Black and White and the Inverted Spectrum

Justin Broackes

To the traditional idea of an undetectable Spectrum Inversion some have added the idea of Inverted Earth. Perhaps ironically, the new combination of ideas is even harder to make coherent—particularly if we try to imagine an inversion of experience of Black and White being supposedly counteracted by an environmental switch of white things for black and vice versa. White and black exhibit numerous asymmetries—in their connections with illumination, shadow, and visibility—that rule out their being reversed. And, since the most saturated yellow is light and the most saturated blue dark, yellow and blue could not be reversed unless light and dark could be—which they cannot. (A similar problem befalls red-green inversion: the best reds are dark, but the best greens are of medium lightness.) The difficulties suggest some more general morals for how we might think of the role of ‘qualia’ in colour perception.

‘Would it be conceivable for someone to see as black everything that we see as white, and vice versa?’ Wittgenstein

The familiar suggestion that one person might see yellow where another saw blue, without there being any detectable difference in what they said or did, is usually traced to Locke’s Essay (2.32.15). The basic idea can be found, I think, about two thousand years earlier, in the Cyrenaics, the ‘minor Socratic’ school that defended a form of philosophical hedonism, and which flourished around 330-270 BCE in what had been the Greek colony of Cyrene in North Africa:

For just as [—to consider cases of jaundice and ophthalmia and the normal—] some people are affected yellowly, and others crimsonly, and others whitely, from one object, so also it is likely that those who are in a normal condition will not be affected similarly from the same things, owing to the differing construction of their senses, but the pale-eyed will be affected in one way, the blue-eyed in another, and the dark-eyed in a different way. So that we set up names that are common for the facts, but we have affections that are peculiar to each of us.

---

2 Sextus Empiricus, Adversus mathematicos VII.197-198 ( = Adversus dogmaticos (Against the Logicians) I.197-198). Note the adverbial theory of perception.
In some sense the problem goes back to Plato’s *Theaetetus* and, beyond that, to the issue of conflicting appearances in Protagoras; but what is remarkable here, and characteristic for the later debate, is the idea that experiential variation is compatible with complete agreement in using the same words regularly for the same things.

But is that idea in fact coherent? A nice way to examine the question is in terms of ‘Inverted Earth’—a place where grass is red and fire-engines green—an idea developed by Ned Block, following Gilbert Harman. We are to suppose a subject having her colour-experience ‘inverted’, and then being transported to Inverted Earth where, supposedly, the environmental changes would counteract the experiential changes—and the person end up being unaware of any change at all. Block claims advantages for the Inverted Earth story over the simple inverted spectrum story. I shall argue, however, that the advantages are not what they seem; but the new story has one virtue—that of revealing vividly, if unintentionally, the difficulties of making either the old or the new suggestion coherent in detail.

1. *Spectrum inversion and Inverted Earth.* The core of the spectrum inversion story, in one version, is this. At stage 1, $x$ and $y$ are speakers of English with apparently normal colour vision; at stage 2, we place ‘color-inverting lenses’ in $y$.

---

3 On which and much else, see Miles Burnyeat’s wonderful lecture ‘Conflicting Appearances’, *Proceedings of the British Academy* 65 (1979), pp. 69-111.

4 I follow the main lines of the presentations of Sydney Shoemaker and Ned Block. See S. Shoemaker, ‘The inverted spectrum’, *Journal of Philosophy* 79 (1982), pp. 357-81, and N. Block, ‘Inverted Earth’, *Philosophical Perspectives* 4 (1990), pp. 53-79, both repr. in N. Block, O. Flanagan & G. Güzeldere, *The Nature of Consciousness* (Cambridge, Mass.: MIT Press, 1997), pp. 643-62 and 677-693. Page numbers in parentheses in the main text (accompanied by ‘a’ or ‘b’ to indicate column one and column two) refer to Shoemaker’s and Block’s articles as reprinted in this collection. I shall follow the common usage of Peacocke’s term ‘red’ (and similarly, ‘green’, etc.), as intended to apply to the kind of experience produced when a red object is presented to a normal observer under normal circumstances. (Peacocke himself prefers to talk of red ‘regions of the visual field, rather than of red’ experiences.) See C. Peacocke *Sense and Content* (Oxford: Oxford University Press, 1983), ch. 1. To simplify the presentation, I shall assume that at stage 1, red things produce red’ experiences in ordinary Earthlings, and produce red’ experiences in ordinary Inverted Earthlings. Block does not assume such uniformity among ordinary people—indeed, he is rather inclined to use the inversion hypotheses to cast doubt on it. But the assumption does not affect the core of the argument.
Red things now produce green’ experiences in him, and he is initially tempted to call them ‘green’. At stage 3, y adapts to the situation, and comes to call red things (which now produce green’ experiences in him) ‘red’ after all. At stage 4, ‘he has amnesia about the period before the lenses were inserted’, and thus finally becomes ‘functionally totally normal’ (681b). At this final stage, when x and y both look at a red thing, they are in qualitatively different but functionally identical states.

The Inverted Earth story is similar.5 (I shall concentrate on the ‘intrapersonal’ case, where the Inverted Earth story has the supposed advantages over traditional Spectrum Inversion.6) At stage 1, Earthlings x and y have apparently normal colour vision. At stage 2, we place ‘color-inverting lenses’ in y. But we also transport him to Inverted Earth, where grass is red and fire-engines green (taking care also to change his body pigments, 683a). ‘The inverting lenses cancel out the inverted colors’ (683a), so he notices no oddity in himself or his environment. Green things (like fire-engines) now produce red’ experiences in him, and (following what seem old habits) he calls them ‘red’. Since the local inhabitants speak Inverted English, in which red things are called ‘green’ and green things ‘red’, y sounds like a local, and notices no mismatch between his classifications and theirs. At stage 3, differences in y’s linguistic background continue to influence the reference of his terms. (He says ‘red’ of green things like the locals; but he still means red by the term, whereas they mean green.) It is perhaps only after ‘50 years have passed’, at stage 4, that his ‘intentional contents

6 The application of these terms isn’t ideal. It makes good sense to distinguish intrapersonal and interpersonal comparisons; but dividing ‘cases’ or ‘inversions’ into the intrapersonal and the interpersonal is not so clear. An ‘intrapersonal’ inversion occurs when x, accustomed to seeing colours one way, comes to see them another way. An ‘interpersonal’ inversion involves a difference in how x and y see colours, and writers often reserve the term for cases where this is supposed to have involved no change within x or within y, but a difference between them since birth. But, of course, in a case of ‘intrapersonal’ inversion, there will as a result (if talk of interpersonal identity and difference in experience makes sense at all) also be a difference (interpersonal) between x and others who previously had the same kind of sensory system as her. Block does also describe an ‘interpersonal’ Inverted Spectrum story (where inverting lenses are inserted into one of two twins at birth), and an interpersonal Inverted Earth story (where the affected twin is also transported to Inverted Earth); but he mentions no special advantages for the latter over the former.
would shift’ (683b)—and he finally comes to mean what the Inverted Earthlings do, as well as sounding like them. At this final stage, therefore, when \(x\), the unchanged Earthling, looks at a red fire engine, and \(y\) looks at a green one, they are in qualitatively identical states. (Both have red’ experiences.) But they are ‘functionally and intentionally’ in different states (cp. 684b-685a). \(x\) is in a state standardly produced by red things and leading (among other things) to predications of the term ‘red’ of them; \(y\) is in a state standardly produced by green things and leading to predications of the term ‘red’ of them. On \(x\)'s lips, ‘red’ means red, on \(y\)'s lips, ‘red’ now means green. Hence Block’s conclusion: ‘the functional and intentional theory of qualia is refuted’ (685b).

The Inverted Earth story is similar to its ancestor: in both, \(y\) receives the spectrum-inverting lenses; the central difference is that, by being taken to Inverted Earth, \(y\) is prevented from discovering the transformation. There is a further difference, of less importance. In the Spectrum Inversion story, \(y\)'s linguistic practice adapts as he advances from stage 2 to stage 3, but only through some disruption. (He has to start calling things that produce green’ experiences ‘red’, thus apparently regaining his status as a normal speaker of English.) By contrast, the second story involves linguistic change without apparent disruption: merely living in what seems an unchanged environment, \(y\) supposedly comes to use the word ‘red’ for green—and thereby changes from speaking English to speaking its Inverted equivalent.\(^7\)

Block claims an advantage for the second story. The first scenario, as described, is liable to the challenge (that he associates with Dennett and, in a different form, with Georges Rey) that we should not ‘tak[e] the subject’s memory reports at stage 3 at face value’ (682a). The subject may claim that grass now looks in colour as fire-engines used to look—but one might wonder if the subject was simply confused about how he saw things in the past. Block concedes: ‘there is some justice in these complaints (though I disagree with them)’ (682a). The Inverted Earth scenario avoids ‘this weakness’ (684a): for, in former case, ‘the subject’s internal disturbance renders his first-person reports vulnerable to doubt’, whereas in the latter case the subject has ‘no internal disturbance’, so there is no cause to suspect his memory reports (684a).

I am not sure about the disadvantage of the former story. Dennett’s worry about memory is in origin a worry about claims of experiential change

\(^7\) For these reasons, my labelling of stages, which secures parallelism in certain respects, exposes dissimilarities in other respects. There is no real change between stages 2 and 3 in the second story. Perhaps the second story would be strengthened by introducing amnesia at stage 4 too, so that \(y\) is not distracted by any memories from before the change.
unconnected in any way with changes in linguistic or behavioural practice. (Cp. Wittgenstein: ‘Always get rid of the idea of a private object in this way: assume that it constantly changes, but that you do not notice because your memory constantly deceives you.’ By extension, one might (though Wittgenstein does not) worry about the interpretation of a case where a subject woke up—apparently unchanged from the night before—and merely began to claim that grass now looked red and fire-engines green. (This is the kind of worry Shoemaker mentions and insures against by making his favoured inversion occur by three partial inversions of just a third of the experienced spectrum (645b).) But neither Wittgenstein nor Dennett nor Shoemaker would credit a sceptical concern about memory in a case where the subject has just undergone a surgical intervention involving the insertion of ‘color-inverting lenses’ (or even a ‘neural converter’, ‘hooked into the optic nerve behind the retina’ which ‘changes the signals that represent red to signals that represent green, and so forth’, 688a, cp. 692a n.17). Such intervention in the sensory system would only be expected to result in a change in experience—just as much as the putting-on of an external apparatus of colour-inverting TV screens—and the subject’s behaviour, linguistic and non-linguistic, would be perfectly capable of confirming the change.

The memory worry about the first scenario is therefore, I think, illusory. If it were well-motivated, then it would, I think, apply equally to the second scenario: if we had reason to think that y, with the insertion of the mystery device, merely underwent a memory confusion, then taking him to Inverted Earth where such confusion would remain unnoticed, would hardly quell the worry. Memory worries, as I have said, have more bite with respect to a case where y wakes up one morning after no intervention, claiming colour-experience inversion; but those worries have equal bite if we take y on a flight to Inverted Earth and find that, still with no intervention, he reports that everything looks normal to him. The very fact of his lack of surprise would for us be a sign that something strange had happened to him quite apart from the interplanetary trip we’d all taken—though what, would remain to be investigated. Block talks as if the memory worry were fuelled by ‘the subject’s internal disturbance’ or ‘confusion’ (684a), which is present only if the subject remains on Earth. But that is not the case: the existence of some important change is evident to theorists from the subject’s behaviour,

whether he is on Earth or on Inverted Earth; the subject’s *disquiet*, if he is on Earth, merely confirms the change, as does his *lack of disquiet* if he is on Inverted Earth; and in neither case does it settle the interpretation of the disturbance—though any sensible friend of Wittgenstein’s would in the versions Block describes credit the inversion-of-colour-experience explanation rather than the disturbance-of-memory rival.

Block, like others before him, is keen to use the inversions as counter-examples to functionalism. I am not sure they work as such, though I shall here raise only a quick question about the use of them against the idea that ‘long-arm’ functional roles can uniquely characterize qualia. The standard spectrum inversion story is supposed to provide a counterexample: $x$ and $y$ at stage 4, when looking at a red thing, are supposed to be in states that are qualitatively different, but functionally identical. And Block takes the Inverted Earth example to provide another counterexample: when $x$ and $y$ look respectively at a red and a green thing, they are in states that are qualitatively identical, but functionally different.

Neither counter-example is compelling. In the former case, amnesia is supposed to ensure that $y$ at stage 4 becomes functionally identical with $x$ after all. But there are certain input-output relations that would still tell $x$ and $y$ apart. Tell $y$ the whole truth about his past, for example, and he could reasonably become anxious about what he really means by ‘red’; tell $x$ the whole truth, and he could *not* reasonably become anxious about what he really means by the term. There are further questions one might raise about whether amnesia can really ensure a complete linguistic break with the past; but I shall save them for now.

In the inverted earth story, there is no apparent shock to the subject, and no amnesia. If anything, this makes less persuasive the claim that $y$’s terms will change their reference (so he becomes ‘intentionally inverted’). My conception of red (like my concepts of natural kinds, though in different ways) includes something like a demonstrative element. Things are what we mean by red, I believe, (though I do not intend this as a definition) if they look red to what are actually normal observers in what are actually normal conditions in our home environment. ‘Home’ for most of us is, in this respect, a large place—extending across the Earth and beyond, during all of our lives, and long before. But one can change one’s home, and come to use new colour concepts, referring to different properties. If after insertion of ‘colour-inverting lenses’ $y$ is now to use the term ‘red’ for things that look red to what are—in his new condition—actually normal observers, in what are—in his new condition—actually normal conditions, then he must, so to speak, change his home. But if there is no apparent cut-off point between the old home and the new, there is no way for $y$ to hive off the period before the change. And without such a division, his later thoughts will be bound
to be infected with conceptions and information from earlier times. ‘This is red’, he thinks, of a green fire-engine after the change; ‘it’s the same colour as the engines we used to have when I was a boy’. And nothing clearly divides this into linguistic ‘understanding’ and factual ‘background’. As causal theories of reference remind us, in the absence of a firm reanchoring, the original application of a term will continue to exercise an influence.\footnote{Block later modifies the Inverted Earth thought-experiment (following a suggestion of Bob Stalnaker’s) to include, after all, awareness of the move. (See N. Block, ‘Mental Paint and Mental Latex’, in E. Villanueva (ed), \textit{Perception (Philosophical Issues, 7)} (Atascadero, CA: Ridgeview, 1996), pp. 19–49.) The modified version avoids the present objection; it also for the same reason, however, abandons the alleged advantage of the Inverted Earth story over the original Inverted Spectrum.}

2. \textit{Could you see white where I see black and vice versa?} The deep question is whether an undetectable inversion of the colours is possible at all. I shall offer a two-stage argument: first, that a suitable inversion of white and black is impossible; secondly, that if an inversion of white and black is impossible, then so is the intended inversion of red and green, and yellow and blue.

Suppose we insert a white/black inverter into \(y\)’s visual system: white things now look black to her, and \textit{vice versa}. We take her to Inverted Earth (or rather, a modification of that place): things that are normally black on Earth are white there, and \textit{vice versa}; and the light and dark shades of all the colours are correspondingly inverted. Asphalt (or its substitute) is white, paper black, and Coca Cola advertises not in dark red but in light.

There are difficulties with the scenario that derive from everyday facts about black and white and their relation to light. White things reflect a fair amount of incident light, and black things do not; and the differences on Inverted Earth will show up even to a person with an inverter in her visual system. White things will not look as black things used to look, nor black things look as white. The result is that transportation will not cancel out the oddity caused by the ‘inverter’ in the visual system.

That is the argument in outline; now for some details. Some kind of inverter is of course possible. The obvious suggestion would be of a device operating to yield experience like that of a photographic negative (though as yet the hues are not to be inverted). Perhaps we could have ‘inverting lenses’ or a computer-inverted image presented to the retina. Perhaps in line with Block’s suggestion (692a n.17), we could intervene behind the retina to ‘chang[e] the

\[10\]
signals that represent [black] to signals that represent [white]’ and vice versa. The
difficulty, however, is that the result would not be like the perception of white
things replacing black, and *vice versa*.

I shall mention three difficulties with this first stage, before the move to
Inverted Earth. Firstly, would, if left on earth, still be able to tell black and white
apart. White things reflect a fair amount of light incident upon them; black things
do not. And this not just a piece of *recherché* scientific fact: the behaviour is
distinctive and influences our identification of the colours—it shows up in familiar
things that white things do, and that people can do with them. (a) Matt white
things, since they reflect nearly all incident light, have a varied appearance that is
highly sensitive to the variations in that light: shadows cast by other things show
up clearly on a white surface. Matt black things, by contrast, reflect little incident
light—hence it makes little difference to their appearance whether shadows are
falling upon them or not: little light is reflected anyway. (b) A similar fact is that it
is easier to see in a room with only a weak source of light if the walls are white or
pale than if they are black. (c) Water colour painting works well on white paper,
and not on black: a partly transparent wash makes quite a difference to the
appearance of white paper, but practically none to the appearance of the black,
which reflects little light regardless of the wash. (To change the appearance of the
black, one needs to *cover* it, not just tinge the light being reflected by the paper
behind. Incidentally: transparent things are no harder to identify after the
inversion: looking through them at something is, as always, pretty much the same
as looking directly at the object; and lightly coloured transparent liquids are
almost as easy to identify.)

Secondly and less importantly, there is no upper limit to the amount of
light that can fall on a surface (or reach the eye), but there is a lower limit—
namely, total absence of light. It is not clear, therefore, exactly how the inversion
is meant to go: as light increases in the world, darkness is supposed to increase in
the computer-generated image; yet there is clearly a limit to the darkness we can
have in the image. (With photographic negatives the problem is solved by
introducing one extra degree of freedom: allowing apertures and exposure times
to vary.) We might also wonder how eyes (which we know can be damaged by too
much light) would be supposed to react similarly to an excess of darkness. Is a
sensation of dazzle producible by black? And if not, then could any transposition
to Inverted Earth succeed in counteracting the inversion of experience?

Thirdly, black is, as one might say, the colour of the dark—darkness (or
absence of light) being a condition independently identifiable by one’s inability to
see anything much in it. I mean this not in the sense that one sees black when in
pitch darkness—for that is false. (One sees something more like grey.) Rather, for
example: when from a well-lit room one glances at the windows after nightfall, they look dark, and dark things outside look black.

With all these signs, it should be obvious that \( y \) would certainly not find, if she looked around her on Earth after the inversion, that white things looked like black things and black things looked like white things. If for a moment \( y \) took a sheet of paper to be black, she would be puzzled to find how differently it caught the light as she moved it around, and how clearly shadows showed up on it; how brightly it caught the light near a lamp; and what a good background it made for watercolour painting. And as the sun went down, even if her inverted vision became lighter, she would eventually find herself unable to see anything: her visual field would in the dark be quite uniformly light. With a flashlight, she could cast a patch of ‘darkness’ on the ground; and holding the paper in this paradoxical region of visibility (where different things produce recognizably different appearances—unlike real darkness, where different things all look much the same—), she might notice that the paper was the object most different in appearance from the general background of invisibility. Even if the paper at first looked in some way black, she could not for long take it for a black thing; it would in fact, whether she noticed it at first or not, be behaving like something white. And as she took a pen and found herself easily able to read what she wrote on the paper (though the ink was not a covering paint, as she could tell by dabbing some on the back of her hand) the impulse to treat the paper as white might be almost irresistible.

Now imagine we take \( y \) to Inverted Earth, hoping by an environmental switch to mask the change in her visual system. The first Inverted Earth we try is one where the sun shines as usual, but on objects whose lightness and darkness are the inverse of what we find on Earth. \( y \) will not long be taken in. Looking, with her inverted vision, at the sun or any source of light, she will see it as, strangely, dark. Even if the asphalt on the roads for a moment looks to her black, all the signs mentioned above will give away its true whiteness. (It shows shadows; it has the colour of walls that make for rooms needing relatively little interior lighting; it has the colour of a good watercolour paper.) Holding up a sample of asphalt near a lamp, she will find that, strangely enough, it looks ‘lighter’ the further it is from the lamp. Less light makes it look lighter. And holding up the sample as a barrier blocking the light from the lamp, she will find the back of the sample looking ‘lighter’ than the lamp itself. \( y \) will not find the double inversion undetectable.

Perhaps we should try a more thorough inversion. On the second Inverted Earth, the sun radiates darkness, and it is darkness that makes things visible. (I think of the biblical land where ‘the light is as darkness’, *Job* X.22.) To \( y \), with her inverting lenses, the darkness should look light, and light should look dark. Hence
after all, we hope, apparent normality will be restored. But the design of this infernal region is not entirely clear. Light is, we might say, actually something positive, but darkness is surely a privation. There are sources of light, which radiates in straight lines. Where it is dark, we can bring a lamp to shed light. Darkness does not in the same way have a source. Where there is light, you cannot bring a source to spread darkness. You can bring a shade, of course. But rather than radiate darkness in all directions, this will cast a shadow only in directions tracing back the light. It is easy to understand Locke thinking the ideas of dark and of light were equally positive; but the natural phenomena give plenty of signs that darkness itself is a privation. If this is right, then it is obscure how darkness—or anything functionally like it—could radiate in straight lines, and give rise to vision—even with the aid of an inverting device.

To summarize: If ‘light’ on Inverted Earth is something positive, that moves in straight lines, is obstructed by non-transparent obstacles, and otherwise behaves like light as we know it, then you will be able to tell white things from black regardless of the inversion of her visual system—for white things will look lighter (i.e. more like light sources) than black things. (The white things will behave as diffuse reflectors, etc.) If on the other hand ‘light’ is a mere privation, then it can’t literally be propagated. If, finally, we try taking ‘light’ as something negative, composed, say, of anti-photons capable of annihilating photons, then we may indeed suppose it to be propagated and reflected, and perhaps even to produce perception; but by this time, I am wondering how the interaction of such stuff with surfaces is supposed to produce perception of either their colour or its inversion. We might as well consider infra-red light instead and how it interacts with these surfaces, or even ultrasound—there is no reason to suppose an object’s interaction with those things is an index (even inverted) of its interaction with visible light of 350 to 700 nanometres; no reason therefore to suppose that the new perceptions produced on inverted earth will be corrections to the inversions in the colour-perception system produced at the earlier stage.

The moral so far seems to be that if black-white qualia inversion is possible, it will not take the form of making black things appear white and vice versa; and if it occurs at all, it will not be maskable by a second inversion in the external world. The difficulties here are quite evident from a careful look at photographic negatives. The world they present

I anticipate two objections. First: even if the simple negative-image inversion that I considered fails, could there not be another inversion that really did make black things look white, and white things look black? I think not. Consider asking a computer vision researcher to design a program to transform images, so that transparent objects appeared opaque and opaque objects
appeared transparent. The task is obviously impossible if the aim is to produce images true to what would have been perceivable from the original viewpoint if that change had actually occurred in the scene: for if the opaque objects had been transparent, they would have revealed other objects behind, information on which is typically unavailable in the image provided. (Of course there are further difficulties in the specification of the task: behind the newly-transparent objects, if there are opaque objects, are they too to be turned transparent?) The instruction to turn black things to white and vice versa is no better: if a black tablecloth becomes white, then it will now show up shadows from surrounding objects and light sources; but suitable information on those objects and sources will typically be unavailable in the original image. We might try reducing what we require: not an image true to what would have been perceivable from the original viewpoint if the change had actually occurred, but merely a plausible if speculative image. The task would be extraordinarily hard: the converter would have to be capable of distinguishing, for example, grey things in bright lighting from white things in low lighting, though both might be represented by the same grey in a photographic image—for only the second would be due to be changed. (I presume that mid-grey maps onto mid-grey.) To make such discriminations as accurately as a human, with scenes of everyday complexity, is already well in advance of machine-vision today. The point, however, is not to claim that even the latter, simplified, task is in principle impossible for a machine, but to make clear how complex is the task it would have to be performing: it would have to be sensitive not just to patterns of light-irradiation but to the presence of objects and their colour qualities in the environment. The moral seems to be that there could be a simple inverter introduced into your visual system, but no change in the environment could restore your experience to apparent normality. Conversely, there could be a world where white replaced black and vice versa; but no inverter in the visual system could blind you to that change—other than one that was so clever that it could identify the lightness and darkness of objects and produce images of things with precisely the inverted qualities. And that is a task we have practically no idea how to get a machine to perform. What is more, the idea that there might be interpersonal qualia variation of a type corresponding to variation between the input and output images of such a system is one that, I think I can say with relief, has never to this date been proposed.

The second objection I expect is the following: I have been treating white (in surfaces), for example, as characterized by a propensity to reflect a large proportion of incident light, fairly evenly across the spectrum; but isn’t it true that many people have no such conception of the colour, and some may even have conceptions incompatible with that idea? I would agree; but my suggestion is
that when people recognize a surface as white, they do so thanks to their sensitivity to what it does to the light and how it reflects it, though they may not realize that that is what they’re doing. A person may recognize Jane in the distance by her walk, but not know that it is that feature that is operative; even if she knows that it is her walk that she is noticing, she may be unable to specify what it is about the walk that is distinctive; if she can specify it at all, it’s quite possible that she can specify it only, for example, by imitating it (perhaps poorly), rather than by giving a physicist’s or anatomist’s specification. Similarly, what is distinctive of shiny surfaces, of highly matt surfaces, and of those slightly-glossy surfaces distinctive of Zone Books\textsuperscript{11} or the OUP Very Short Introductions, is clearly, on investigation, a matter of how they reflect the light—though there may be people who can classify surfaces as of those three types while ignorant of, or even denying, that fact. And so it is, I believe, with surface colours too. And that is all I need in order to explain why it is that a person would be unwilling to treat as black a white object of which they merely had a negative image.

3. \textit{From black \& white to colour}. Now for the final part of the argument. Several people have pointed out difficulties in the idea of a human spectrum inversion which derive from the non-symmetry of our actual colour space. Bernard Harrison has drawn attention to the variation in the individuation of shades in different portions of colour space. C. L. Hardin has pointed out two special difficulties with the commonly described spectrum inversion proposal. First: yellow and red, which are warm colours, would be mapped onto blue and green, which are cool—and the warm/cool classification seems a function of qualitative character, and, Hardin believes, physiologically grounded, being found extensively among humans of different cultures. The second difficulty is that blackened yellow—i.e. a form of brown—looks much more different in colour from (lighter) yellow than blackened blue looks from (lighter) blue—yet in the standard spectrum inversion story they

\textsuperscript{11} I am thinking of the list published by MIT Press including Gilles Deleuze’s \textit{Expressionism in Philosophy: Spinoza} (1990) and Georges Canguilhem’s \textit{The Normal and the Pathological} (1989).
(or something similar) would be mapped onto each other.\textsuperscript{12} I shall develop some further objections of the same kind.\textsuperscript{13}

The spectrum-inversion story tells of the insertion of colour-inverting lenses. If we let \textit{y} wake up in her old environment, she would of course (as expected) report the unusual colour appearances of things. The first difficulty is this. As Steven Palmer has pointed out in a fine discussion,\textsuperscript{14} the most saturated yellow is a light yellow; the most saturated blue is a dark blue. If light remains light and dark remains dark, therefore, a simple mapping of yellow to blue would surely produce different judgments of relative saturation: the maximally saturated blue, being seen by \textit{y} as a dark yellow, would not support the judgment that it was maximally saturated. It would be difficult to see how \textit{y} could come to mimic the colour-judgments of the rest of us. We might try adding (as Palmer therefore does) an inversion of lightness and darkness too, to ensure that the saturated blue would be mapped onto a light yellow (not a dark one), which might (with luck) be the most saturated of yellows. But we can see now that that will fail too—given the problems we have explored in the idea of any such inversion of black and white.

Suppose we ignore the problems and go back to an exchange of light yellow and light blue in the visual system, and then take \textit{y} to Inverted Earth where the sky is yellow, and the sun is blue. The change in the colours of things is supposed to cancel out the original inversion, thus confirming the success of the original intervention. The difficulty is this. The most saturated yellow is a fairly light yellow; if one substitutes a blue of equal lightness on Inverted Earth, neither the local residents (if they are like us except in hue-language) nor the visitor with the inverting lenses would call it highly saturated. So the double inversion would not at all make \textit{y} think she was back to normality. And if instead we exchange light yellow for dark blue, and in general light for dark in the visual system, then we already know from the previous section that no environmental change can compensate for the change.

One last attempt to defend an inversion hypothesis: What about simply swopping red and green, reflecting those hues in the yellow-blue axis? (Palmer calls this ‘by far the most plausible’ of the three inversions that he believes are


\textsuperscript{13} See also Hardin, \textit{Color for Philosophers}, p. 140.

'logically possible' (op. cit., p. 926).) The same kind of problem arises again. The best reds are dark, whereas the best greens have middle levels of lightness; ‘red’ (by contrast with ‘orange’ and ‘pink’) is applied only to samples at lower lightness levels, but ‘green’ is regularly applied to samples at virtually all lightness levels. If therefore we just mapped red onto green and green onto red, then saturated paradigm greens would end up producing red’ sensations of only medium lightness—which would surely be too light to support a judgement that they are the most saturated of their type. And if we tried to add in a mapping of dark onto some other degree of lightness, we would run into the kinds of problem we have already explored—not to mention the additional problems that would come from mapping dark onto a medium level of lightness. It seems, therefore, that, of the three supposedly undetectable inversions that Palmer presents as being ‘logically possible’—(1) an inversion of the red-green axis, (2) an inversion of yellow-blue and black-white axes together, and (3) an inversion of all three axes—, all are, in fact, impossible. The first two we have just seen directly to be impossible, and the third is the combination of the first two.

What morals should we draw? A range of options seems available. In the discussion of white and black, my main concern has been not with qualia inversion, but with the inversion of colour experience. A response friendly to qualia would be this: I have shown that experience of white and experience of black can’t be undetectably inverted, but qualia might none the less be invertible. If this is so, however, there are still some striking consequences for our conception of colour perception: for the having of qualia will be further away from colour perception than is usually thought. Seeing a surface as white is more like seeing a surface as silvery, for example, than we tend to think—it is not individual qualia but, at the very least, patterns of qualia that count for the perception of these qualities. (Some might go further and say it is not even patterns of qualia that count, but rather, awarenesses of ways in which the surfaces catch or change the light.) To talk of qualia of white may make no more sense than to talk of qualia of silveriness or of shininess. This should not surprise us. We adapt very well to many variations in lighting, noticing both sameness of colour and variation in the light: this perception of the sameness of colour cannot be understood in terms of the persistence of a single kind of quale. If we talk of qualia at all, they will be

15 ‘Unlike the other colors, green and blue are reported at virtually all lightness levels. ... The best greens are seen at the middle lightness levels. ... Red, brown, and purple are reported only at the lower lightness levels.’ Peter K. Kaiser & Robert M. Boynton, *Human Color Vision* 2nd edn. (Washington, DC: Optical Society of America, 1996), p. 504.
features of experience that one ‘looks through’, so to speak, in order to see the \textit{colours} of things, just as much as one looks through them to see their shapes and sizes.

The talk of white’ experiences—as experiences that a white thing would produce in normal observers under normal circumstances—if it succeeds in characterizing a single precise qualitative state, gives a special status to one particular appearance, which white things actually have only under the most stringent and rare of conditions. Red$_{24}$, presumably, will be the colour that things have if they produce in normal observers under normal circumstances red$_{24}$' experiences. Yet it would not be surprising if things that are red$_{23}$ also produced red$_{24}$' experiences, under conditions that were by ordinary standards quite ‘normal’. Only under specially specified conditions would we be, in the most direct fashion, aware of a thing’s proper colour. But rather than treating human beings as measuring colours, so to speak, simply against their own responses, we could instead treat human beings as measuring colours against each other—measuring objects against objects—not forgetting of course that the comparison (and reading of the measurements, so to speak) is itself mediated by human sensibility.\footnote{For a development of this suggestion, treating colours as something like 'phenomenal natural kinds', see my article ‘Colour, World and Archimedean Metaphysics: Stroud and the Quest for Reality’, \textit{Erkenntnis} 66 (2007), at end of § 4.} We identify red$_{24}$ not by seeing what experience it produces under ideal circumstances, but by comparing it with other things—some of whose colours we know. And the business of learning the concept of a particular colour would require not acquaintance with a particular kind of sense-datum or quale, but acquaintance with the colour itself—that is, with a certain natural phenomenon in the world. (The conception I have of red could differ from the conception of ‘red’, that I would have had if instead I’d been brought up in a world where green things were producing my ‘qualia’ of red. There would be something like a demonstrative element in the former conception that would be different from the parallel element in the latter conception.) We thereby detach the colours from the mythical conception of ‘the proper way’ to see the colours, and the ‘normal conditions’ under which, so to speak, they shine in their true colours.

Of course others would draw a more radical moral: that there are no such things as colour qualia at all. In vision, we represent things as having colours and other features, as laid out in a visual field characterized (among other things) by the apparent directions in egocentric space of those things from us; and a
representational conception of our visual experience, it might be argued, is adequate and complete. The reasons for thinking this go beyond the present discussion, though the discussion lends some support to it. A large part of the inverted qualia story is perfectly compatible with this conception: a person could find herself apparently seeing red where she used to see green. What would not be compatible is the idea that this could happen completely undetectably and in ways that were not essentially connected with her behaviour and representation of the world. And the difficulties we have found both with the traditional and the Inverted Earth examples give support to the conclusion that the latter idea does not describe a coherent possibility.

Of course there are larger questions about what the motivation is for the rival conceptions in this area. In his presentation of the Inverted Spectrum story, Ned Block considers objections concerning memory, and then—rather than address them head-on—explains that he can complicate the example in such a way as to evade the problems.\(^\text{17}\) I have argued that the new Inverted Earth case actually fails to evade the original objections—to the extent that those objections actually had merit. What we really need, however, is not the accumulation of increasingly complex thought-experiments, but a reconsideration of the motivations lying behind the various conflicting intuitions. I have left that mostly for another occasion, but I hope I have brought to the surface one quite major point of principle: that both the inverted spectrum and the Inverted Earth stories tend to neglect the very special character of surface colours—which is so striking in the case of white and black. On investigation, these, we see, are not so plausibly characterized by qualia produced as by the kinds of way in which surfaces interact with light. When we reflect on that fact, the possibility of a coherent and in principle undetectable inversion of colour-perception seems remote; and if we give a role to qualia in the production of colour-perception, it will be a surprisingly subterranean one.\(^\text{18}\)

\(^{17}\) ‘I believe that these criticisms can be defeated on their own terms, but instead of trying to do that here, I propose to take a different tack.’ (682a)

\(^{18}\) For comments and discussion I am grateful to John Campbell and James Pryor.