chapter 4

The Primary Quality View of Colour

The Location Problem for Colour

THERE is an important sense in which we know the live possibilities as far as colour is concerned. We know that objects have dispositions to look one or another colour, that they have dispositions to modify incident and transmitted light in ways that underlie their dispositions to look one or another colour, that they have physical properties that are responsible for both these dispositions, and that subjects have experiences as of things looking one or another colour. We also know that this list includes all the possibly relevant properties. Some say that the completeness of this list is an empirical discovery of science; others that the view that it might have turned out that redness, say, is a feature of reality additional to, and different in kind from, those listed—a non-dispositional, intrinsic feature of surfaces quite distinct from their physical properties—is some kind of conceptual confusion. Either way, the list is complete. Also, we have words for the listed properties—I used them in giving the list. But these words are not colour names as such; they are rather terms for dispositions to look coloured and affect light, for the physical property bases of these dispositions, and for certain perceptual experiences. Colour thus presents a classic example of the location problem. The colours must, if they are instantiated anywhere, be findable somehow, somewhere in accounts that mention dispositions to look coloured and affect light, the physical bases of these dispositions, and colour experiences; it must be the case that some of these properties have colour names as well as names from our list. Our question is, which ones?

My answer is the 'Australian' view that colours are physical properties of objects: certain physical properties of objects have colour names as well as their physical property names. This view is sometimes known as the primary quality view of colour, although

the idea is not that colours are identical with complexes of primary qualities in a sense tied to Locke's famous list, but rather that they are identical with complexes of certain of the properties the physical sciences appeal to, or will appeal to, in their causal explanations of things' looking coloured.

How might you argue for this view, or indeed for any view about which properties are named by the colour terms? You might, of course, stipulate that in your mouth the word 'red', say, names the disposition to look red, or perhaps that it names the relevant feature of the experience that we call something's looking red to one, but that would hardly address the question of which property the word 'red' names in the mouths of others, and, more generally, in the mouths of the folk. In order to address that question, we need to start with what we find most obvious about colour. Accordingly, I start by emphasizing what seems most obvious about colour, the fact that is sufficiently central to count as defining our subject. We will see how this fact, when combined with what science tells us, forces us to identify colours with certain physical properties. I then note some properties of the resulting account of colour, including how it accommodates what is right about the dispositional view of colour. The final part of the chapter is concerned with certain well-known objections to the primary quality view that arise, as is only proper given our starting-point, from folk views about colour that seem, when combined with certain empirical facts, to be inconsistent with identifying colours with physical properties.

THE PRIME INTUITION ABOUT COLOUR

The Visually Conspicuous Nature of the Colours

There is something peculiarly visually conspicuous about the colours. Redness is visually presented in a way that having inertial mass and being fragile, for instance, are not. When we teach the meanings of the colour words, we aim to get our hearers to grasp the fact that they are words for the properties putatively presented in visual experience when things look coloured. By contrast, the term 'square' picks out a property that is *only* visually conspicuous

in objects that are coloured (in the wide sense in which anything not completely transparent is coloured).

However, although colours present themselves in visual experience in a peculiarly conspicuous way, we do not use 'red' as the name of the experience itself, but rather of the property of the object putatively experienced when it looks red. For we examine objects to determine their colour; we do not introspect. We look out, not in. Moreover, we hold objects up to the light and look carefully before ruling on their colour; and we regard the opinions of others, particularly others visually better placed than we are, as relevant to arriving at the right judgement concerning an object's colour. In sum, the ways we arrive at judgements about the colours of objects have the distinctive hallmarks of the ways we arrive at judgements about the nature of the objects we interact with. Our judgements of colour seek to conform themselves to the nature of these objects, despite the fact the colour an object seems to have has special authority in determining the colour it is.

We can sum this up by saying that some such clause as:

> 'red' denotes the property of an object putatively presented in visual experience when that object looks red

is a subject-determining platitude for red. Let's call this platitude, and the corresponding platitudes for yellow, green, and so on, the prime intuition about colour. The prime intuition is simply that red is the property objects look to have when they look red—and if this sounds like a triviality, as surely it does, that is all to the good. It is evidence that we have found a secure starting-place.

Causation and Presentation

Despite its trivial sound, our prime intuition tells us something important about the metaphysics of colour when we combine it with plausible views about what is required for an experience to be the presentation of a property.

The question: How must experience *E* be related to property *P* to count as the presentation of *P*, or, equivalently, to count as *E* representing in experience that something is *P*? is a notoriously difficult one. Nevertheless, part of the story is relatively uncontroversial. A necessary condition for *E* to be the presentation of *P* is that there be a causal connection in normal cases. Sensations of

heat are the way heat, that is, molecular kinetic energy in the case of objects whose molecules move, typically presents itself to us; and essential to this is the fact that molecular kinetic energy typically causes sensations of heat in us.

What is controversial is what is sufficient for *E* to be the presentation of *P*. We know that mere causal connection is not enough: there are far too many normal causes of any given experience. However, for present purposes we can largely set to one side the hard question of what has to be added to causation to get presentation. We can work with the rough schema: redness is the property of objects which typically causes them to look red in the right way, where the phrase 'the right way' is simply code for whatever is needed to bring causation up to presentation, for whatever is needed to make the right selection from the very many normal causes of a thing's looking red. In particular, the rough schema gives us enough to show that the dispositional theory of colour is mistaken, or so I will now argue.

THE CASE AGAINST THE DISPOSITIONAL THEORY OF COLOUR

Background on Causation

Before I present the case against the dispositional theory of colour based on the prime intuition, we need to note that properties can be causes.

How things are at one time causally affects how things are at future times. How much coffee I drink at dinner affects how much sleep I get that night; the film *The Way We Were* is about how the way its protagonists were in their youth led to how they became in middle age; how steep an incline is, is responsible for how short of breath a climber is; and so on and so forth. But talk of how things are is talk of properties; thus, to the (considerable) extent that these examples strike us as commonplaces, it is a commonplace that causation relates properties.

A good question is how to integrate this commonplace into the familiar events framework for thinking about causation. We might construe events (in the sense relevant to causation) as property-like entities. Or we might distinguish two kinds of things that can stand

in causal relations: events considered as concrete entities to be placed in the category of particulars, and, secondly, certain properties of these events. There would then be two subjects for discussion: which events cause which events, and which properties of these events are responsible for their standing in these causal relations. For it is because of the properties the events have that they stand in the causal relations that they do stand in, and, moreover, we can distinguish which properties of some cause-event matter for which properties of some effect-event—the steepness of the incline matters for how short of breath the climber is, but the colour of my sweater is neither here nor there.

It does not matter for our purposes which strategy is the right one. What matters is that properties are causes, however this fact should be integrated into our talk of events causing events.[1] With this background we can now present the case against the dispositional theory of colour.

Dispositions are not Causes

The dispositional theory of colour is mistaken because dispositions are not causes, and, in particular, are not causes of their manifestations. Their categorical bases do all the causing, where by the categorical basis of a disposition in some object, I mean the property of the object responsible for its having the disposition; that is, the property that is responsible for the object's being disposed to behave in the way definitive of the disposition in question. Consider, to illustrate the point, a fragile glass that shatters on being dropped because it is fragile, and not (say) because of some peculiarity in the way it is dropped. Suppose that it is a certain kind of bonding *B* between the glass molecules which is responsible for the glass being such that if dropped, it breaks. Then the dispositional

[1] I here skate over a large debate. For further references and more argument for the view I favour, see Frank Jackson, 'Essentialism, Mental Properties and Causation', *Proceedings of the Aristotelian Society*, 95 (1995): 253–68. For a recent statement of the other side, see Donald Davidson, 'Thinking Causes', in John Heil and Al Mele, eds., *Mental Causation* (Oxford: Clarendon Press, 1993), 3–17. Of course, when I say that properties are causes I do not mean that property *universals* are causes. When the squareness of a child's building block causes it to bump when rolled, the squareness of *my table* has nothing to do with it. I mean that how things are at certain times and places are causes.

property of being fragile is the second-order property of having some first-order property or other, bonding *B* as we are supposing, that is responsible for the glass being such as to break when dropped. And the first-order property, bonding *B*, is the categorical basis of the fragility. But then it is bonding *B*, together with the dropping, that causes the breaking; there is nothing left for the second-order property (second-order in the sense of being the property of having a property), the disposition itself, to do. All the causal work is done by bonding *B* in concert with the dropping. To admit the fragility also as a cause of the breaking would be to admit a curious, ontologically extravagant kind of overdetermination.[2] Or consider what happens when a signal is amplified by an amplifier. Surely what causes the signal to increase is not the amplifier's being an amplifier, but rather whatever features the amplifier's designers put into it that make it an amplifier.

Peter Menzies has pointed out that cases where different dispositions have the same basis raise a problem here.[3] A well-known example is the opacity and electrical conductivity of many metals. The basis for the different dispositional properties of opacity and conductivity is, roughly, the way free electrons permeate the metal; nevertheless, an explanation in terms of a metal's opacity is clearly not the same as one in terms of its conductivity. For instance, the behaviour of a galvanometer would not normally be explained by the opacity of a metal rod, but might well be explained by its conductivity. But I have to say that the cause is the same in both cases, so how can I account for the difference in explanation? I have to

[2] The thesis that dispositional properties, and functional properties in general, are not causes has been much discussed recently in connection with the question of the causal efficacy of content, see e.g. Ned Block, 'Can the Mind Change the World', in G. Boolos, ed., *Meaning and Method: Essays in Honor of Hilary Putnam* (Cambridge: Cambridge University Press, 1990), and Frank Jackson and Philip Pettit, 'Functionalism and Broad Content', *Mind*, 97 (1988): 381–400. I set aside what to say about the causal role of 'bare' dispositions, if such are possible. All the dispositions we are concerned with here are not bare; they all have bases to cause their manifestations.

[3] In discussion; the example is David Lewis's in another context. Ned Block has objected (in correspondence) that cases where different dispositions appear to have the same basis, and, more generally, cases where different functional roles appear to be occupied by the same state, turn out, on examination, to involve subtly different bases and states. But it would be strange if having learnt the lesson of multiple realizability that the same role may be filled by different states, we turned around and insisted that the converse—different roles filled by the same state—is impossible.

say that when we explain by citing a disposition, we are doing two things together: we are saying that the basis of the disposition, be it known or not, did the causing, and that what got caused has a special connection with the manifestation of the disposition. When conductivity explains the behaviour of the galvanometer, the behaviour of the galvanometer will have a special connection to a manifestation of conductivity that it lacks to any manifestation of opacity; this is why it is right to cite conductivity, and wrong to cite opacity, as the explanation of the galvanometer's behaviour. Thus, we cite electrical conductivity as the explanation when a current flow plays a special role in the path to what happens, and cite opacity when a failure of light to pass through something plays a special role in the causal path to what happens.

It follows, therefore, from the prime intuition that the colours are presented in colour experience, and so are causes or potential causes of things' looking one or another colour, that the colours are not dispositions to look coloured. They are instead the categorical bases of dispositions to look coloured. Moreover, the categorical bases of the dispositions are, we know, one or another complex of physical properties of the objects, perhaps in conjunction with their surroundings.

We can spell the argument out thus:

Pr. 1 Yellowness is the property of objects putatively presented to subjects when those objects look yellow. (Prime intuition)

Pr. 2 The property of objects putatively presented to subjects when the objects look yellow is at least a normal cause of their looking yellow. (Conceptual truth about presentation)

Pr. 3 The only causes (normal or otherwise) of objects' looking yellow are complexes of physical qualities. (Empirical truth)

Conc. Yellowness is a complex of the physical qualities of objects.

And likewise for all the colours.

The obvious analogy is with heat. Feelings of heat are the putative presentations in perceptual experience of heat. Thus, heat is

not the disposition to cause *inter alia* sensations of heat, but rather what causes the sensations of heat and the various phenomena associated with heat. But what does the causing in the right way is molecular kinetic energy. Thus, heat is molecular kinetic energy.[4]

Are the Bases Themselves Dispositions?

Bill Lycan (among many) has objected that there is no interesting distinction in kind between 'categorical' basis and disposition, and, more generally, between what occupies a functional role and the functional role occupied.[5] When we specify what fills a functional role, we simply specify some further functional property. Suppose, for example, we find that the causal basis for the disposition to look yellow in some object is a certain surface molecular configuration. Aren't molecules, Lycan would ask, in part defined in terms of the *role* they play in physical theory? Moreover, a molecular configuration can be multiply realized. Many different arrangements of molecules and their sub-molecular constituents will make up the same configuration. But, first, the question of the nature of some property is distinct from the question of the nature of the language we may use to pick it out. Non-functional and non-dispositional properties can be, and very often are, picked out via what they do—for example, in the words 'the body shape that disposes to heart attacks'. Any specification of the causal basis of the disposition to look yellow that colour science comes up with will most likely contain dispositional and functional terms—they are endemic—but it does not follow that the basis is itself a disposition. Secondly, there are two distinct senses in which a state or property may be multiply realized. The multiple realisability distinctive of dispositional and functional properties is a matter of the possibility of a number of different states doing the very same causal job. This is quite different from the fact that nearly all states are multiply realizable in the sense that they can be regarded as being, to some degree or other, disjunctive, and, accordingly, as realizable by virtue of one or another disjunct obtaining. The body

[4] I follow the usual 'convention' of ignoring molecular potential energy, and generally of grossly simplifying the science.

[5] William G. Lycan, *Consciousness* (Cambridge, Mass.: MIT Press, 1987), see ch. 4.

shape that disposes to heart attacks can be realized in many ways, but this does not mean that shape is a dispositional or functional property.

SOME FEATURES OF THE PRIMARY QUALITY VIEW OF COLOUR

I now note some properties of the primary quality account, and most especially how it accommodates the data that so famously point towards the dispositional theory, before we turn to a consideration of three objections to it.

First, the primary quality account should regard attributions of colour as *relativized to a kind of creature and a circumstance of viewing*. The primary quality account is the result of combining a causal theory of colour—the view that the colours are the properties that stand in the right causal connections to our colour experiences—with empirical information about what causes colour experiences. And a causal theory of colour takes as fundamental: colour for a kind of creature in a circumstance.

The relativity to kinds of creatures arises from the fact that which properties of the world around us stand in the right relations to certain experiences for those experiences to count as presentations of the properties is, in part, a matter of how the creatures having the experiences are, just as which kinds of intruders a burglar alarm latches onto is in part a matter of how the alarm is made, and which weather conditions a barometer records is in part a matter of how the barometer is calibrated.

The relativity to circumstances of viewing arises from the fact that the very same thing may look different colours in different circumstances, and yet there may be no substantial reason to favour one appearance over the other. For example, the coloured patches in many magazines look red from normal viewing distances but are revealed as made up of small magenta and yellow dots on closer inspection.[6] Some insist that the red appearance is an illusion. The patches are really magenta and yellow. This response faces two problems. First, it means that we are under illusion much

[6] I take the example from Mark Johnston, 'How to Speak of the Colors', *Philosophical Studies*, 68 (1992): 221–63.

more often than we naturally suppose. A lot of things look very different colours when viewed close up. Secondly, it is hard to say in any non-arbitrary way what the right viewing circumstances for the 'real' colour of an object are, and yet we know that just about any object will look different colours depending on how closely it is viewed. Famously, blood does not look red under a microscope, and nothing looks any colour under an electron microscope. Moreover, the situation is quite different from one in which a change of circumstance actually affects the object seen in some significant way. Then the right thing to say is that the object changes colour as we go from one circumstance to the other—the situation is, in principle, no different from what happens when we paint a white object red, except that the viewing circumstance 'does' the painting. But the coloured patches in the magazines do not alter as we viewers peer more closely at them. Nor does blood change when viewed through a microscope. What we need to say, accordingly, is that the colour something has in—in the sense of relative to—one circumstance may differ from the colour it has in another, where viewing distance is part of the circumstance, and that each colour is equally 'real'.

In any case this is what the causal theory must do. For it is plausible that both the looking red from a normal viewing distance and the looking made-up-of-yellow-and-magenta-dots from close up are colour experiences that count as presentations of features of what is seen. Although what must be added to causation in order to get presentation is controversial, there is a fair degree of agreement about the general shape of what is needed. We need clauses requiring that there be a systematic dependence between the nature of the experience and the nature of what is experienced, a dependence that allows us to think of the experience as tracking the nature of what is experienced, and it is plausible that there will be such dependencies both between the red-look at a reasonable distance and a patch's surface, and between the assemblies-of-yellow-and-magenta-dots look from close up and (some different feature of) the patch's surface.[7]

In sum, the causal theory should take as basic: colour for *S* in circumstance *C*, as is made explicit in the following schema:

[7] And, for those who like teleological theories of content, we could add the relevant observations about selectional history.

> *O* is red at *t* for *S* in *C* iff there is a property *P* of *O* at *t* that typically interacts with *S* in circumstances *C* to cause *O* to look red in the right way for that experience to count as the presentation of *P* to *S*.

As we are humans, we are naturally interested in redness for humans, and for humans whose perceptual faculties are working normally or properly—just as we are more interested in poisons for humans (what is poisonous for us) than in poisons for Martians (what is poisonous for them). Thus, we typically count things as red just if they have a property that interacts with normal humans to make the objects look red in such a way that their so looking counts as the presentation of the property to normal humans. Also, there is a wide range of circumstances we count as normal for viewing the world, in the sense of being circumstances that reveal the nature of it to us. For instance, seeing something from somewhere between a third of a metre and ten metres, in daylight during most parts of the day, or in typical indoor lighting, are all good for detecting the shape, distance away, size and relative position of the objects around us, and it is the colour of objects in such normal circumstances that especially interests us. We know that visual perception in these circumstances tells us more about the nature of the objects around us than what happens when we look at them at dusk, or from a kilometre away, say. Moreover, mostly objects look much the same colour in all the circumstances we count as normal. The aforementioned coloured patches are something of an exception. Accordingly, from now on I will be concerned principally with colour in a thoroughly anthropocentric sense tied to normal humans in normal circumstances. Thus, we can mostly work in terms of the following clause:

> *O* is red at *t* iff there is a property *P* of *O* at *t* that typically interacts with normal human perceivers in normal circumstances to make something that has it look red in the right way for that experience to count as the presentation of *P* in that object,

and its partners for the other colours. But the fact remains that the fundamental notion is that of the colour of *O* at *T* for S in C.[8]

[8] As far as I know, there are not equally good candidates for being normal human percipients whose colour perceptions deliver sharply different answers as to

Secondly, the clause specifying when something is red can be thought of as a piece of *reference-fixing* or as piece of *meaning-giving*.[9] If it is a piece of reference-fixing, the question of whether an object is red in some counterfactual world will turn on whether it has redness the way things actually are—that is, has a property which makes things look red in the right way in the actual world. If the clause is a piece of meaning-giving, what matters is what the property does in the counterfactual world—to be red in a world is to have a property that makes things look red in the right way in that world. I suspect that speakers of natural language vacillate between these two readings, depending in part on the persuasive powers of the philosophy tutor they are discussing the issue with. In either case, we can think of our clause as a priori, and its a priori nature constitutes our honouring of the commitment to the relativity of colour to viewers and circumstances.

Thirdly, the primary quality theory has an advantage over the dispositional theory of colour distinct from the causal advantage pressed earlier. For the primary quality theory can handle in a straightforward way a well-known problem for the dispositional theory.

The problem arises from cases where, by virtue of an object's relatively intrinsic nature, it would look a certain colour to persons with normal colour vision in normal circumstances, and yet it does not count as being that colour. There are many fanciful examples in the literature but here is one I owe to David Lewis; it is, by the standards that operate in philosophy, a real-life example. Suppose *O* has a surface property *W* that in itself would cause things to look white in the right way—perhaps the property is one that does the job in normal white paper. Suppose that *O* also has a

the colour of the objects around us. But if there are we would have to, on the appropriate occasions, relativize to one or another human percipient. If Jonathan Bennett, 'Substance, Reality, and Primary Qualities', repr. in C. B. Martin and D. M. Armstrong, eds., *Locke and Berkeley* (New York: Doubleday, 1968), pp. 86–124, is right about phenol-thio-urea, we do need to do this for taste. However, though phenol tastes bitter to about 75 per cent and is tasteless to about 25 per cent of otherwise comparable human tasters, the explanation may be (I understand) that what is being tasted is not phenol itself but a by-product produced only in certain mouths, in which case it is not true that one and the same substance has a dramatically different taste to equally normal tasters. Rather, one and the same substance causes different and different tasting substances in different mouths.

[9] Kripke, *Naming and Necessity*.

property that has no relevant effect on *W* except when *O* is in normal circumstances, but when it is, this 'stand-by' property *S* affects this property of *O*, perhaps by eliminating it or perhaps by modifying its normal action, in such a way that *O* looks black. In short, *O* is a piece of photo-sensitive paper—paper that is white in the dark but turns black on exposure to light.

What makes it true that the paper is white before exposure to light? Not the fact that it looks or would look white before exposure—before exposure it does not look any colour; and not the fact that it would look white were it seen in normal circumstances—as they involve exposure to light, it looks black in those circumstances. True, there is a short time lag before photo-sensitive paper turns black, but it is too short to see. (We may suppose—the example is only real-life by philosophical standards!) And yet clearly the paper is white before it is exposed to light. To say otherwise is to commit oneself implausibly to telling photographers who say that photo-sensitive paper *turns* black—and so was not black to start with—that they are wrong.

The primary quality theorist handles this example by drawing on the fact that there are two properties in play: property *W* of the paper's surface, and the stand-by property which operates very quickly, when normal viewing circumstances arrive. (If *S immediately* eliminates or modifies *W—per impossible*, as causation takes time in the real world—it is no longer intuitive that the paper is white until the normal circumstances arrive; we simply have a case where, though *W* by itself makes something white, the conjunction of *W* and *S* makes something black.) As long as the paper has an unmodified instance of *W*, the primary quality theorist can count it as white, because it has a property that normally disposes things to look white in normal circumstances. Thus, until the 'interfering' takes place, the paper counts as white. You might reasonably urge that this means that until *W* is eliminated or modified, the paper itself can be said to be disposed to look white in normal circumstances.[10] But the key point for us is that the story about why the

[10] Exactly what to say turns on how to handle 'finkish' dispositions—dispositions that tend to go away when the occasion for their manifestation arrives. These cases were raised many years ago by C. B. Martin, though, to my knowledge, he did not publish on the subject until his 'Dispositions and Conditionals', *Philosophical Quarterly*, 44 (1994): 1–8. For a response to the problem of finkish dispositions that would count the paper itself as disposed to look white until *W* is

paper counts as white until normal viewing circumstances arrive, turns on the role of the causal basis of the disposition, not on the question of whether the paper itself would look white in those normal viewing circumstances.

Fourthly, the primary quality theory can, as we said at the beginning, honour the dispositionalist's insight that there is something a priori, or somehow truistic, about the connection between being red and being appropriately disposed to look red.

Although the theory identifies colours with physical properties and so makes them objective and observer-independent, it is not an objective, observer-independent matter which physical properties (if any) are which colours. The basic idea can be illustrated with the example of the most dangerous chemical structure for humans. This structure is an objective, observer-independent property. For instance, on some ways of measuring toxicity it is, I understand, the structure of plutonium, and the structure of plutonium is an objective, observer-independent property. Nevertheless, what makes it true that plutonium is the most dangerous substance is of course a highly relative matter. It concerns the effect that plutonium has on humans, and that is in part a function of how humans are made. Likewise, on the causal theory of colour, which physical properties (if any) are which colours is an observer-dependent matter. It turns on whether the physical properties or property complexes in question have the right kinds of causal effects in the right kinds of ways on normal observers in normal circumstances to count as being presented in experience when things look one or another colour. David Hilbert has a good name for this kind of theory. He calls it *anthropocentric* realism.[11] The colours *per se* are observer-independent properties, but *which* observer-independent properties they are is not observer-independent.

What has masked the possibility of this kind of theory is the tendency to define the notion of a dispositional property in terms of the a priori nature of the relevant biconditional; to say, roughly, that Φ is a dispositional property iff some such biconditional as '*x* is Φ iff *x* is of a nature such that *x* does such-and-such in so-and-so

eliminated or modified, see David Lewis, 'Finkish Dispositions', *Philosophical Quarterly*, 47 (1997), 143–58.

[11] David R. Hilbert, *Color and Color Perception: A Study in Anthropocentric Realism* (Stanford, Calif.: CSLI, 1987).

circumstances' is a priori.[12] But it is a priori that *x* has the most poisonous structure if and only if (roughly) *x* has a nature such that ingesting *x* has certain effects. Nevertheless, the most poisonous structure is not a disposition. It is the structure of plutonium.

What makes a property a disposition is that it itself is *essentially* linked to the production of certain results in certain circumstances, not whether some open sentence concerning it is a priori. And it is indeed a consequence of the causal theory that redness, for instance, is not essentially linked to looking red. Not just because of the possibility of 'defeaters', but because the, or any, property that typically makes things look red might fail to do so in some other world, just as the structure of plutonium might have been harmless to humans. In my role as a fence-sitter on whether the relevant causal roles (in part) fix the reference or give the meaning of the colour terms, I say nothing about whether things with these properties count as red in these worlds; what is clear and what matters for us is that the very properties that make things look red might not have.

Finally, I should note that the primary quality cum causal theory as presented here ducks an important issue. It refers to colour experiences under their colour-experience names, it says nothing illuminating about how to understand colour experience. Once upon a time I was convinced that any adequate account of colour experiences required reference to *qualia* understood as properties over and above those that appear in the physicalists' story about our world. Nowadays I am much more sympathetic to physicalism.

OBJECTIONS TO THE PRIMARY QUALITY THEORY

The primary quality theory of colour is built on the folk axiom that colours are the properties putatively presented in the experience of things looking coloured. The obvious question to ask then is whether there are other claims that are equally part of the folk theory of colour, and which, in one way or another, undermine the view that colours are physical properties. As I said at the beginning of this chapter, there is an important sense in which we know

[12] Roughly—in view of the Martin point referred to in n. 10.

all the possibilities as far as colour is concerned—we know what the possibly relevant properties are, and we know how to name them—and the issue that remains is—to say it in Lewis-speak—which of the possibly relevant properties deserve the names of the colours in addition to the names they already have. And this is a question that can only be settled by consulting the folk theory of colour.

It has variously been suggested that the primary quality theory conflicts with (at least) three central tenets of the folk theory: the first is variously known as transparency or revelation, the idea that our experience of colour reveals its essential nature; the second is, in Keith Campbell's words, the axioms of unity;[13] and the third is the doctrine that different colours are strongly incompatible. The rest of this chapter will be mainly concerned with the first two suggested folk constraints on colour, and especially with whether they constitute objections to identifying the colours with physical properties. I will though say a little about strong incompatibility.

The Objection from Revelation

If colours are physical properties, it must be conceded that the way they look does not reveal their essential nature. When something looks red, it does not look one or another physical quality (or complex of physical qualities). You cannot see 'through' the experience to the nature of what is being experienced. Thus, if it is part of folk theory that the experience of colour reveals in itself the nature of colour, that colour is transparent in this sense, the primary property view must be false. And a number of philosophers have indeed suggested that it is part of the folk theory of colour that colour experience is transparent in the sense of revealing the essential nature of colour. For instance, Galen Strawson says that 'color 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, phenomenal-quality experience, given only the qualitative character that that sensory experience has*'.[14] If

[13] Keith Campbell, 'Colours', in Robert Brown and C. D. Rollins, eds., *Contemporary Philosophy in Australia* (London: Allen & Unwin, 1969), 132–57.

[14] Galen Strawson, '"Red" and Red', *Synthese*, 78 (1989): 193–232, at 224, author's italics. Revelation is defended under the name of transparency by John

Strawson is right, colours, or at least colours as the folk conceive them, are not physical properties.

But is revelation really part of the folk theory of colour? There seem to me three reasons for denying that it is. First, it is hard to believe that our experience of colour is that different from our experience of heat. Perhaps before we had any idea of what heat was, some were tempted to say that sensations of heat revealed the full nature of heat, that heat is precisely that which is fully transparent to us when something feels hot. After all, that it feels hot was the main thing most people knew about heat, just as the main thing that is currently common knowledge about redness is that it makes things look red. However, our very preparedness to identify heat with molecular kinetic energy when the empirical evidence came in shows that this opinion was merely opinion. We did not hesitate to identify heat with something whose full nature is manifestly not given to us in the experience of heat.

Secondly, the folk allow that we can misperceive colour, that colour illusion is possible. But that is to draw a distinction between colours as they really are and colours as they appear to be, and that is to concede that the colours have a nature that outruns our experience of them.[15]

Finally, the prime intuition requires treating our experience of colour as typically caused by colour, and it is part of the folk notion of causation that causes and effects are distinct. But if our experience of colour is distinct from what it is an experience of, how could it transparently reveal the nature of colour? The folk thus know something about colour that tells them that revelation could not possibly be true. Of course, this last argument has force only if—unlike Mackie—we work on the general presumption that the folk are not badly confused.[16] If we incline to the view that the

Campbell, 'A Simple View of Colour', in John Haldane and Crispin Wright, eds., *Reality, Representation, and Projection* (New York: Oxford University Press, 1993), 257–68.

[15] A point made by Michael Smith, 'Colour, Transparency, Mind-Independence', in John Haldane and Crispin Wright, eds., *Reality, Representation, and Projection* (New York: Oxford University Press, 1993), 269–77.

[16] I have in mind Mackie's tendency to favour error theories that attribute to the folk seriously erroneous conceptions. See J. L. Mackie, *Ethics* (Harmondsworth: Penguin, 1977), for his error theory of (folk) value, and J. L. Mackie, *Problems from Locke* (Oxford: Clarendon Press, 1976), for his error theory of (folk) colour.

folk often are badly confused, a proponent of revelation can reply that here is an illustration of this very tendency to be confused. I think that the folk are smarter than that, but if you are of Mackie's mind, you can think of the last point as telling us how to restore consistency to the folk conception of colour: the way to do it is to drop revelation.

It might be suggested that although we should reject revelation, we should, nevertheless, try for a theory of colour that respects it as much as possible. Thus, Mark Johnston argues that the major advantage of a dispositional theory of colour over a primary quality theory—be it of our causalist variety or not—is that it gives enough to revelation to avoid sceptical worries that any primary quality theory necessarily engenders. He argues that the dispositional theory of colour secures an important cognitive value that the primary quality theory denies.

> Vision can be a mode of revelation of the nature of visual response-dispositions. It cannot be a mode of revelation of the properties that the Primary Quality Theorist identifies with the colors. Since we are inevitably in the business of refiguring our inconsistent color concepts, we should make the revision which allows us to secure an important *cognitive value*—the value of acquaintance with those salient, striking and ubiquitous features that are the colors.
>
> The point here is not simply that the Primary Quality Account does not satisfy even a qualified form of Revelation. What is more crucial is that as a result, the account does not provide for something we very much value: acquaintance with the colors. The ultimate defect of the Primary Quality View is therefore a *practical* one. From the point of view of what we might call the ethics of perception, the Secondary Quality Account is to be preferred. It provides for acquaintance with the colors.[17]

I think that this misunderstands the nature of the issue between primary quality cum causal theories and dispositional theories. There is, as we emphasized before, no deep metaphysical dispute between primary quality theorists and dispositionalists. The dispute is over whether the dispositions to look coloured or the physical quality bases of those dispositions should be tagged as the colours; the dispute is ultimately over the distribution of names among putative candidates. And how we answer this *labelling* question can have no cognitive, epistemic or practical significance.

[17] Johnston, 'How to Talk of the Colors', 258, my emphases.

If we reject revelation, we must reject the view that different colours are strongly incompatible in the sense of its being part of our very concept of different colours that they are *essentially* incompatible. *If* it is a priori that no object is red and green (all over, for a given *S* at a given time, and in a given circumstance), it will be because it is a priori that what is required by way of action (on *S* etc.) for an object to count as red all over (for *S* etc.) cannot coexist with what is required by way of action (on *S* etc.) for an object to count as green all over (for *S* etc.). It will be like the impossibility of a substance being both poisonous and harmless to the very same population in the very same circumstances. But this is consistent with red and green themselves being compatible (though if they were ever together in an object, it would be wrong to call them 'red' and 'green'). What is ruled out by the denial of revelation is that it is a priori that the properties themselves are essentially incompatible, for that would require embracing some form of revelation into the essential nature of the colours. It may, of course, be a posteriori that red and green are essentially incompatible, but this is something primary quality theorists can happily accept. They can allow that it may turn out that the physical properties identical with red and green are mutually exclusive, as would be the case if one is having a 'grain' greater than *x* and the other is having a 'grain' less than *x*.

The Objection from the Axioms of Unity

The axioms of unity say that redness is the property common to all red things; that blueness is the property common to all blue things; and so on and so forth. If (a) the axioms of unity are a central part of the folk theory of colour, (b) a certain view about causation by disjunctive properties is correct, and (c) a certain empirical claim is correct, then the primary quality view is false.[18] Let me spell all this out.

Disjunctive properties can be causes. For instance, Tom's being taller than Dick may cause Tom to be chosen for the basketball team ahead of Dick. Equally, Tom's living next door to Dick may be the cause of Tom's knowing a lot about Dick. In both cases the cited cause can be thought of as disjunctive in the sense that it can

[18] I am indebted to Michael Watkins for pressing me on this point.

be realized in many different ways. Tom's being taller than Dick is a matter of Tom's being 200 cm and Dick's being 199 cm, or Tom's being 199 cm and Dick's being 198 cm, or . . .; and living next door can be thought of as a disjunction of the many significantly different ways of living next door. Indeed, it is arguable that most things we cite as causes are more or less disjunctive. When we cite the depth of the wound as responsible for the death of the victim, it is typically not the absolutely precise depth of the wound that matters but rather the fact that the wound's depth falls within a certain range of depths, any of which counts as deep. Nevertheless, excessively disjunctive properties cannot be causes. Indeed, we cannot even say that they are causes, properly speaking. Consider, for instance, the sentence 'Either arsenic administered by Harry or cyanide administered by Mary caused the death'. Surely we only make sense of this sentence by reading it as 'Either arsenic administered by Harry caused the death or cyanide administered by Mary caused the death'. When we are confronted with a claim that appears on the surface to cite an excessively disjunctive property as a cause, we make sense of it by reading the claim as one about one or another of the disparate disjuncts being the cause.

Now consider an example of Johnston's. Let us suppose that what makes a canary look yellow is a different property, P_1, from the property, P_2, that makes the relevant section of a colour photograph of the canary look yellow.[19] What should primary quality theorists identify as yellowness? The axioms of unity imply that they cannot say that P_1 is yellowness in the bird, whereas P_2 is yellowness in the area on the photograph. They must rather say that yellowness in both bird and photograph is the shared disjunction P_1 or P_2—or more generally that yellowness is the disjunction of all the physical property complexes that make things look yellow in the right way, but we will suppose that the disjunction of P_1 with P_2 covers all the cases.

Finally, suppose that P_1 and P_2 are very different, so different that the disjunction P_1 or P_2 counts as excessively disjunctive. Now the causal theory is in trouble. For it is built on the intuition that yellowness is what causes things to look yellow, and so cannot afford to identify yellowness with an excessively disjunctive property.

[19] Johnston, 'How to Speak of the Colors'.

How should we reply to the objection from the axioms of unity? We might follow Saul Kripke's lead and think of the colours as kinds. We might think of the word 'red' as denoting the kind *K* that a good number of exemplars of red things share and which causes them to look red (in the right way).[20] We then declare anything which is *K*, whether or not it looks red, to be red, and declare things which are not *K* but look red in normal circumstances to be 'fool's red'. Thus, if Johnston is right about the difference between what makes a canary look yellow and what makes the colour photograph of a canary look yellow, at least one of the canary and the photograph is fool's yellow. This approach might or might not be combined with the view that the colour terms are rigid designators. That is, we might understand

> 'red' denotes the (causally relevant) kind common to the red-looking exemplars of red things

as giving the meaning or as fixing the reference of 'red'. On the first understanding, the denotation specification applies world by world. The red things in a world *w* are the things that belong to the kind common to the red-looking exemplars of red things in *w*. But the red things in one world need not belong to the same kind as the red things in some other world. On the second understanding, 'red' is a rigid designator. The red things in a world *w* are the things that belong to the kind common to the red-looking exemplars of red things in the actual world, and so 'red' will denote the same kind in every world. (The latter is, I take it, what Kripke had in mind.)

I do not think that either version of the kind view is part of the folk theory of colour. Whether or not it turns out that there is some feature common to most things that look red, or most things that are, for whatever reason, counted as the exemplars of red things, a feature of sufficient note to count as marking out a kind which explains their looking red, I do not think our talk about red in any way presupposes that there is.

[20] See Kripke, *Naming and Necessity*, 128 n. 66 and 140 n. 71. Kripke's view is sometimes reported as that the colour terms mark out *natural* kinds. However, as Graham Oppy convinced me, it is not clear that he wants (or wanted at that time) to hold that all yellow things, say, have in common something significant enough to be regarded as collecting them into a natural kind.

In the case of terms like 'water' and 'gold' it is plausible that we take it for granted that there is something important that might be properly regarded as a kind, indeed a natural kind, distinctive of the exemplars of water and gold. As a result, the contention that it is part of their meaning that they denote kinds is plausible. But the diversity of kinds of things that look red—sunsets, ripe tomatoes, blood, feathers—along with the notorious variability of apparent colour, facts with which the folk have been long familiar, predispose the folk to expect that there may well not be any single kind distinctive of the things we use the word 'red' for. In short, the folk are too sensible to have presupposed something as risky as that there is a distinctive kind in common to things we call 'red'.

This is consistent with allowing that we might, after the event, give kind membership an important role in determining colour. Suppose it turns out that most of the things that look red to normal perceivers in normal circumstances do so because of some commonality that we may reasonably think of as marking out a kind. Then we might say that other things that look just as red to normal perceivers, in circumstances equally regardable as normal, but which are not of the kind in question, are fool's red.

The best reply to the objection from the axioms of unity is, I suggest, to urge that the disjunction is not excessively disjunctive. Even if most red things do not belong to a kind responsible for them normally looking red, there will turn out to be, all the same, sufficient similarity between what typically makes things look red to allow us to identify red with a disjunctive property that is sufficiently unified to count as a cause. For it is hard to believe that there is not enough rhyme or reason to things looking red given the evolutionary importance of colour vision, the role of colour difference in the detection of shape, the phenomenon of colour constancy (the fact that apparent colour is relatively invariant under changes in intensity of illumination), and the phenomenon of colour stability (the apparent colour of things in a given circumstance is fairly constant over time) to unify the disjunction. It makes good empirical sense that something physically interesting (which may well not have the status of marking out a kind except under extremely relaxed standards for kind-hood) unifies the various red-looking things over and above their being red-looking, and that colour vision is there in order to enable us to process this information, and that the same is true for the other colours.

It is sometimes thought an insuperable difficulty for this claim that we know that the physical nature of the light entering our eyes from objects that look the same colour varies greatly, and yet this physical nature is the relatively immediate cause of how the objects look. There is, it is said, no rhyme or reason to be found in the physical causes of one and the same colour judgement. For example, C. L. Hardin observes that 'apart from their radiative *result*, there is nothing that blue things have in common . . .'.[21] But consider an analogy pressed by Hilbert.[22] He points out that quite different factors are involved in our being able to see how far away things are. A major one is the information that comes from the fact that we have binocular vision, but you can still tell how far away things are with one eye closed or after losing the sight of one eye. This means that the very same judgement of visual depth may be driven by very different properties of the light that enters our eyes (and, if it comes to that, of the light that leaves the object). But it would be wrong to think that there is a disjunction problem here. The disparity in the nature of the light that enables us to make some given judgement of depth is irrelevant. What is relevant is the fact that there is a unifying distal property of the objects, namely, how far away they are, which our visual system disentangles from the otherwise disparate nature of the light it receives.

The issue then in the case of colour is whether there is a unifying distal property. Now there is some reason to hold that triples of integrated reflectances correlate closely with perceived colour. The fine detail is not important here, and, needless to say, it is controversial. But roughly a triple of integrated reflectances is the result of taking the reflectance—that is, certain proportions of reflected light to incident light—over three band-widths, scaling, and then summing. The result correlates closely with the apparent colour of reflecting surfaces.

What is more, these triples capture the similarity relations that are part of the folk theory of colour. The triple for orange, for instance, is closer to the triple for red than it is to the triple for blue. Hilbert infers that we should identify the colours with the relevant

[21] C. L. Hardin, 'Are "Scientific" Objects Coloured?', *Mind*, 93 (1984): 491–500, my emphasis.

[22] David R. Hilbert, 'What is Color Vision?', *Philosophical Studies*, 68 (1992): 351–70.

values of these triples, but here we have to be careful. Hilbert, as I understand him, thinks of the triples as themselves dispositional properties—as an object's disposition to reflect light displaying the relevant value of the triple. This is how he allows objects to have colours in the dark, and how he avoids having to say that light *creates* an object's colour rather than, as we folk want to say, *revealing* its colour. (There is no *actual* value of interest for the triple for an object in the dark.) But I cannot follow him in identifying the colours with these dispositions. I have to think of the value of the triple for a given colour, red, say, as what unifies the possibly highly disjunctive basis that is responsible for the disposition to look red in normal circumstances. It is what prevents the basis counting as excessively disjunctive. David Braddon-Mitchell drew my attention to a nice example here. Vitamins are a pretty heterogenous lot, but vitamin deficiency counts as a cause because there is a unity in the way lack of a vitamin acts on us. In the same way we should say that the reflectance triple story is one about how the possibly highly heterogeneous bases of the disposition to look red in different objects form a sufficiently unified disjunction to count as the normal cause of looking red.[23]

A nice feature of seeing the unity in causes as a matter of unity in the triples is that it squares with our allowing that the yellow-and-magenta-dotted look of an area of a colour magazine seen close up and its red-look seen from a normal viewing distance should *both* be thought of as revealing colour: one reveals the colour from close up, the other the colour from a normal viewing distance. For a triple of integrated reflectances is a holistic property of an area—an area as a whole may have a different triple value from that possessed by some sub-area. Thus, we can maintain that we are latching onto a physical property of the area when we view from a normal distance, and a different physical property of a sub-area of the larger area when we view from close up, because the categorical basis underlying the triple of integrated reflectances for an area will not in general be the same as that underlying the integrated triple of reflectances for a sub-area of the area.[24]

[23] I am much indebted here to comments by David Lewis, David Braddon-Mitchell, and Ian Gold on earlier fumblings with this point.

[24] Essentially the same account applies to the difference between blood seen

It is, however, unlikely that these possibly disjunctive bases will reflect the similarities and differences among the colours in the way that Hilbert's identifications arguably do. There is no reason to think the physical property we are latching onto when some particular thing looks red is similar to that we are latching onto when some particular thing looks pink, for example. This looks like trouble. For it is plausible that colour experience, in addition to representing objects as having properties which are causally responsible for these objects looking coloured, also represents these properties as occupying certain places in the three-dimensional colour array (red is opposite green, orange is nearer red than green, etc.). I think, though, that we need to ask: In what sense does, for instance, looking red represent objects as having a property more like the property looking orange represents them as having than does looking green; in what sense is orange as represented in experience more like red as represented in experience than it is like green as represented in experience?

A clearly wrong answer would be to say that it is somehow 'more' true or more obvious that orange is a different colour from green than that it is a different colour from red. It is certainly true and completely obvious *both* that red is different from orange and that red is different from green. The only alternative seems to be to borrow, in one form or another, from behavioural psychology by analysing the needed sense in terms of *jnds* (just noticeable differences). Roughly, the sense in which orange is closer to red than it is to green lies in the fact that it takes more jnds to get from orange of a given saturation to green of the same saturation than to get to red of the same saturation. But in *that* sense, or anything roughly like it, the physical properties do stand in the right similarity relationships. They induce the relevant behavioural relationships. More generally, the point is that if we can, as seems plausible, understand the three-dimensional array, the colour solid, in terms of suitably scaled jnds, then the nature of the array will not be trouble for the primary quality view.

However, none of this means that I can duck the question of what to say if it turns out that although there are some underlying

with the naked eye and blood seen through a microscope. See Hilbert, 'What is Color Vision?'.

unities among the objects that typically look red—it would be incredible if there weren't—there is no single principle unifying them. The reflectance triples story, let us say, turns out to have major holes. Perhaps the red-looking objects naturally divide into two groups: in one group the categorical basis for looking red in the right way is one kind of structure S_1, and for the other it is some quite different structure S_2, and there is no way, in terms of reflectance triples or whatever, of seeing any sort of unity here.

In thinking about this case we should remember the example of jade. Jade, it turned out, comes in two quite different forms (nephrite and jadeite), but this did not lead us to deny the existence of jade. It led us to say there are two kinds where we might have thought that there was only one. Likewise, if it turns out that there is no way of treating what makes tomatoes look red and what makes sunsets look red as different manifestations of some disjunctive but not excessively disjunctive common feature, we should say that the red of sunsets is a different property from the red of tomatoes just as New Zealand jade is a different kind from Chinese jade (though the two reds will occupy the same spot in the colour solid, of course). We should, that is, modify the axioms of unity. Redness is not the property in common to red things. Rather there are two rednesses, and red things have one or other of the two rednesses. I think that the folk would happily say this, and so that folk theory implicitly allows us to modify the axioms of unity. Indeed, I think that we could live with considerably more than two rednesses. What would be intolerable would be if it turns out that there are no interesting distinctive distal commonalities underlying similarities of apparent colour. For then what would be called for is not some more or less radical modification of the axioms of unity, but a total abandonment of them. If this turned out to be the case, I think that we would have to declare colour a pervasive illusion. Nothing is coloured, just as nothing has impetus in the sense given to it in medieval physics. Certain things appear to have impetus, which is how medieval physics made its mistake, but nothing really has it. We would have to say the same for colour.

The next two chapters are concerned with the location problem for ethics.