics. What is needed are explanations that "will go in any world (whatever the microstructure) in which those *higher level* structural features are present" (Putnam 1975: 296); explanations, that is, that will ground generalisation. Granting that there is an intimate relation between explanation and causation – that, at least for the most part, explanation is *causal* explanation (see, for example, Lewis 1986) – there is no reason to deny that colours are causally efficacious.[28]

740nm and below 1000nm). See Thompson (1995: 146).

[28] Explanations in terms of colour properties may, for example, bear the relation to explanations in terms of fundamental physics sketched by Campbell: explanations in terms of colour properties are explanations because they "add modal data to a description of the physical sequence" initiated by an object's micro-physical structure that terminates in (the physical event of) an experience of redness: they say that "in nearby worlds in which the physical character of the thing was varied but its redness maintained, an experience of redness was still the upshot" (1993: 184).

But if this is correct, then the dispositionalist's attempt to draw a substantial metaphysical distinction between colours and shapes on the basis of causal efficacy, and with it his response to the Argument from Physics, fails.

That this *is* correct is borne out by Jackson's causal argument against the metaphysical dispositionalist: if "dispositions are not causes", then colours, identified with dispositions, will be causally inefficacious in the visual process. This leaves the dispositionalist unable to account for the "prime intuition" about colour, namely that, for example, yellowness is the property of objects putatively presented to subjects when those objects look yellow (1996, 2000). Jackson takes this to show that colours must therefore be primary qualities of objects. We have already seen that this cannot be correct, however.

In light of the failure of both physicalism and dispositionalism to satisfactorily afford a response to premiss (2) of the Argument from Physics, we should expect that if anything is wrong with the eliminativist's argument, it will be premiss (1). To see whether this is in fact so, a more detailed investigation into the import of this claim is required.

CHAPTER FOUR: The Physical World and the Objective World

According to premiss (1) of the Argument from Physics, a physical description of the world exhausts what exists objectively – everything else is mere appearance. This can be paraphrased as the claim that 'the objective world *is* the physical world', which in turn admits of two interpretations. Either, it can be read as a stipulative definition, where the objective world is *defined* as the physical world. Or, it can be read as a substantive identity claim to the effect that the objective and the physical worlds are one and the same. The advantage of the definitional reading is that it only requires us to give content to the notion of 'the physical world'; the content of this notion can then be used to fix that of 'the objective world'. However, for the purposes of the Argument from Physics, this reading renders (1) equivalent to the pleonasm,

> (1') A purely physical description of the world exhausts all that exists physically,

which entails nothing about the significance of being describable in purely physical vocabulary. Instead, we need to interpret 'the objective world *is* the physical world' as a substantive identity claim. And in order to assess this claim, it is necessary to have a clear understanding of *both* the statement's relata.

4.1 The Physical World:

The most natural way of giving content to the notion of 'the physical world' is in terms of the description of the world that we are given by the purely physical sciences. Roughly, on this view, all that reality consists in is the interaction, in accordance with the laws of physics, of 'physical entities'.

To say that the physical world is the world populated by physical entities, however, is so far unilluminating. Specifically, it is necessary to give content to 'physical' in 'physical entity' without creating a definitional circle so small that a physical entity is defined as something that is described by the purely physical sciences. One strategy for breaking this circle is to ostensively identify the extension of 'physical entity': physical entities are those entities that are

mentioned by the physical sciences. So, for example, physical entities include the microscopic electrons of particle physics and the macroscopic galaxies of cosmology. There is, however, a fundamental problem with this approach. The problem is that the range of entities mentioned by the physical sciences is decidedly limited; as Stroud says:

> If being expressed in the language of physics were a condition of a statement's being part of a description of the independent [objective] world, it would not be part of the independent world that there are such things as mountains on earth, or fish in the sea. 'Mountain', 'earth', 'fish', and 'sea' are not terms of physics. (2000: 54)

So unless we consider not physics itself, but some general extension of it, it looks as though the Argument from Physics, if sound, will prove far too much. Not only will it show that colours do not exist objectively, but by parity of reasoning that tables, books and people don't exist objectively, either. And this conclusion is not the eliminativist's: eliminativism is the doctrine that the *material objects* that populate the physical world are not in fact coloured, not that there exist no material objects either.

A solution to this problem may, however, be extracted from ideas present in the discussion immediately preceding the passage quoted from Stroud. There, Stroud attempts to define 'the physical world' in terms of 'physical properties' – those properties that feature in physical descriptions, laws and explanations (2000: 49). He rejects this definition on the grounds that explicating 'physical description, law and explanation' requires a notion of physical fact, in which the adjective 'physical' is still undefined. However, his eventual solution of ostensive identification could just as easily be used at this point: physical properties are those properties mentioned in *this* set of descriptions, laws and explanations. If we now define 'physical entities' as the bearers of physical properties such as mass, velocity, spin and so on, we arrive at an account of the world that encompasses tables, galaxies, electrons, quarks etc., thereby restricting the scope of the Argument from Physics. But one major question still remains: what is meant by 'the physical sciences'?

First of all, there is a question of scope. Taking as our starting point physics, how many of the other branches of science, if any, is 'physical science' intended to cover? The answer to this question will depend upon your views on theory

reduction, specifically whether there is any sense in which all scientific disciplines are reducible to physics, and if so, if it is a strong enough sense to give substance to the idea of 'the unity of science'. The prospects for this project, however, look bleak; the prudent approach, when we speak of 'physical science', will therefore be to restrict ourselves to physics.

Even so, the question of what is meant by 'physics' still remains. Let us assume that we can give a list cataloguing all the disciplines that are currently accepted branches of physics (e.g. mechanics, quantum mechanics, cosmology, acoustics etc). Can we use this list to fix the meaning of 'physics'? This suggestion is problematic on two counts. First, some or all of the branches in this list may contain a significant number of falsehoods. This worry need not depend on the validity of the pessimistic meta-induction for the falsity of all current scientific theories, according to which the falsity of all previous scientific theories licences the inductive inference that all present and future scientific theories will also be false, but might simply be based on modesty with respect to our cognitive powers. Even those philosophers of science who accept "the miracle argument" for scientific realism - the abductive inference to the truth of our scientific theories from their success - are usually reluctant to draw any stronger conclusion than that current physical theory is *approximately* true; that is to say it might, and probably does, contain some falsehoods.[1] Second, even granting that our current physical theory is entirely correct, we could still doubt its completeness if we thought that it could be extended to describe phenomena that presently lie beyond its scope (including phenomena that we do not even currently realise exist); and this surely seems a possibility.

Both these problems suggest that we cannot straightforwardly identify the physical world with the description of the world that current physical theory gives us; rather, we need to identify the physical world as 'that which is described by the entirely correct and complete physical theory, whatever that may be'.[2]

---

[1] Understanding a theory as approximately true if it contains more truths than falsehoods is a gross over-simplification, but one that will suffice for ease of exposition.

[2] Melnyck (1997) believes that if we base our physicalist principles on current physical theory, like that theory, our principles will very likely be false. Nevertheless, he argues, this should not stop us from believing them, as they will be at least as good as their relevant rivals. Accepting this, however, does not help the eliminativist. It may be reasonable, and even rational, to believe something that is unlikely to be true; but in admitting that our current physics is probably false we directly undermine premiss (2). Poland's response to this problem is similarly unhelpful. Poland rejects the *a posteriori* approach to identifying the physical discussed here, insisting that *a priori*

In respect of the Argument from Physics, however, this immediately raises problems. If we do not yet know what properties (and with them entities) the entirely correct and complete physical theory will quantify over, no direct comparison between the physical world with the objective world will be possible. This threatens to leave the truth of premiss (1) radically underdetermined, which means that even if the argument *is* sound, we will not be justified in believing it to be so: even if our current physical theory were in fact entirely correct and complete, the epistemic problem that we could not know it to be so would still remain – if modesty with respect to our cognitive powers in light of past failure is ever a good attitude to take, then it is always a good attitude to take.[3] Assessment of premiss (1) will therefore have to be indirect, focussing on structural similarities between the world as it is described by physics and the world as it is objectively. In order to do this, however, we first need to get clearer about what is meant by 'the objective world'.

4.2 The Objective World:

Intuitively, the 'objective world' is that which exists independent of any thought or experience of it; in Williams's term it is "what is there *anyway*" (1978: 64). This, however, is vague: to assess the identity claim made in premiss (1) of the Argument from Physics and thereby determine the strength of the eliminativist's conclusion, we need to give the notion some determinate content.

One way of giving content to the notion of the objective world is in terms of Williams's "absolute conception of reality". The absolute conception of reality is a characterisation of the world 'as it is in itself' – a characterisation, that is, that is not dependent on any particular "point of view", where "point of view" covers amongst other things spatial and temporal location, partiality, self-interest, and

---

constraints must be imposed on any such account (for a partial list of these constraints, see 1994: 123). Even if we accept his account of physics as "the branch of science concerned with identifying a basic class of objects and attributes and a class of principles that are sufficient for an account of space-time and of the composition, dynamics, and interactions of all occupants of space-time" (1994: 124), it gives no more than nominal support to premiss (2), as Poland himself freely admits that "no one knows" what theory answers to this characterisation, and consequently whether it will quantify over colours.

[3] This objection is distinct from that presented by Hempel's dilemma: if based on present physics, physicalism is false; if based on some future physics, physicalism is contentless. Even if we allow that physicalism does not need to be identified relative to an actual body of physical theory to have a determinate content, we are still no closer to knowing what this content is.

social and historical perspective.[4] In localised cases, failure to account for a subject's point of view can breed error and confusion. Taking appearances at face value, if, for example, you look at a coin from an oblique angle it appears elliptical. This appearance, however, is merely the result of your spatial location in relation to the coin; the coin is not, *in reality*, elliptical. Similarly, if you say "It is raining", and then wait a few minutes until it has stopped and say "It is not raining", your utterances are not contradictory; they are just uttered from different temporal locations.

An absolute conception of the world avoids perspectivalism of this sort. More importantly, it also aims to overcome "any systematic bias or distortion or partiality in our outlook as a whole, in our representation of the world" (Williams 1978: 66). To achieve this aim, Williams assumes that the absolute conception must not include any "anthropocentric" concepts, possession of which depends on possession of peculiarly human characteristics.[5]

Our colour concepts are paradigmatically anthropocentric. Even amongst animals that are physiologically similar to humans in respect of their visual systems, there can be significant differences in discriminatory capacities, which in turn suggests that these animals do (or at least would, if they had the conceptual capacity) conceptualise colour differently. Honeybees, for example, that like humans are "trichromats" having three different types of cone visual pigment in their eyes, are sensitive to spectral light between 300 nms and 650 nms. Humans, on the other hand, are sensitive to spectral light between 400nms and 700nms. This means that honeybees are able to discriminate hues at the ultraviolet end of the spectrum that are alien to humans, and conversely, that humans can discriminate hues at the visible end of the spectrum indistinguishable to bees. Each, therefore, has colour concepts not possessed by the other.

---

[4] Moore (1997: 2-6) distinguishes between "points of involvement", defined in terms of concerns, interests or values, and "points of view", which are defined in terms of location in the broadest possible sense (of which points of involvement form a proper subset, along with spatial and temporal location). On the basis of this distinction, he goes on to distinguish "objective" judgements, which depend on no particular point of involvement, from "absolute" judgements, which depend on no particular point of view. I shall use the terms "objective" and "absolute" interchangeably, however, as the distinction is not one that my overall project requires.

[5] "The suggestion is that there are possible descriptions of the world using concepts which are not peculiarly ours, and not peculiarly relative to our experience" (1978: 244). See also Nagel, "[to conceive of the true nature of objects] means not only not thinking of the physical world from our own point of view, but not thinking of it from a more general human perceptual point of view either" (1986: 14).

The situation is more extreme when it comes to species that are not relevantly similar to us with respect to visual apparatus. There is evidence to suggest that the visual experiences of dichromats, such as squirrels and rabbits, whose eyes contain only two types of cone visual pigment, and tetrachromats, such as pigeons and ducks, whose eyes contain oil droplets in addition to visual pigments and for whom there are four functionally significant oil droplet-visual pigment combinations, are radically different from our own. Arguably, it is not that these animals are more or less sensitive to the hues that we see, but rather that their subjective colour space is entirely different and does not admit of any mapping onto our own at all.[6] These remarks strongly suggest that our conceptualisation of colour depends directly on our physiological makeup, and therefore that our colour concepts are anthropocentric in the sense indicated above. Consequently if Williams is right, they cannot figure in a representation of reality which is corrected for the peculiarities of particular observers.

The absolute conception yields a very strong understanding of 'objectivity'. It also raises a number of important issues. First, we need to ask whether Williams, in requiring that the absolute conception be a representation of the world that is independent of the peculiarities of any particular observer, goes far enough? A negative answer to this question is suggested by the thought that we can distinguish not just two, but *three* grades of what might be called 'perspectival involvement':

(i) the view from no-*one*,

(ii) the view from no-*thing*,

(iii) the view from no-*where*.

At grade (i), our representation of the world does not involve essential reference to the point of view of any particular human observer. Grade (ii) goes further and corrects our conception of reality for peculiarly *human* features – it is at this stage that we remove any "anthropocentric" concepts from our representation of the world. But arguably, it is grade (iii) to which we need to go in order to frame a truly absolute conception of reality, for it is only at this stage that we arrive at a

[6] An experiment by Wright,and Cummings, discussed in Thompson (1995: 151), found that pigeons treat wavelengths falling either side of 540nms as belonging to different hue categories, whilst humans do not.

conception of the world that is entirely independent, not only of the peculiarities of all observers, but more importantly of all observa*tion*.

Purely for the purposes of eliminativism, it is necessary that the absolute conception be framed at the third grade of perspectival involvement, because it is only at this grade that *being coloured* drops out of the picture altogether: at grade (ii), even though objects do not have colours – are not yellow, red etc – they are still *coloured* (even if other species conceptualise colour differently, what they're conceptualising is still colour).[7] A more general reason for insisting that the absolute conception be framed at the third grade of perspectival involvement becomes apparent, however, when the second important issue raised by the absolute conception is considered: whether, that is, given the inherent perspectivalism of our position, such a conception is even in principle possible. An argument, first stated by Williams (1978: 64-6), and later developed by Moore (1997: Chapter 4), suggests that some such absolute conception must be possible; the argument, however, only conclusively establishes the possibility of an absolute conception of the world framed at grade (iii).

Suppose A and B both possess true representations of the world from different points of view. Given that both of their representations are true, there must be some way of representing this fact, i.e. representing the similarities and differences between their representations and showing how both are nevertheless representations of the same thing. But if we can form a meta-representation that relates A and (an arbitrarily selected) B to the world, then it should be possible to form a meta-representation that relates A and *every* possible representation to the world. Combining this with the meta-representations that relate B to every possible representation of the world, C to every possible representation of the world, and so on, we arrive at a single conception of reality from no particular point of view: the absolute conception of reality.

It is important to stress that this absolute conception can only be achieved within the context of a rationalistic epistemology.[8] The point is brought out well by Strawson's statement of the problem confronting the proponent of the absolute conception (1987: Appendix B). Knowledge is the result of a subject's interaction

[7] An account of colour according which yellowness and its shades are not colours is defended by Averill (1985). There is no reason to classify this account as eliminativist, however.
[8] See, for example, Williams (1978), McGinn (1983: Chapter 7), Nagel (1986: 15).

with their environment. Yet plausibly, we only count as knowing something if our putative knowledge does not contain elements that depend essentially on the particular way in which we are sensorily affected by the world. This is because it is logically possible that there should exist a different experiencing being with a radically different nature who differs significantly in the way in which he is affected by the world; and *prima facie*, it seems wrong to say that there is a 'the right way' of being affected by the world, such that my judgements are knowledge and his are not. Consequently, if *any* representation of the world necessarily depends on how its subject is affected, then there can be no such thing as a correct representation of it.

The usual response to this objection is to challenge the empiricist presupposition that knowledge can only result from a subject's interaction with their environment, and maintain instead that some non-sensory account of the world is possible: if our representation of the world does not depend on the peculiarities of our nature, then the argument above gets no foothold. And for this response to work, the absolute conception of reality must be framed at the third grade of perspectival involvement.

The problem now, however, is that it threatens to put the absolute conception beyond our grasp. Even if valid, all Williams' argument establishes is the *in principle* possibility of an absolute conception: nothing has been to said to imply that we could ever actually achieve it. Indeed, it should be obvious just how mammoth a task this would be: to arrive at the absolute conception we would need to integrate every possible way in which the world could be represented, and at the very least, it seems likely that there are ways of representing the world of which we cannot even conceive. So *we*, at least, could not arrive at an absolute conception of the world.

As with 'the physical world', then, giving determinate content to 'the objective world' understood in terms of the absolute conception of reality looks, at least at present, unfeasible. This makes it doubly hard to determine the truth of premiss (1) directly. Indirectly, however, there is good reason to suppose that the two notions would, at least in principle, coincide. For example, both physics and the absolute conception aim at universality; as Moore says, it is a "deep prejudice" (in the non-pejorative sense) of scientific enquiry that:

With regard to its basic workings, the physical world (nature) looks the same from every point of view. There are no privileged positions in the world from which its natural laws are peculiarly perspicuous, or peculiarly evident, or peculiarly simple (1997: 28).

Moreover, it is plausible to suppose that both are (at least in principle) sufficient to explain *everything* that goes on in the world, in something like Lewis's weak sense of 'explanation' discussed above (given the epistemic problems associated with physicalist accounts of colour, of course, it is important to stress the *in principle*[9]).

But even if we grant all this, what follows? First, these considerations do not establish the truth of the identity claim that the objective world *is* the physical world, as they are consistent with the weaker claim that the physical world is merely a proper subset of the objective world: that describability in purely physical vocabulary is a sufficient, but not a necessary condition, for inclusion within the objective world. This claim is too weak for the eliminativist, as it does not preclude the possibility that colours constitute part of the objective residue left over once everything that comprises the physical world has been accounted for. Unfortunately, the weaker claim is *at least* as attractive as the stronger identity claim. Williams, for instance, owes some explanation of why the objective world is the sparsely populated environment he envisages, as opposed to a world in which *all* the perceptible properties exist, for which understanding how the world is in itself consists in grasping the relations that these properties bear to each other.[10] Williams assumes that a property exists objectively if its perception does not depend on characteristics peculiar to any particular observer. This assumption, however, is problematic. Imagine a world in which colour perception is uniform across the animal kingdom and therefore where colour is not an anthropocentric concept. Does colour exist objectively in this world, despite that fact that some creatures have no visual experience? To suppose not would appear to support the more general conclusion that if some creatures are incapable of any perceptual experience whatsoever, then *no* properties exist. This, however, is surely absurd. The problem for Williams lies in saying why this is so: colour concepts, even in

[9] Williams himself concedes that explaining the place in the world of "psychological phenomena such as the perception of secondary qualities, and further, of cultural phenomena such as the local non-absolute conceptions of the world...is a programme for philosophy" (1978: 300-1).

[10] On this understanding the objective world would still be beyond our grasp, given that we are unable to conceive of all the different ways in which the world could be perceived.

this world, are anthropocentric to the extent that they depend upon having the relevant perceptual mechanisms at all. Moreover, the conclusion that it is only the weaker claim that follows receives further support from a closer consideration of the argument Williams uses to establish the in principle possibility of an absolute conception. The argument starts by taking any two *true* representations of the world and proceeds from there. But these representations can only be true if the world really is (in some sense) coloured. So why not require that this is represented in the absolute conception?

Besides, even if we grant the truth of the identity claim that 'the objective world *is* the physical world', it does not help the eliminativist because it to prove too much. Specifically, it turns out that it is not just colours that don't exist objectively; the same is true of pretty much everything else we commonly take to exist. Excluding any concept from the absolute conception, possession of which depends on possession of peculiarly human characteristics, will leave us with a distinctly limited set of concepts. Indeed, there is a sense in which *all* of our concepts are anthropocentric. Privileging concepts for medium sized objects, such as tables and chairs, for instance, is a clear prejudice of medium sized observers: assuming that tiny organisms had the capacity to conceptualise the world, it is highly unlikely that they would have concepts for the medium sized objects that we do. By Williams's principles, it would appear to follow that the objective world therefore does not even contain medium sized objects. And although not obviously absurd *per se*, this conclusion is not the eliminativist's: eliminativism is the supposedly surprising doctrine that the material objects that populate the physical world are not in fact coloured. Its interest lies in the fact that it challenges common sense only in one very specific respect. However, if not only are colours not part of the furniture of the universe, but neither is furniture (nor anything else we might ordinarily care about), then eliminativism is internally unstable: it is simply not possible to maintain both that material objects exist but that colours don't. The eliminativist can only conclude that material objects are not, despite appearances, coloured if he also concludes that these objects are not, despite appearances, objects. And though not logically incoherent, this conclusion is *much* less appealing. Premiss (1) of the Argument from Physics is therefore unacceptably strong.

CHAPTER FIVE: Non-reductive Realism

Consistent with the outright rejection of premiss (1) of the Argument from Physics are two prominent non-reductivist theories of colour: Campbell's simple view and non-reductive dispositionalism.[1] Consideration of these theories, in light of previous findings, will enable us finally to determine the nature of colour.

5.1 Non-reductive Dispositionalism:

Westphal's (1987) account of colour is a non-reductivist variant on the physicalist dispositionalism of Byrne and Hilbert (1997b) *et al*, which identifies colours with dispositions to affect the light, but which does not require that these dispositions be specified in purely physical terms: Westphal's account is neutral between physicalistic and phenomenalistic interpretations. His strategy is to take the surface reflectance profiles (SRPs) that determine perceived colour (unsuccessfully exploited by the physicalist dispositionalist) and then show how they can be reinterpreted without reference to any technical, physical, notions. Specifically, this involves replacing the x-axis value 'wavelength' with 'illuminant colour' and the y-axis value 'percentage of illuminant reflected' with 'brightness', such that instead of measuring the percentage of illuminant reflected at each wavelength, we judge how bright an object looks under different lighting conditions.[2]

The individual colours that objects instantiate are then defined in terms of the colour of light that those objects 'darken' or absorb. White objects, for example, which tend to diffusely reflect a high proportion of the light at each wavelength can be defined, phenomenalistically, as those objects that do not "darken the light" in any illumination. The chromatic colours, on the other hand, which reflect different percentages of light at different wavelengths, each darken different coloured lights and are defined accordingly: yellow objects, for example, which

[1] Byrne and Hilbert (1997c: xxi) contrast "reductive" dispositionalism, in which the same undefined colour term appears on both sides of the biconditional, with "non-reductive" versions of dispositionalism in which it doesn't. On my usage, however, "reductive" and "non-reductive" mark the distinction between those versions of dispositionalism that treat colours as falling within the compass of the physical, and those versions of dispositionalism that do not.

[2] Using the eye to make colour comparisons in this way is called "colorimetry".

reflect only a small amount of the blue light that strikes them compared to the light of the other colours that they reflect, can be defined as those objects that darken (or absorb) more blue light than light of any other colour (i.e. they appear darkest when viewed under a blue illumination); green objects are those that darken more red light than light of any other colour; and so on.

Presented thus, this account requires refinement. As Broackes points out (1992: 213), Westphal's proposal for phenomenalising an object's SRP faces difficulties when the illuminant colour is not generated by a single light, but rather by a mixture of lights. So, for example, under light of 570nms an object may reflect 85% of the incident light; however, under a mixture of 550 and 600 nms light, which looks exactly the same colour as a pure light of 570nms, the object may reflect only 75% of the incident light. Consequently, there is no 1-1 correspondence between illuminant colour and percentage of the illuminant reflected. This problem, however, merely calls for a slight complication in Westphal's account (as, indeed, it requires in the physical dispositionalist's account). For single lights, Westphal's account is fine as it is. For mixed lights, however, a range of different SRPs are needed. This is not a problem, however, (at least not in theory) so long as the lights that comprise these mixtures are themselves characterised in phenomenal terms.

A more serious difficulty concerns Westphal's characterisation of an object's colour in terms of the light that that object darkens (i.e. fails to reflect). The SRP for grass has three troughs: at roughly 400, 480 and 650 nms.[3] Of these the last is the least pronounced; yet, if Westphal's definition is correct, a green object should be one that darkens more red light than any other type of light. The problem is that grass actually darkens (fails to reflect) *more* blue and blue-green light than it does red light, and therefore Westphal's definition is incorrect. (The problem is exacerbated by the fact that after the trough at around 650nms, the SRP for grass peaks sharply to the extent that by 700nms – the boundary between the visible and infra-red spectra – the percentage of light reflected is almost twice that of the light that is reflected in the 'green' part of the spectrum. Given that this light is from the 'red' part of the spectrum, it looks more accurate to say that

[3] See Broackes (1992: 212, figure 11.2)

green objects actually reflect *more* red light than they do any other kind of light, and therefore that Westphal's definition is straightforwardly false.)

Again, however, the problem is easily remedied, this time by following Broackes' proposed variation on Westphal's account, and instead of talking about the light that objects *fail* to reflect, defining colours in terms of the light that objects *succeed* in reflecting:

> red surfaces are those that when illuminated with normal white light tend to reflect light that is, phenomenally, red (whatever the spectral composition of that light) – where for accuracy that must be interpreted: the surfaces tend in normal white light to reflect light that results in normal people when normally affected by it having a perception as of a red material object (1992: 214).

Leaving the precise details of this account aside, however, defining the colours in this way brings out "the significant fact that the [surface reflectance] curves can be obtained by direct observation, independently of the wavelength theory" (1987: 32), thus yielding what Westphal calls "a certain sort of relational common sense realism about the being coloured of coloured objects and materials" (1987: 7). What this shows is that colours are properties that can be specified independently of scientific theory and that the "relationship between colorimetry and the physics of radiation illustrates perfectly the celebrated commonsense conception of autonomous explanation advanced by Putnam" (1987: 76): explanations in terms of colorimetry bring our features of situations in which colour perception occurs, holding regardless of the underlying micro-physical facts. As such, explanations of colour perception in terms of colorimetry are epistemologically more perspicuous than those in terms of the physics of radiation. Consequently, they avoid the phenomenological objections to physicalist theories of colour discussed in Chapter 2: the problem for the physicalist was to identify similarity relations amongst the colour *properties*, conceived of in physical terms, that explain the similarity judgements that we make about them on the basis of visual experience. It was argued there that no such relations could be identified: divorcing the description of the nature of colour properties from human responses leaves our similarity judgements mysterious; explicitly introducing reference to such responses, on the other hand, explains the grounds of our similarity judgements only derivatively, or else by accruing large explanatory debts elsewhere. The advantage of non-reductive

realism is that whilst the identification of colour properties involves ineliminable reference to human experience, it does not do so derivatively.

This point will receive fuller explanation in the next section. Before it does, however, a more fundamental interpretative issue must be addressed. Westphal and Broackes identify colours with dispositions to affect the light. What exactly this means, however, bears closer scrutiny. As a response to the Argument from Physics, non-reductive dispositionalism does not require that there be any metaphysical distinction between colours and their grounds: non-reductive dispositionalists reject the deference to physical theory that forces more traditional dispositionalists to draw a distinction of kind between colours and shapes. *A fortiori*, in light of the problems identified with drawing a distinction of this sort, it is prudent for the non-reductive dispositionalist to suppose that there is in fact no such distinction: that those properties that exist are metaphysically on a par. The non-reductive dispositionalist, that is, should be a property monist.

Property monism can take one of two forms. According to causal powers theories of properties, properties (of whatever kind) are exhausted by their causal role: there is nothing more to property existence than the causal powers exhibited by a property's bearer. In contrast, according to more substantial theories of properties, properties inhere in objects independent of the causal interactions in which they dispose their bearers to partake: properties are the grounds of an object's behavioural dispositions, and are not exhaustively constituted therein. The distinction may be illustrated by reference to the weak and strong metaphysical interpretations of the dispositionalist biconditional, identified in Chapter 3. According to causal powers theories of properties,

(D) x is F iff were x to φ, then it would ψ,

exhaustively characterises the property F; it tell us what F-ness *is*. According to more substantial theories of properties, however, (D) tells us merely something *about* F-ness; that as it happens, F-instantiating objects ψ when φ. In characterising colours as *dispositions* to affect the light, non-reductive dispositionalism is most naturally understood as presupposing a causal powers theory of properties. Although epistemologically advantageous, however (see, for example, Shoemaker 1980: 236-8 and Blackburn 1990: 64), causal powers theories of properties are conceptually counter-intuitive.

Bennett, for example, claims that "western science has for centuries proceeded on the assumption ["or regulative principle"] that wherever (1) ["If x were F, it would be H] is true (2) [There is some non-dispositional φ such that: x is φ, and it is a causal law that if anything is both φ and F then it is H"] is also true" (1971: 104). Whether or not this is a regulative principle that physicists *still* adhere to (in relation to the electrical and gravitation fields posited by modern physics, for instance, this might be doubted), at the very least, Evans's more guarded remark that, "a deep conceptual prejudice of ours...is offended by dispositional properties without categorical grounds" (1980: 276), certainly *does* seem correct: 'bare' dispositions are conceptually very difficult to accept. The problem with causal powers theories of properties is that they seem unable to respect this conceptual prejudice of ours: according to the causal powers theorist, properties are dispositional 'all the way down'; dispositional properties are grounded by properties that are themselves essentially dispositions.

Diagnosis of the debate between the causal powers theorist and the substantival property theorist in light of previous findings suggests a possible source of confusion that might explain the causal powers theorist's temptation to disregard this prejudice: like certain dispositionalists about colour, causal powers theorists appear to take too seriously conceptual truths, inferring fallaciously that because our *concepts* of properties are determined by the behaviour that bearers of these properties are disposed to exhibit, the *properties* to which these concepts refer are themselves exhaustively constituted by this behaviour. Conceptual prejudices may, of course, be prejudices in the pejorative as well as the non-pejorative sense of the word. Nevertheless, being able to explain away the attraction of a position that is otherwise problematic does seem to stack the odds firmly against this view.

These considerations would need be developed in a lot more detail to prove conclusive. But they at least suggest that our account of colour should presuppose a more substantial theory of properties than that offered by the causal powers theorist.

Granting this has important implications for non-reductive dispositionalism: non-reductive dispositionalism ceases to be an account of what colours *are* and becomes instead an account of what they are like. The claim is no longer that, for example, greenness (metaphysically speaking) *is* a disposition to reflect light that

is phenomenally green, but rather that 'x is green iff it is disposed to reflect light that is phenomenally green' merely represents a truth *about* greenness.[4]

The advantage of conceiving of the interrelation between theories of properties and theories of colours in theses terms is that we are now in a position to explain the attraction of the dispositional theses considered in earlier chapters: whilst there may only be one uniquely adequate description of what a thing *is*, there can nevertheless be many truths *about* it. From the perspective of a substantial theory of properties, dispositional theories of colour therefore cease to be in direct competition with each other: not only is it a truth about the property green that it causes its (material object) bearers to reflect light that is phenomenally green (or, if Westphal is correct, darken the red light that strikes it), so it is equally true that it causes its bearers to reflect more light in the middle-wavelength part of the visible spectrum than in the long-wavelength part, and roughly the same amount of light in the short-wavelength part of the spectrum as in the long- and middle-wavelength parts put together. Similarly, it is true that in reflecting light of this composition, objects instantiating greenness cause normal observers in standard conditions to undergo experiences that represent that object as green; which in turn is equivalent to saying that these objects cause normal observers in standard conditions to undergo experiences in which the red-green channel codes green and the yellow-blue channel codes zero. These claims do not compete for the same logical space, but are rather pitched at different explanatory levels. Depending on the context, one may be more appropriate than another, but ultimately none is 'more correct' than any of the others; they just give us different types of information about colour properties.

5.2 The Simple View:

Whilst the realisation that premiss (1) of the Argument of Physics is unworkably strong blunts the eliminativist's attack, and dispositional theories of colour can be exploited to tell us a great deal *about* what the colour properties whose existence is thereby assured are like, we have still to determine what colours *are*. The

[4] Westphal does not himself think that his theoretical definitions are adequate for the purposes of reductively defining colours (1987: 34), although his reasons for this differ from those presented here.

"simple" view of colour, proposed by Campbell (1993), fills this lacuna.[5] According to the simple view, colours are simple properties that ground an object's colour-related dispositions and whose real nature is transparent to us.

In claiming that colours are properties that ground an object's colour-related dispositions, the simple view of colour, like the primary quality view of colour, appears to get the metaphysics of properties right. It does not, for example, presuppose the qualitatively uneconomical, counter-intuitive and ultimately unfeasible metaphysical distinction between the dispositional and the categorical required by the metaphysical dispositionalist. Neither does it apparently call for the questionable property monism - in the form of a causal powers theory of properties - that fits most naturally with the physicalist and non-reductive dispositionalists' characterisation of colours as *dispositions*. In stark contrast to the primary quality view of colour, however, the simple view is able to meet the adequacy constraint on putative theories of colour that they explain the similarity judgements concerning the colours we are disposed to make by assuming that the intrinsic nature of colours is transparent to us: the reason why we judge that (for example) red is more similar to orange than it is to blue, according to the simple view, is that red just *is* more similar to orange than it is to blue, and moreover that this is a fact of which we can be directly aware.

The claim that the nature of colour is laid bare to us in visual experience is not, of course, uncontroversial. Evans, as we have already seen, for example, objects that it is unclear what the difference between metaphysical dispositionalism and theories of colour that presuppose the intelligibility of conceiving of colours as abiding properties of objects that resemble our experiences of them is supposed to consist in, given that we are unable to make sense of an exemplification of a property of experience in the absence of any experience (1980: 272-3). Similarly Smith objects that if colours were transparent then colour illusion and unperceived colour would both be inexplicable:

> what it is about colour that makes colour illusion and unperceived colour possible...is not manifest to us in colour experience because colour experience merely gives us the 'experience' side of the equation, and what we want is an account of why the 'experience' side of the equation may yet be an unreliable indicator of an object's colour (1993: 274).

[5] A similar account is defended by Yablo (1995).

Neither objection, however, is conclusive.

Smith's objection, for example, confuses awareness of the intrinsic nature of a property with awareness of everything there is to know about a property. An account of why the 'experience' side of the equation may yet be an unreliable indicator of an object's colour will depend upon factors extrinsic to the object: whether it is being viewed under suitable lighting conditions, from a suitable position and by a suitable observer. If these environmental conditions do not hold, then there is a good chance that the experience will be illusory. It is no part of the simple theorist's claim, however, that whether optimal conditions for viewing a particular colour obtain is manifest in experience. The simple theorist's view is merely that the *intrinsic* (i.e. non-relational) nature of colour properties is laid bare in experience, not that no further knowledge *about* these properties is possible. The simple theorist is therefore committed only to transparency as it is characterised by Russell: "so far as concerns knowledge of the colour itself, *as opposed to knowledge of truths about it*, I know the colour perfectly and completely when I see it and no further knowledge of it itself is even theoretically possible" (1912: 47, emphasis added). In particular, he should not accept the thesis that is implied by Strawson's statement of the doctrine: "colour words are words for properties which are of such a kind that their *whole* and essential nature as properties can be and is fully revealed in sensory-quality experience given only the qualitative character that that experience has" (1989: 224, emphasis added). The transparency of colour, that is, is perfectly consistent with the exploitation of dispositional theories to tell us facts *about* the conditions under which colours can best be seen; transparency is a thesis of limited scope.

The considerations adduced by Evans are more persuasive, but establish (as we have already seen) only the truth of the conceptual claim that our colour concepts are concepts *as of* dispositional properties; it does not follow from this, however, that the properties of which these concepts are concepts of actually *are* dispositions. Just because we cannot conceive of the existence of abiding colour properties over and above the types of experience they dispose their bearers to produce, it does not follow that there are no such properties. Indeed, as an inference to the best explanation in light of the failure of both physicalism and metaphysical dispositionalism, it looks as though in this respect, conceivability is simply not a reliable guide to possibility.

Granting that colours are simple properties whose intrinsic nature is transparent allows us to reject the eliminativist's conclusion that material object's are not, despite appearances, coloured. It is does not, however, uniquely determine the picture of reality with which we are left. Specifically, a question still remains concerning the relation that these simple colour properties bear to the natural physical properties. Consistent with the simple view are two models of the interrelation between physics and colour. According to the first, colours are emergent properties that, though perhaps dependent upon the instantiation of natural properties for their existence, are otherwise entirely independent of physical properties.[6] According to the second, in contrast, colour properties supervene upon physical properties: there is no possible world in which the perfectly natural properties are exactly as they are in this world but in which the colour properties that objects instantiate differ.[7]

Unlike the second model of the relationship between physics and colour, the first treats colour properties and physical properties as equal partners; it accords no special status to the physical. Such a position is radically non-reductive – there does not even exist a dependence relation between the non-physical on the physical. As such, however, it is far too extreme; it fails to present a satisfactorily integrated account of the world.

It was argued in Chapter 2 that colours are not adequately describable within a physical vocabulary; the 'adequately', however, deserves emphasis. It is not that there is *no* sense in which colours are physical properties of objects, but rather that describing them in physical terms is epistemically imperspicuous. Colours, that is, do not 'float free' of physical properties, but co-exist as part of the same world. To suppose otherwise severs the causal relationship between physical properties and colour properties: if colours can vary independently of physical properties then a difference in neither will causally explain a difference in the other, and we will have to conclude that the two causal systems are unconnected.

---

[6] This view is attributed to Broad by Tye (2000: 148). This attribution, however, may not be entirely accurate. Whilst Broad concedes that the relation between the world of physics and the world of sensible appearance is not a straightforward one, he is optimistic that someday someone will unite these worlds into "the one whole of Nature" (1923: 548). Whether this unification will involve the discovery that the world of appearances supervenes on the physical world is, presumably, left open.

[7] See, for example, Campbell (1993: 178).

But this is absurd.[8] The world that is described by fundamental physics is exactly the same world as that which is described by biology, colour science, sociology etc.; physics just offers us information about a different (albeit more basic) aspect of this world.

The position in which we find ourselves is perhaps best illustrated in relation to the pessimistic realism of Stroud, discussed in Chapter 1.2. Stroud's position with respect to the question of the interrelation between the physical and the non-physical is agnosticism. Due to the inherent paucity of our epistemic position, he believes that we are unable to determine what relation the one bears to the other: whilst he argues that there is absolutely no reason to accept the Argument from Physics, and with it the eliminativist's claim that colours do not exist, he believes that neither is there any reason to draw the converse conclusion that material objects really are coloured. The result is dissatisfaction.

Rejecting Stroud's pessimism concerning the prospects for human knowledge, the proponent of the simple view of colour, argues that not only can we be assured of the existence of colour, but we can even determine what colours *are*: colours are properties whose intrinsic natures are transparent, that ground an object's colour-related dispositions. Moreover, on this view, we are not destined to ignorance concerning the relationship between the physical and the non-physical: the relationship is one of interdependence; the non-physical supervenes upon the physical. And this allows the simple theorist to respect the intuition that made the problematic premiss (1) of the Argument from Physics look at all plausible in the first place. Quine once asked, "Why...this special deference to physical theory?", replying that "nothing happens in the world, not the flutter of an eyelid, not the flicker of a thought, without some redistribution of microphysical states" (1981: 98). To respect this dictum, we do not need to assume, like the eliminativist, that only those properties that can be described in terms of fundamental physics exist. We do not even need to assume that there is any substantial metaphysical distinction between the properties of fundamental physics and properties like colours. It is enough, to respect this intuition, that there be a suitable dependency relation between colours and the properties of

[8] As Locke observed, for example, altering the physical properties of an almond (by pounding it) causes an alteration in the almond's colour properties (1975: 2.8.20).

fundamental physics. Once we accept this, there is no reason to doubt that colours are simple transparent properties of objects.[9]

[9] I would like to thank the AHRB for financial support whilst writing this thesis and my supervisor, Michael Martin, for many invaluable comments and suggestions.

Bibliography

Adams, E. 1975. *The Logic of Conditionals*. Dordrecht: Reidel Publishing Company.

Armstrong, D.M. 1968. *A Materialist Theory of Mind*. London: Routledge.

- 1978. *A Theory of Universals, Volume 2*. Cambridge: Cambridge University Press.

- 1987. "Smart and the Secondary Qualities", reprinted in Byrne and Hilbert 1997a.

Armstrong, D.M., Martin, C.B. and Place, U. 1996. *Dispositions: a Debate*, T. Crane (ed). London: Routledge.

Averill, E.W. 1985. "Color and the Anthropocentric Problem", reprinted in Byrne and Hilbert 1997a.

Bennett, J. 1971. *Locke, Berkeley, Hume: Central Themes*. Oxford: Clarendon Press.

Berkeley, G. 1998. *A Treatise Concerning the Principles of Human Knowledge*, J. Dancy (ed). Oxford: Oxford University Press.

Blackburn, S. 1990. "Filling in Space", *Analysis 50*, 62-5.

- 1993. "Circles, Finks, Smells and Biconditionals", in Tomberlin (ed), *Philosophical Perspectives 4*. Northridge, Calif.: Ridgeview.

Boghossian, P.A. and Velleman, J.D. 1989. "Colour as a Secondary Quality", reprinted in Byrne and Hilbert 1997a.

- 1991. "Physicalist Theories of Color", reprinted in Byrne and Hilbert 1997a.

Bolton, M.B. 1983. "Locke and Pyrrhonism: The doctrine of primary and secondary qualities", in Burnyeat (ed), *The Skeptical Tradition*. Berkeley: University of California Press.

Broackes, J. 1992. "The Autonomy of Colour", reprinted in Byrne and Hilbert 1997a.

Broad, C.D. 1923. *Scientific Thought*. New York: Humanities Press.

Byrne, A. and Hilbert, D.R. 1997a. *Readings On Color, Volume 1: The Philosophy of Color*. Cambridge, Mass.: MIT Press.

- 1997b. "Colors and Reflectances", in Byrne and Hilbert 1997a.

- 1997c. "Introduction", to Byrne and Hilbert 1997a.

Campbell, J. 1993. "A Simple View of Colour", reprinted in Byrne and Hilbert 1997a.

Cosmides, L. and Tooby, J. 1995. "Foreword", to Baron-Cohen, *Mindblindness*. Cambridge, Mass.: MIT Press.

Cuneo, T. 2001. "Are Moral Qualities Response-Dependent?", *Nous 35*, 569-91.

Edgington, D. 1986. "Do Conditionals have Truth-Conditions?", reprinted in Jackson (ed), *Conditionals*. Oxford: Oxford University Press.

Evans, G. 1980. "Things Without the Mind – A Commentary upon Chapter Two of Strawson's *Individuals*", reprinted in *Collected Papers*. Oxford: Clarendon Press.

Galileo, G. 1957. "The Assayer", reprinted in Drake (trans), *Discoveries and Opinions of Galileo*. New York: Anchor Books.

Hacker, P.M.S. 1987. *Appearance and Reality*. Oxford: Blackwell.

Hardin, C.L. 1993 (2nd edition). *Color for Philosophers*. Indianapolis: Hackett.

- 1997. "Reinverting the Spectrum", in Byrne and Hilbert 1997a.

Harman, G. 1990. "The Intrinsic Quality of Experience", in Tomberlin (ed), *Philosophical Perspectives 4*. Northridge, Calif.: Ridgeview.

- 1996. "Explaining Objective Color in Terms of Subjective Reactions", reprinted in Byrne and Hilbert 1997a.

Hilbert, D.R. 1987. *Color and Color Perception*. Stanford: Center for the Study of Language and Information.

Hilbert, D.R. and Kalderon, M.E. 2000. "Color and the Inverted Spectrum", in Davis (ed), *Vancouver Studies in Cognitive Science 9*. Oxford: Oxford University Press.

Hinton, J.M. 1967. "Visual Experiences", *Mind 76*, 217-227.

Hirsch, E. 1993. *Dividing Reality*. Oxford: Oxford University Press.

Jackson, F. 1977. *Perception: A Representative Theory*. Cambridge: Cambridge University Press.

- 1996. "The Primary Quality View of Colour", in Tomberlin (ed), *Philosophical Perspectives 10*. Cambridge, MA: Blackwell.

- 2000. "Philosophizing About Color", in Davis (ed), *Vancouver Studies in Cognitive Science 9*. Oxford: Oxford University Press.

Jackson, F. and Pargetter, R. 1987. "An Objectivist's Guide to Subjectivism About Colour", reprinted in Byrne and Hilbert 1997a.

Johnston, M. 1992. "How to Speak of the Colors", reprinted in Byrne and Hilbert 1997a.

Kripke, S. 1980. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.

Kuehni, R. 1997. *Colour: An Introduction to Practice and Principles*. New York: John Wiley and Sons.

Lewis, D. 1983. "New Work from a Theory of Universals", reprinted in Mellor and Oliver (eds), *Properties*. Oxford: Oxford University Press.

- 1986. "Causal Explanation", in *Philosophical Papers Volume 2*. Oxford: Oxford University Press.

- 1997. "Finkish Dispositions", *Philosophical Quarterly 47*, 143-58.

Locke, J. 1975. *An Essay Concerning Human Understanding*, P.H. Nidditch (ed). Oxford: Oxford University Press.

Mackie, J.L. 1976. *Problems from Locke*. Oxford: Clarendon Press.

- 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.

Martin, C.B. 1994. "Dispositions and Conditionals", *Philosophical Quarterly 44*, 1-8.

McDowell, J. 1985. "Values and Secondary Qualities", in Honderich (ed), *Morality and Objectivity: A Tribute to J.L. Mackie*. London: Routledge.

McGinn, C. 1983. *The Subjective View*. Oxford: Clarendon Press.

- 1996. "Another Look at Colour", reprinted in *Knowledge and Reality: Selected Essays*. Oxford: Clarendon Press.

Mellor, D.H. 1974. "In Defence of Dispositions", reprinted in *Matters of Metaphysics*. Cambridge: Cambridge University Press.

- 2000. "The Semantics and Ontology of Dispositions", *Mind 109*, 575-580.

Melnyck, A. 1997. "How to Keep the 'Physical' in Physicalism", *Journal of Philosophy 94*, 622-37.

Moore, A.W. 1997. *Points of View*. Oxford: Oxford University Press.

Mumford, S. 1998. *Dispositions*. Oxford: Oxford University Press.

Nagel, T. 1986. *The View from Nowhere*. Oxford: Oxford University Press.

Peacocke, C. 1983. *Sense and Content*. Oxford: Clarendon Press.

Poland, J.S. 1994. *Physicalism*. Oxford: Clarendon Press.

Price, H.H. 1947. "Obituary Notice: H.A. Prichard", *Proceedings of the British Academy*, 331-50.

Prior, E. 1985. *Dispositions*. Aberdeen: Aberdeen University Press.

Putnam, H. 1975. "Philosophy and our Mental Life", in *Mind, Language and Reality, Volume 2*. Cambridge: Cambridge University Press.

Quine, W.V.O. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.

- 1969. "Natural Kinds", reprinted in Stalker (ed), *Grue!* Illinois: Open Court.

- 1981. "Goodman's *Ways of Worldmaking*", reprinted in *Theories and Things*. Cambridge, Mass.: Harvard University Press.

Reid, T. 1983. *Essays on the Intellectual Powers*, in Lehrer and Beanblossom (eds), *Thomas Reid's Inquiry and Essays*. Indianapolis: Bobbs-Merrill.

Russell, B. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press.

Ryle, G. 1949. *The Concept of Mind*. Harmondsworth: Penguin.

Shoemaker, S. 1980. "Causality and Properties", reprinted in Mellor and Oliver (eds), *Properties*. Oxford: Oxford University Press.

- 1994. "Phenomenal Character", reprinted in Byrne and Hilbert 1997a.

Smart, J.J.C. 1975. "On Some Criticisms of a Physicalist Theory of Colors", reprinted in Byrne and Hilbert 1997a.

Smith, M. 1993. "Colour, Transparency, Mind-Independence", in Haldane and Wright (eds), *Reality, Representation, and Projection*. Oxford: Oxford University Press.

Smith, P. 1987. "Subjectivity and Colour Vision", *Proceedings of the Aristotelian Society Supplementary Volume 61*, 245-64.

Snowdon, P. 1981. "Perception, Vision and Causation", *Proceedings of the Aristotelian Society 81*, 175-192.

Strawson, G. 1987. *The Secret Connexion*. Oxford: Oxford University Press.

- 1989. "Red' and Red", *Synthese 78*, 193-232.

Stroud, B. 2000. *The Quest for Reality*. Oxford: Oxford University Press.

Thompson, E. 1995. *Colour Vision*. London: Routledge.

Travis, C. 1996. "Meaning's Role in Truth", *Mind 105*, 451-66.

Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge, Mass.: MIT Press.

Westphal, J. 1987. *Colour: Some Philosophical Problems from Wittgenstein*. Oxford: Blackwell.

Wiggins, D. 1987. "A Sensible Subjectivism?", in *Needs, Values, Truth*. Oxford:

Blackwell.

Williams, B. 1978. *Descartes: The Project of Pure Enquiry*. London: Penguin.

Wright, C. 1992. *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press.

Yablo, S. 1995. "Singling Out Properties", in Tomberlin (ed), *Philosophical Perspectives 9*. Cambridge, MA: Blackwell.

Bibliography

Adams, E. 1975. *The Logic of Conditionals*. Dordrecht: Reidel Publishing Company.

Armstrong, D.M. 1968. *A Materialist Theory of Mind*. London: Routledge.

- 1978. *A Theory of Universals, Volume 2*. Cambridge: Cambridge University Press.

- 1987. "Smart and the Secondary Qualities", reprinted in Byrne and Hilbert 1997a.

Armstrong, D.M., Martin, C.B. and Place, U. 1996. *Dispositions: a Debate*, T. Crane (ed). London: Routledge.

Averill, E.W. 1985. "Color and the Anthropocentric Problem", reprinted in Byrne and Hilbert 1997a.

Bennett, J. 1971. *Locke, Berkeley, Hume: Central Themes*. Oxford: Clarendon Press.

Berkeley, G. 1998. *A Treatise Concerning the Principles of Human Knowledge*, J. Dancy (ed). Oxford: Oxford University Press.

Blackburn, S. 1990. "Filling in Space", *Analysis 50*, 62-5.

- 1993. "Circles, Finks, Smells and Biconditionals", in Tomberlin (ed), *Philosophical Perspectives 4*. Northridge, Calif.: Ridgeview.

Boghossian, P.A. and Velleman, J.D. 1989. "Colour as a Secondary Quality", reprinted in Byrne and Hilbert 1997a.

- 1991. "Physicalist Theories of Color", reprinted in Byrne and Hilbert 1997a.

Bolton, M.B. 1983. "Locke and Pyrrhonism: The doctrine of primary and secondary qualities", in Burnyeat (ed), *The Skeptical Tradition*. Berkeley: University of California Press.

Broackes, J. 1992. "The Autonomy of Colour", reprinted in Byrne and Hilbert 1997a.

Broad, C.D. 1923. *Scientific Thought*. New York: Humanities Press.

Byrne, A. and Hilbert, D.R. 1997a. *Readings On Color, Volume 1: The Philosophy of Color*. Cambridge, Mass.: MIT Press.

- 1997b. "Colors and Reflectances", in Byrne and Hilbert 1997a.

- 1997c. "Introduction", to Byrne and Hilbert 1997a.

Campbell, J. 1993. "A Simple View of Colour", reprinted in Byrne and Hilbert 1997a.

Cosmides, L. and Tooby, J. 1995. "Foreword", to Baron-Cohen, *Mindblindness*. Cambridge, Mass.: MIT Press.

Cuneo, T. 2001. "Are Moral Qualities Response-Dependent?", *Nous 35*, 569-91.

Edgington, D. 1986. "Do Conditionals have Truth-Conditions?", reprinted in Jackson (ed), *Conditionals*. Oxford: Oxford University Press.

Evans, G. 1980. "Things Without the Mind – A Commentary upon Chapter Two of Strawson's *Individuals*", reprinted in *Collected Papers*. Oxford: Clarendon Press.

Galileo, G. 1957. "The Assayer", reprinted in Drake (trans), *Discoveries and Opinions of Galileo*. New York: Anchor Books.

Hacker, P.M.S. 1987. *Appearance and Reality*. Oxford: Blackwell.

Hardin, C.L. 1993 (2nd edition). *Color for Philosophers*. Indianapolis: Hackett.

- 1997. "Reinverting the Spectrum", in Byrne and Hilbert 1997a.

Harman, G. 1990. "The Intrinsic Quality of Experience", in Tomberlin (ed), *Philosophical Perspectives 4*. Northridge, Calif.: Ridgeview.

- 1996. "Explaining Objective Color in Terms of Subjective Reactions",

reprinted in Byrne and Hilbert 1997a.
Hilbert, D.R. 1987. *Color and Color Perception*. Stanford: Center for the Study of Language and Information.
Hilbert, D.R. and Kalderon, M.E. 2000. "Color and the Inverted Spectrum", in Davis (ed), *Vancouver Studies in Cognitive Science 9*. Oxford: Oxford University Press.
Hinton, J.M. 1967. "Visual Experiences", *Mind 76*, 217-227.
Hirsch, E. 1993. *Dividing Reality*. Oxford: Oxford University Press.
Jackson, F. 1977. *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- 1996. "The Primary Quality View of Colour", in Tomberlin (ed), *Philosophical Perspectives 10*. Cambridge, MA: Blackwell.
- 2000. "Philosophizing About Color", in Davis (ed), *Vancouver Studies in Cognitive Science 9*. Oxford: Oxford University Press.
Jackson, F. and Pargetter, R. 1987. "An Objectivist's Guide to Subjectivism About Colour", reprinted in Byrne and Hilbert 1997a.
Johnston, M. 1992. "How to Speak of the Colors", reprinted in Byrne and Hilbert 1997a.
Kripke, S. 1980. *Naming and Necessity*. Cambridge, Mass.: Harvard University Press.
Kuehni, R. 1997. *Colour: An Introduction to Practice and Principles*. New York: John Wiley and Sons.
Lewis, D. 1983. "New Work from a Theory of Universals", reprinted in Mellor and Oliver (eds), *Properties*. Oxford: Oxford University Press.
- 1986. "Causal Explanation", in *Philosophical Papers Volume 2*. Oxford: Oxford University Press.
- 1997. "Finkish Dispositions", *Philosophical Quarterly 47*, 143-58.
Locke, J. 1975. *An Essay Concerning Human Understanding*, P.H. Nidditch (ed). Oxford: Oxford University Press.
Mackie, J.L. 1976. *Problems from Locke*. Oxford: Clarendon Press.
- 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
Martin, C.B. 1994. "Dispositions and Conditionals", *Philosophical Quarterly 44*, 1-8.
McDowell, J. 1985. "Values and Secondary Qualities", in Honderich (ed), *Morality and Objectivity: A Tribute to J.L. Mackie*. London: Routledge.
McGinn, C. 1983. *The Subjective View*. Oxford: Clarendon Press.
- 1996. "Another Look at Colour", reprinted in *Knowledge and Reality: Selected Essays*. Oxford: Clarendon Press.
Mellor, D.H. 1974. "In Defence of Dispositions", reprinted in *Matters of Metaphysics*. Cambridge: Cambridge University Press.
- 2000. "The Semantics and Ontology of Dispositions", *Mind 109*, 575-580.
Melnyck, A. 1997. "How to Keep the 'Physical' in Physicalism", *Journal of Philosophy 94*, 622-37.
Moore, A.W. 1997. *Points of View*. Oxford: Oxford University Press.
Mumford, S. 1998. *Dispositions*. Oxford: Oxford University Press.
Nagel, T. 1986. *The View from Nowhere*. Oxford: Oxford University Press.
Peacocke, C. 1983. *Sense and Content*. Oxford: Clarendon Press.
Poland, J.S. 1994. *Physicalism*. Oxford: Clarendon Press.
Price, H.H. 1947. "Obituary Notice: H.A. Prichard", *Proceedings of the British*

Academy, 331-50.
Prior, E. 1985. *Dispositions*. Aberdeen: Aberdeen University Press.
Putnam, H. 1975. "Philosophy and our Mental Life", in *Mind, Language and Reality, Volume 2*. Cambridge: Cambridge University Press.
Quine, W.V.O. 1960. *Word and Object*. Cambridge, Mass.: MIT Press.
- 1969. "Natural Kinds", reprinted in Stalker (ed), *Grue!* Illinois: Open Court.
- 1981. "Goodman's *Ways of Worldmaking*", reprinted in *Theories and Things*. Cambridge, Mass.: Harvard University Press.
Reid, T. 1983. *Essays on the Intellectual Powers*, in Lehrer and Beanblossom (eds), *Thomas Reid's Inquiry and Essays*. Indianapolis: Bobbs-Merrill.
Russell, B. 1912. *The Problems of Philosophy*. Oxford: Oxford University Press.
Ryle, G. 1949. *The Concept of Mind*. Harmondsworth: Penguin.
Shoemaker, S. 1980. "Causality and Properties", reprinted in Mellor and Oliver (eds), *Properties*. Oxford: Oxford University Press.
- 1994. "Phenomenal Character", reprinted in Byrne and Hilbert 1997a.
Smart, J.J.C. 1975. "On Some Criticisms of a Physicalist Theory of Colors", reprinted in Byrne and Hilbert 1997a.
Smith, M. 1993. "Colour, Transparency, Mind-Independence", in Haldane and Wright (eds), *Reality, Representation, and Projection*. Oxford: Oxford University Press.
Smith, P. 1987. "Subjectivity and Colour Vision", *Proceedings of the Aristotelian Society Supplimentary Volume 61*, 245-64.
Snowdon, P. 1981. "Perception, Vision and Causation", *Proceedings of the Aristotelian Society 81*, 175-192.
Strawson, G. 1987. *The Secret Connexion*. Oxford: Oxford University Press.
- 1989. "Red' and Red", *Synthese 78*, 193-232.
Stroud, B. 2000. *The Quest for Reality*. Oxford: Oxford University Press.
Thompson, E. 1995. *Colour Vision*. London: Routledge.
Travis, C. 1996. "Meaning's Role in Truth", *Mind 105,* 451-66.
Tye, M. 2000. *Consciousness, Color, and Content*. Cambridge, Mass.: MIT Press.
Westphal, J. 1987. *Colour: Some Philosophical Problems from Wittgenstein*. Oxford: Blackwell.
Wiggins, D. 1987. "A Sensible Subjectivism?", in *Needs, Values, Truth*. Oxford: Blackwell.
Williams, B. 1978. *Descartes: The Project of Pure Enquiry*. London: Penguin.
Wright, C. 1992. *Truth and Objectivity*. Cambridge, Mass.: Harvard University Press.
Yablo, S. 1995. "Singling Out Properties", in Tomberlin (ed), *Philosophical Perspectives 9*. Cambridge, MA: Blackwell.