

This article was originally published in *Philosophical Perspectives* 21, 1, 2007: 73-115, edited by John Hawthorne. This is a revised version, Block, N. (2009). *Wittgenstein and Qualia. Permutations: Essays on Hilary Putnam. M. Baghramian. Oxford, Oxford University Press, 275-317.*

10

WITTGENSTEIN AND QUALIA¹

Ned Block

Introduction

The representational content of an experiential state can often be described in public language partly in terms of qualities of objects that bear some salient relation to the state, for example, “looks red” or “feels like sandpaper” or “smells rotten,” but it is the contention of this chapter that public language terms including terms for such properties of objects do not fully capture the contents. Qualia in my terminology are ways things look red or feel like sandpaper or smell rotten. If things can look red in more than one way – as I will be arguing – “looks red” does not fully capture the content of the state. Qualia can be referred to in public language, for example as the quale I get when I see green things. But that reference to the quale does not fully capture its content – it does not capture the individuating particularity about the way that I see green that is different from the way others may see green. Similarly, I can refer to the content of a thought as “the content of the thought I had at 11:33 AM” without fully capturing its content.²

We have a notion of the content of an utterance or thought which can be fully captured without specifying anything about modes of presentation. If I say “Napoleon is buried in Paris,” we can fully capture an important content of that utterance by saying it is the content that Napoleon is buried in Paris, even though that way of putting it does not attempt to specify my eccentric ways of thinking of Napoleon and Paris. But the experiential aspect of the content of an experience just is a kind of mode of presentation, so a notion of the content of an experience that left out the mode of presentation would be defective.

Note that I am assuming color realism, i.e. that there are facts about the colors of things and those facts involve things actually having colors, for example that fire hydrants are often red and grass is often green. This position contrasts with error theories that say that nothing really is colored but rather color involves an erroneous projection of mental properties onto the world (Boghossian and Velleman 1989, 1991).

NED BLOCK

In other work, I have defined qualia as qualities of experience that cannot be defined in terms of their representational,³ functional⁴ or cognitive properties. And I have used the inverted spectrum hypothesis (1990, 1994, 1999; Block and Fodor 1972) in those arguments. The word “qualia” is often used to indicate a putative feature of our experience whose existence is disputed. Here the focus is on the ineffability of experience, and so it is convenient to use the term “qualia” to denote features of experience that have the kind of ineffability just mentioned.⁵

Two inverted spectrum hypotheses

Wittgenstein’s “Notes for Lectures on ‘Private Experience’ and ‘Sense Data’” (Wittgenstein 1968) were written in English, apparently between 1934 and 1936 (Rhees 1968). Wittgenstein is notoriously difficult to interpret, even to the extent that scholars cannot agree whether claims that are clearly formulated in his writings are being asserted or denied or something else. I will put forward an interpretation of Wittgenstein’s view of the inverted spectrum hypothesis and of the nature of sensory experience, not as a proposal that meets the standards of Wittgenstein scholarship, but rather as a suggestion in a recognizably Wittgensteinian framework that is worthy of discussion (and refutation) on its own merits. When I attribute a view to Wittgenstein, you may wish to understand that as “Wittgenstein-according-to-one-non-expert.” The seeming endorsement of the possibility or at least coherence of one kind of spectrum inversion (the kind I am calling “innocuous”) begins in this passage:⁶

The normal use of the expression “he sees [red] where ... ” is this: we take it as the criterion for meaning the same by ‘red’ as we do, that as a rule he agrees with us in giving the same names to the colours of objects as we do. If then in a particular instance he says something is red where we should say that it’s green, we say he sees it different from us.

Notice how in such a case we would behave. We should look for a cause of his different judgment, and if we had found one we should certainly be inclined to say he saw red where we saw green. It is further clear that even before ever finding such a [[284]] cause we might under circumstances be inclined to say this. But also that we can’t give a strict rule for ...

Consider this case: someone says “it’s queer/I can’t understand it/, I see everything red blue today and vice versa.” We answer “it must look queer!” He says it does and, e.g., goes on to say how cold the glowing coal looks and how warm the clear (blue) sky. I think we should under these or similar circumst[ances] be incl[ined] to say that he saw red what we saw [blue].⁷ And again we should say that

WITTGENSTEIN AND QUALIA

we know that he means by the words ‘blue’ and ‘red’ what we do as he has always used them as we do.

(Wittgenstein 1993: 231)

I will be arguing that the kind of inverted spectrum hypothesis that Wittgenstein endorses (as possible or at least coherent) in this passage commits him to something he would not agree with, so a defender of Wittgenstein might respond that after proper consideration, Wittgenstein would not have endorsed it.⁸ However, I think what Wittgenstein describes here is obviously coherent. Further, subsequent technological developments have shown that something near enough is possible, even technically feasible now. Colors are easily reversed in digital television. I myself have appeared, inverted in color, in an interview on a German television station. And it is feasible right now for virtual reality goggles to make use of such technology in producing that inversion experience in a subject. With a relatively small investment, such virtual reality goggles could be produced now. I think most vision scientists would agree that the same transformations could in principle be accomplished by circuits embedded between the eye and the brain, although no one knows how to do this now. (I will describe this science fiction scenario as having “wires crossed” in the visual system: a terminology that I think was first used by Putnam (1981: 80).)

Later in these notes (Rhees 1968: 316, Stern: 285), Wittgenstein confirms the endorsement and introduces the version of the inverted spectrum hypothesis that he rejects – what I will call the “dangerous” type. (This is my terminology, not Wittgenstein’s – I mean dangerous for a Wittgensteinian point of view.)

... We said that there were cases in which we should say that the person sees green what I see red. Now the question suggests itself: if this can be so at all, why should it [not] be always the case? It seems, if once we have admitted that it can happen under peculiar circumstances, that it may always happen. But then it is clear that the very idea of seeing red loses its use if we can never know if the other does not see something utterly different. So what are we to do: Are we to say that this can only happen in a limited number of cases? This is a very serious situation. – We introduced the expression that A sees something else than B and we mustn’t forget that this had use only under the circumstances under which we introduced it.⁹ Consider the prop[osition]: “Of course we never know whether new circ[umstance]s wouldn’t show that after all he saw what we see.” Remember that this whole notion need not have been introduced. “But can’t I imagine all blind men to see as well as I do and only behaving differently; and on the other hand imagine them really blind? For if I can imagine these possibilities, then the

NED BLOCK

question, even if never answerable makes sense.” Imagine a man, say W., now blind, now seeing, and observe what you do? How do these images give sense to the question? They don’t, and you see that the expression stands and falls with its usefulness.

The idea that the other person sees something else than I, [[317]] is only introduced to account for certain expressions: whereas it seems that this idea can exist without any reference to expressions. “Surely what I have he too can have”.

(Wittgenstein 1993: 285)

Note the words “We said that there were cases ... ” suggesting that those cases are coherently describable and perhaps possible. (Some may say that those words are not meant to be understood as Wittgenstein’s own view, but it is hard to read this material without seeing that Wittgenstein supports the idea that, in certain circumstances, there would be a use to the idea that the other person sees things as green what we see as red.) What is the difference between the innocuous and dangerous cases? In the innocuous case, colors are inverted but certain properties of them – warm and cool – are not, and this happens suddenly to someone whose color experience has been normal before. So it is detectable because of two significant changes: (1) Things the subject knows to be red such as hot coals and blood look green and things the subject knows to be green such as grass look red and (2) Blood suddenly seems cool colored and grass seems warm colored. The dangerous scenario is widespread and is not behaviorally detectable.

There is an immediate problem in figuring out what is supposed to be happening in the dangerous scenario. The most straightforward way for the innocuous case to become widespread would be if the odd thing that happened to the subject of the innocuous inverted spectrum scenario simply happened repeatedly, the result being many inverted people who saw blood as green and cool colored and grass as red and warm colored and were aware of this fact about their vision. But that is not what Wittgenstein intends in the dangerous scenario, since he takes the dangerous scenario to be one that is behaviorally undetectable. He says about the dangerous scenario that “we can never know if the other does not see something utterly different.”

A proponent of qualia could suppose that the odd thing that happened to the subject in the innocuous scenario happens in the dangerous scenario at or before birth, making the inversion widespread but not detectable. For the inverted people have not experienced a change and they have learned to use color words and the word “cool” in conjunction with the color experiences they get on seeing, for example, grass. So the proponent of qualia could take the dangerous scenario to be one in which warm and cool are inverted along with the colors themselves (Block 1990: footnote 16; Levine 1991), and in which the subject has no memories of things having looked differently in the past. What Wittgenstein seems to be saying is that the legitimate use of the

WITTGENSTEIN AND QUALIA

location “The person sees green what I see red” would not apply to such a scenario.

The inner arena model

This paper started life as a comment on a chapter of Paul Horwich’s book in progress on Wittgenstein (Horwich, forthcoming). Horwich – like Wittgenstein – contrasts a form of inverted spectrum involving experiential contents that can be captured in ordinary language with another form of inverted spectrum which cannot, and so, in my terms involves qualia. I am grateful to Horwich and his chapter for inspiring me to try to elaborate the case for qualia from considerations of inverted and shifted spectra – the subject of this chapter.

Horwich, interpreting and elaborating Wittgenstein’s ideas, argues that (what I would call) the dangerous inverted spectrum hypothesis presupposes a private arena model according to which experiences are observed directly by the person whose arena it is, but only indirectly by others. And observation of experience is not just a matter of having the experience. On Horwich’s account, the private arena model motivates a conception of experience as essentially subjective and ethereal and not constituted by anything objective and concrete. He contrasts these ideas unfavorably with the claim that the meaning-constituting use of “I am in pain” should be taken instead to be an expressive use in which those words are substituted for a natural expression of pain. On this account, the reference of “pain” is not involved in fixing the meaning of the term. He also links the private arena model to the idea that there must be determinate answers to questions about the sensations of others, e.g. whether computers can have pains. A related view of the “Cartesian Theater” has been advocated by Dennett (1991).

I will argue that an inner arena model has little to do with the argument for inverted or shifted spectra.

It is not easy to find contemporary philosophers who have argued for the possibility of an inverted spectrum and who can also reasonably be said to accept a private arena model. D. M. Armstrong (1968; Bacon, Campbell, and Reinhardt, 1993, reply to Martin) accepts an observational model of introspection and something that might be called an inverted spectrum, but the issue is complicated by his eliminativism about experiential qualities. Sydney Shoemaker (1982) has argued strongly for the possibility of an inverted spectrum but equally strongly against an observational model of introspection (1994b; 1996b, including footnotes not in the journal version). As Shoemaker notes, it is not clear that even Locke accepted an observational model of perception, given that his theory of perception involves causal signals from external objects affecting sense organs so as to produce ideas in our minds that represent those external objects. This is a reliable but fallible process as contrasted with introspection, in which for Locke there is no gap between appearance and reality.

NED BLOCK

Shoemaker gives a detailed analysis of what is wrong with various observational models of introspection. For example, perception involves an organ of perception but there is no organ of introspection; perception involves experiences of what is perceived but there are no experiences of experiences in introspection; there are no different introspective perspectives on the same experiences in the way that there are different perceptual perspectives on the same perceived object; the relation between the pain and the introspective belief about it is more intimate than an observational relation: the pain is more like a part of the introspective belief than a reliable cause of it; the objects of perception often exist completely independently of the perception of them and the perceivers who perceive them, but what is introspected cannot be independent of the introspection and introspector in this way. Shoemaker emphasizes a version of the self-intimating nature of experience in relation to introspection which has no analog in observation. His overall viewpoint provides a serious challenge to the idea that a dangerous inverted spectrum necessarily involves an observational model of introspection.

I am another example of someone who argues for an inverted spectrum but does not accept the observational model of introspection. I agree with the points just mentioned about what is wrong with the observational model of introspection, but unlike Shoemaker, I think that one can attend to one's own experiences. (And in that respect, my view is more like an observational model than Shoemaker's.) Shoemaker's view is shared by Fred Dretske (1995), Gilbert Harman (1990), Michael Tye (1995, 2000) and many others who advocate what G. E. Moore termed the diaphanousness (or sometimes the transparency) of experience. Harman (1990) puts the point by saying that the more one tries to attend to one's experience of the tree, the more one attends to the real tree instead. Although Moore is sometimes cited as the originator of this point, he did not actually accept it. I have heard him quoted saying "... the moment we try to fix our attention upon consciousness and to see what, distinctly, it is, it seems to vanish: it seems as if we had before us a mere emptiness. When we try to introspect the sensation of blue, all we can see is the blue; the other element is as if it were diaphanous." (Moore 1903) But these words are followed by what I regard as a more significant truth: "Yet it can be distinguished if we look attentively enough, and know that there is something to look for." (See Amy Kind's treatment in her 2003.)

If we want to appreciate the case for the claim that we can attend to our own experiences, we would do well to avoid examples like Harman's in which one is looking at a tree, trying to attend to one's experience instead of the tree. An example I have used (Block 1995): you are engaged in an intense conversation, realizing suddenly that for some time, you have been hearing street-drilling outside and have even raised your voice to compensate for it. At the moment of the sudden realization, you switch at least some of your attention to the experience of the sound, and not just to the sound

WITTGENSTEIN AND QUALIA

itself. For another type of example, consider the difference between blurry vision as of something, say a movie, whose lines may be clear and crisp, as contrasted with clear and crisp vision of a blurry movie (Block 1996, 2003). There is an introspectible difference that cannot be appreciated without attending to the experience as well as to what the experience is of. What we experience is closely tied to what concepts we bring to bear in experience and introspective experience is no exception. Of course these examples are the beginning of an argument, not the end.¹⁰

Another way in which I differ from the private arena model as sketched by Horwich is that I do not take experience to resist objective concrete constitution. The view of qualia advocated here is, at least according to me, fully compatible with physicalism. I also doubt that the private arena model motivates the claim that experience resists objective concrete constitution. I take what Dennett (1991) calls the Cartesian materialist picture to be one that combines a private arena view with materialism.

Still another way in which I differ from the private arena model at least in Horwich's version is that I do not think that questions about other minds need have determinate answers. An example that I have given elsewhere (Block 2002) is that belief in qualia does not require a determinate fact of the matter as to whether fish have qualia.

Horwich takes the difference between the innocuous and dangerous inversion hypotheses to be the difference between a case that can be described in public language with expressions like "smelling a rose," "painful" and the like, and a case that requires a private way of describing experience not expressible in public language. In my terms, the difference is qualia. He takes the rules of use for mentalistic expressions to dictate that your experience when you look at red things in normal circumstances is just like mine. There can be no further question of what it is like for you to see red as compared with me. To ask what looking red is like is like asking how long five feet is. Here, I disagree, as will become clear in the sections below on shifted spectra and inverted spectra. Horwich argues that the dangerous form of the inverted spectrum hypothesis derives from the internal arena model, especially the conflation of first-person sensation reports – which he construes as expressions of sensation (on the model of natural expressions of pain) – with observations. As I have mentioned, many contemporary advocates of inverted spectra do not accept key tenets of the internal arena model and the observational view of introspection. Of course, what one explicitly advocates and what one presupposes are two different things. In introducing a materialistic version of the "Cartesian Theater," Dennett allows that "Perhaps no one today explicitly endorses" such an idea. Indeed, "Many theorists would insist that they have explicitly rejected such an obviously bad idea" (Dennett 1991: 107). But, he insists, the picture he has sketched nonetheless plays a behind-the-scenes role. And perhaps Horwich would suppose something of this sort too. But Horwich

NED BLOCK

and Dennett give no reasons for thinking that actual contemporary arguments for the possibility of dangerous inverted spectra really do presuppose the mistaken models.

In my view, Wittgensteinians who emphasize the observational model of qualia are barking up the wrong tree. The view of qualia that leads to the epistemic problems that exercise Wittgensteinians is that there are determinate facts of qualia independently of our cognitive access to those facts. A Wittgensteinian might call it the “inner light” model.¹¹

Intrasubjective vs. intersubjective

Let us now return to the issue of what the difference is between the innocuous and dangerous inverted spectrum scenarios. Philosophers have made much of the difference between an intrasubjective spectrum inversion – in which a person at one time is said to be inverted with respect to the same person at another time – and an intersubjective spectrum inversion in which one person is said to be inverted with respect to another (Putnam 1981; Shoemaker 1982, 1996a). The key dialectical difference can be explained as follows. Suppose we have a pair of identical twins at birth, one of whom has had the wires crossed in his visual system. The twins are raised normally and acquire color terminology in the normal way. On the point of view that I favor, it is possible (as far as we know) that the way red things (which they agree are red) look to one twin is the same as the way green things (which they agree are green) look to the other. Now a vulnerability in this line of thought stems from the following objection (Block 1990, 1994; Harman, 1990).

Notice that it is not possible that the brain state that one twin has when he sees things that both twins call “red” is exactly the same as the brain state that the other twin has when he sees things that they both call “green.” At least, the total brain states can’t be exactly the same, since the first causes the subject to say “Its red,” and to classify what he is seeing as the same color as blood and fire hydrants, whereas the second causes the other twin to say “Its green,” and to classify what he is seeing with grass and Granny Smith apples. Here is an example. Suppose that the color-relevant brain state that the normal twin (the one who has had no operation) has when he sees red things and that the abnormal twin has when he sees green things is R-oscillations in area V4, whereas the color-relevant brain reaction in the normal twin to green and to the abnormal twin to red is G-oscillations in area V4. If the proponent of the possibility of an inverted spectrum says this is evidence for inversion, the objector can say that phenomenal properties should not be thought to be based in brain states that are quite so “localized” as R-oscillations in V4 or G-oscillations in V4, given that R-oscillations lead to different reports in the different twins. Rather color experience should be seen as based in more holistic brain states that include the brain bases of reporting and classification behavior. (I am assuming

WITTGENSTEIN AND QUALIA

agreement on an extremely weak form of physicalism.)¹² Thus the objector whom I am thinking of will want to say that one twin's holistic brain state that includes R-oscillations and the other twin's holistic brain state that includes G-oscillations are just alternative realizations of the same experiential state: that experiential state has a disjunctive realization. So the fact that red things cause R-oscillations in one twin but G-oscillations in the other doesn't show that their experiences are inverted.

Of course the G to R transition within a single person is a transition in the way things look. However, that experiential difference has only been demonstrated intrasubjectively, keeping constant the larger brain state that specifies the roles of R- and G-oscillations in classifying things. The R/G difference in each subject grounds a color experience difference, but this gives us no knowledge of cross-person comparisons. The objector can insist on typing brain states for interpersonal comparisons holistically. And most friends of the inverted spectrum are in a poor position to insist on typing experiential states locally rather than holistically, given that they normally emphasize the "explanatory gap," the fact that there is nothing known about the brain that can adequately explain the facts of experience. So the friend of the inverted spectrum is in no position to insist on local physiological individuation of qualia. At this stage, the debate seems a standoff, and that is where the intrasubjective inverted spectrum comes into the picture.¹³

The intrasubjective inverted spectrum scenario can be seen as a way for the defender of the inverted spectrum to evade this objection, for if the change happens in the life of an individual person, we have introspective and behavioral evidence of an inversion and so can avoid the issue of whether brain states should be thought of in a localistic or holistic manner (Block 1990, 1994; Shoemaker, 1982).

I agree with this superiority of the intrasubjective inverted spectrum scenario, and will be pursuing it in detail later. But for now, my point is that the intra vs. inter difference is not directly involved in the difference between the innocuous and the dangerous inverted spectrum. What is involved, as I will be arguing, is whether the inverted pair are both normal, since that is what is relevant to the expressibility of phenomenal character in ordinary language.¹⁴

Normality

So what is the difference between the "innocuous" inverted spectrum case, the one Wittgenstein regards as at least coherent, and the "dangerous" case, the one that he rejects? One difference is that the innocuous case occurs suddenly and "under peculiar circumstances": the subject agrees that fire and the sky now "look queer"; whereas in the dangerous case "it may always happen," even not under peculiar circumstances. Another difference is that the innocuous case is behaviorally detectable because the subject says, e.g., coal looks coal colored, whereas what Wittgenstein is suggesting in the dangerous case

NED BLOCK

is an inverted person who is not behaviorally detectable. I would guess that the behavioral undetectability of the dangerous inversion was extremely important to Wittgenstein's rejection of it. (This comes out in the second quote from Wittgenstein together with the passages in the text surrounding it.)

However, there is another difference that does not depend on behavioral undetectability, and this difference is the one I am going to focus on here. I will argue that the dangerous scenario can be used to argue for qualia even if it is behaviorally detectable. Qualia, you will recall, are features of experience that are not expressible in ordinary language, including terms for properties of objects. One cannot express what red things look like colorwise in normal circumstances using color terms as in "looks red," since things that look the way red things look to me (in normal circumstances) may look to you the way green things look to me. So among qualia are ways things can look red. Something might look red to two people in different ways. And I think a crucial difference between the innocuous and dangerous cases relevant to qualia is whether normal people (or one normal person at different times) in normal circumstances can be said to be color inverted or shifted with respect to another normal person in normal circumstances. In the innocuous case, the color inverted person is abnormal. But if "one section of mankind has one sensation of red and another section another" (Wittgenstein 1958, section 272, p. 95) then it would seem that normal people can have inverted color experience with respect to one another, and that is enough to produce a "very serious situation."¹⁵ Here is why.

It is incoherent to suppose that there are normal people in normal viewing circumstances for whom red things look green. To see this, suppose you are one of those putatively normal people for whom red things always look green. Suppose that you and I agree that there is a difference between us in ways that things look color-wise, and that the difference can be described by the locution: "Things we agree are red look to you the way things we agree are green look to me." I say to you that whereas red things look red to me, red things look green to you. You can object to the idea that red things look green to you. You can reply: "Who says that red things look green to me? Why are you the one for whom red things look red? I can with equal justification say that red things look red to me and green to you!"

If we are both normal, you have as good a case as I do. Since we are both normal, the right response is that neither you nor I have a superior claim to be the one for whom red things look red and green things look green. The situation is relevantly symmetrical. What we should say instead is that the way red things look to me is the same as the way green things look to you, and in allowing that there are ways things look that cannot be expressed in terms of "looking red" or "looking green" or any phrase of the form "looking F" where F is a color name, we step into the realm of qualia. (Given what I mean by "qualia," this is definitional.)

WITTGENSTEIN AND QUALIA

Using the familiar “what is like” terminology (Farrell 1950; Nagel 1974), the point is that if we acknowledge the existence of an innocuous inverted spectrum scenario, we can say that green is what it is like for the abnormal person to see red, whereas for the normal person, red is what it is like to see red. So the innocuous scenario does not require us to suppose that there are color experiences that cannot be expressed in terms of properties of things. However, if we allow the existence of a dangerous scenario, in which normal perceivers are inverted with respect to one another, we cannot say of either of them that green is what it is like to see red. If we acknowledge the existence of an inverted spectrum in this sense, we have to agree that no color name expresses what it is like for either one of the inverted people to see red. There are ways things look red, and those ways are qualia. (Whether the actuality of a dangerous scenario is required for qualia or whether its possibility will do will be taken up later.)

Note that I am not denying that what it is like for them to see red can be referred to in English. For example, we can refer to it by saying “What it is like for that person to see red.” What we cannot find is a color name “F” such that what it is like for one of these people to see red can be expressed in the form “looking F,” and in that sense we can say that the experiential property is an ineffable quale.

To repeat: I am not denying that Wittgenstein was concerned with the behavioral indistinguishability of the persons in the dangerous inverted spectrum case. If there could be an inverted but behaviorally indistinguishable pair of people, then it is hard to see how the difference between them could be relevant to any everyday use of terms. My point is rather that an inverted pair both members of which are normal, is problematic for Wittgenstein even if they are not behaviorally indistinguishable.

Sydney Shoemaker (Shoemaker 1982) distinguishes between intentional and qualitative similarity. If something looks red to you and me, then our experiences are thereby intentionally similar, but if your spectrum is inverted relative to mine, our experiences are thereby qualitatively dissimilar. (Although I have not yet argued for the possibility of the dangerous kind of inversion, I will assume it here to get the conceptual groundwork in place.) If something looks red both to you and to me, our experiences thereby have a shared intentional content, but if our spectra are inverted with respect to one another, and if you are looking at a red thing and I am looking at a green thing, our experiences have a shared qualitative content. Looking red is an intentional content of color experience, not a qualitative content. For both members of the inverted pair, red things look red and green things look green. Color language in application both to the outer and the inner is keyed to intentional contents of experience (Block 1990). States that have qualitative contents are just qualia, and it is the existence of qualitative contents that poses the challenge to expressibility of qualia in language.

NED BLOCK

It may seem from what I have said so far that the issue of qualia is the issue of whether the phenomenal character of experience goes beyond its representational content, that is, the issue of representationalism, or representationism as I call it (Block 1990, 2003). But that is not the case, as can be seen by a brief consideration of a view held by Shoemaker (1994a, 1994b) and Michael Thau (2002). In Shoemaker's version, the view is that when one looks at a ripe tomato, one's experience represents the tomato as having two distinct kinds of color-relevant properties, as being red and as having a certain "phenomenal property." If your spectrum is inverted with respect to mine, and we are looking at different items with complementary colors (e.g. you a red thing, me a green thing), our experiences represent those items as having the same phenomenal property and as having different colors. These phenomenal properties are definable in terms of qualia. As far as what I have said so far is concerned, one might think of them as qualia projected onto objects – but Shoemaker does not think of them that way, since he holds they are causally efficacious, indeed defined in terms of their production of qualia. So Shoemaker is a kind of representationist, and that allows one to see that one can be both a representationist and a believer in qualia. On Shoemaker's view, color experience has two kinds of representational contents. One of them is what I have been and will be calling intentional contents, namely color-representing representational contents, contents that represent something as, e.g., red. But the other kind of representational content (the kind that represents phenomenal properties) is not expressible in public language. The issue of qualia is not the issue of whether the mental properties of experience can be fully captured in terms of what is represented by experience, since it may be that what is represented by experience cannot always be fully captured in public language.

I endorse the principle that normal perceivers in normal circumstances see red things as red, and thus that public language attributions of color experience express intentional contents of experience rather than qualitative contents. We could call this principle the Principle of Normality.¹⁶

Gottlob Frege held a somewhat similar view:

The word 'white' ordinarily makes us think of a certain sensation, which is, of course, entirely subjective; but even in ordinary everyday speech, it often bears, I think, an objective sense. When we call snow white, we mean to refer to an objective quality which we recognize, in ordinary daylight, by a certain sensation. If the snow is being seen in a coloured light, we take that into account in our judgement and say, for instance, 'It appears red at present, but it is white.' Even a colour-blind man can speak of red and green, in spite of the fact that he does not distinguish between these colors in his sensations; he recognizes the distinction by the fact that others make it, or perhaps by making a physical experiment.

WITTGENSTEIN AND QUALIA

Often, therefore, a colour word does not signify our subjective sensation, which we cannot know to agree with anyone else's (for obviously calling things by the same name does not guarantee as much), but rather an objective quality.

(Frege 1884/1953: §26, quoted in Byrne 2006)

And, although my interpretation is not standard, one can see Wittgenstein's beetle in the box passage as compatible with a similar view.¹⁷ The idea would be that the thing in the box – a quale – is irrelevant to the language game, which involves only what I am calling intentional contents (that is, color-representing contents). But qualia can exist and be important for some purposes even if the language game that Wittgenstein has in mind has no need of them.

In a comment on a version of this chapter at Hilary Putnam's 80th birthday conference, Pierre Jacob noted, correctly I think, that belief contents do not mirror intentional contents of perception in the sense in which I am using the term. If our putative inverted twins both see a ripe tomato and have intentional color contents representing it as red (and the same shade of red) but have different qualitative contents, we would be reluctant to say that they have exactly the same color-relevant beliefs about the tomato. For the twins have different qualitative beliefs about the color of the tomato, where a qualitative belief is a belief that includes a color quale (or, alternatively, a belief that represents the tomato as having a certain phenomenal property in Shoemaker's sense of the term). Another way of seeing Jacob's point would be to think of the intentional contents of perception as individuated both in terms of reference and mode of presentation, where qualia are taken as modes of presentation of those referential contents as suggested by Tyler Burge (2003). Nothing hangs on how we use the term "intentional content" (and I will continue to use the term to mean color-representing purely referential content). The important point is that if normal people can be inverted with respect to one another, there is an aspect of color experience that cannot be captured in ordinary language in terms of properties of objects.

The Normality Principle links public language color terminology to intentional rather than qualitative contents of experience. I now turn to the argument for that link, and then to arguments for shifted and inverted spectra.

Color terms express intentional not qualitative contents

Assuming without argument at this point that the dangerous scenario is possible, what exactly is the argument that color terminology ("red," "green," etc.) is keyed to color-representing intentional contents rather than qualitative contents?

In daily life, we do not distinguish between intentional and qualitative content, just as we do not normally distinguish between weight and mass

NED BLOCK

or between rest mass and relativistic mass. Shouldn't we suppose that there is some sort of indeterminacy, as with "mass," where ordinary uses of the term "mass" partially denote both, as argued plausibly by Hartry Field (1973)? In Block (1990), I argued that if spectrum inversion is known to be rife, we should think of our tacit semantic policy as one of using color terms as applied to experience to denote intentional contents of experience, since when we say of someone that the fire hydrant looks red to him, we often know what the intentional content is, but not what the qualitative content is. So how can we be understood as attributing a qualitative content? However, we do not know whether spectrum inversion is actually rife (cf. Byrne 2006: §3.8). Is there an argument from our lack of knowledge whether or not inversion is rife to the same conclusion?

Yes, a similar argument applies. If phrases like "looking red" were intended to apply to qualitative contents, we would have a vulnerability to widespread error that we tacitly assume that we do not have. Of course if everyone has the same spectrum except for a few defectives who see red things as green, we do think that stoplights do not look red to the defectives. But if spectra vary from normal person to normal person, we would not take that to impugn our judgment that stoplights look red. For what would the alternative be? The Principle of Normality reminds us that we can't suppose that stoplights look green to some normal people or yellow to other normal people. So even if spectra vary from one person to another, we must suppose that stoplights look red to all normals, and that shows that "looking red" is keyed to intentional content and not qualitative content.

I have some agreement with Frege when he says "Often, therefore, a colour word does not signify our subjective sensation, which we cannot know to agree with anyone else's." I would take the "cannot know" to mean in practice rather than in principle, since I think perceptual neuroscience is making great strides in the direction of such knowledge. The Fregean point is very easily available to language users. I became aware of it as a child. And my daughter at age seven commented on first hearing the inverted spectrum hypothesis that it explained why some people didn't have purple as their favorite color.¹⁸

Shifted spectra

The chapter now shifts gears to two arguments for the possibility of spectra that are inverted or shifted. I have not said what sense of "possible" is at issue. The relevant notion, I think, is metaphysical possibility, but there is no need to go into the matter because I will be arguing for a stronger thesis. In the case of shifted spectra, my conclusion will be that there is reason to think they are actual and in the case of inverted spectra I will be arguing for nomological possibility with an eye to actuality.

Thus far, I have argued that we do not take our attributions of looking red to be shown false or otherwise problematic by the discovery of qualia,

WITTGENSTEIN AND QUALIA

thus indicating that our tacit semantic policy is to key “looking red” to intentional contents rather than qualitative contents. But I have not yet argued that there really are states with qualitative contents, that is, I have not argued for qualia. In this section I argue that individual variation in the phenomenology of color vision at least sometimes goes beyond variation in the use of public color terminology. I argue that this fact can be used to argue for the existence of qualia and moreover that its appreciation by us provides further support for the claim of the last section that attribution of looking red is keyed to intentional contents rather than qualia. (See Block 1990, 1999 and Putnam’s example of the Ixxzians (Putnam 1999: 162–66).)

The first premise in the argument is the fact that there is enormous variation in the physiology of color vision in the normal population, which leads to enormous variation in color phenomenology. For example, there are large differences among normal perceivers in peak sensitivity of cones in the retina. There is a 51.5%/48.5% split in the population of two types of long-wave cones that differ by 5–7 nm, roughly 24% of the difference between the peak sensitivities of long and middle wave cones (Neitz and Neitz 1998). This characteristic is sex-linked. The distribution just mentioned is for men. Women have smaller numbers in the two extreme categories and a much larger number in between. As a result, the match on a common test “most frequently made by female subjects occurs where no male matches” (Neitz and Jacobs 1986: 625). (See Block 1999 for further explanation.)

Further, variation in peak sensitivities of cones is just one kind of color vision variation. In addition, the shape of the sensitivity curves vary. These differences are due to differences in macular pigmentation, which vary with “both age and degree of skin pigmentation” (Neitz and Jacobs 1986: 624). Hence races that differ in skin pigmentation will differ in macular pigmentation. There is also considerable variation in amount of light absorption by pre-retinal structures. And this factor also varies with age.

I emphasize gender, race and age to stifle the reaction that one group should be regarded as normal and the others as defective.

Premise 1 said that that there is enormous variation in the physiology of color vision in the normal population which leads to enormous variation in color phenomenology. Premise 2 says that the variation in the phenomenology of color vision must obtain even among those whose use of color terms is the same because the borders for our application of color terminology to objects must to some extent be constrained by needs of communicating and cooperating with other people.

Imagine, contrary to fact, that we were all given rigorous training in the application of accepted color terminology to the point where variation in application of color terms among us was minor. Things we almost all categorize as “borderline between red and purple” will still look quite different to us, just as, in the inverted spectrum scenario, a fire hydrant that the members of the inverted pair both categorize as “red” looks different to them. My premise

NED BLOCK

2 is that to some extent that is the way things are. That is, color terminology is to some extent subject to social and linguistic constraint.

How do we know that? Normal people see colors as having the same similarity relations. In the Farnsworth–Munsell 100 hue test, subjects are asked to arrange 100 chips in a circle with one chip fixed as the starting chip. Normal subjects make nearly the same arrangement (Hilbert and Kalderon 2000). One of the things social and linguistic influence does is to impose a set of categories on that similarity space. Before the introduction of oranges into England, the color we call “orange” was considered a shade of “red.” (The OED’s first listed use of “orange” as the name of a color is from 1512.) Cultural and linguistic influences no doubt affect the borders of our color categorizations. And of course culture also affects our shade categories. I am sure that those who were of school age in the USA after 1949 will agree pretty much on the forty shades of Crayolas introduced then and in use (with more added and a few subtracted in 1958, 1972, 1990, 1993 and 1998) since then, such as turquoise blue – despite variation in the way those colors look to us.

The upshot is that there are likely to be many normal cases of same use of public color terminology, different phenomenology – and that is what I am calling shifted spectra. If red things look slightly different to you than to me, there is no saying that either of us perceives more veridically than the other, since we are both normal perceivers, and so there is no way of capturing the difference in external terms. (A number of replies are canvassed in my 2003.) So there are qualia.

An objection to this line of reasoning: red things may look slightly different to you than to me, but that difference may be describable in ordinary language. For example, a red thing may look yellower to you than to me. Here is an argument to support the objector. Many people report slightly different color vision in their two eyes, e.g. things look yellower through one eye than through the other. (Hilary Putnam mentioned that he sees colors differently in different eyes at the Putnam at 80 conference mentioned in note 1.) Suppose Putnam’s left eye could be copied twice and inserted in both eyes of one member of a set of identical twins at birth, Lefty, and copies of his right eye inserted into Righty. If Putnam sees yellower through the left eye, we have reason to think that Lefty sees yellower than Righty. So here is a difference that is effable.

Suppose Lefty and Righty are looking at a wall painted an orange that is perfectly balanced between yellow and red. One possibility is that Righty says it is pure orange but Lefty says the wall is slightly on the yellowish side of orange. However, another possibility – and one that my argument of this section militates towards as a reality and not just a possibility – is that both Lefty and Righty say and think that the wall is pure orange. One would expect that if both were trained on Crayolas! And there is no reason to suppose that the way it looks to them deviates systematically from their thought and talk. The situation is then that they both see it as pure orange

WITTGENSTEIN AND QUALIA

but one sees it as yellower than the other. How can that be? The best way to understand how that could be is to acknowledge that there are ways of seeing orange, that those ways can differ from person to person. The objection that we are considering notes that we can compare these ways in ordinary language. And here is the reply: yes we can compare them in ordinary language, but that does not show that we can fully capture these contents in ordinary language. Lefty and Righty differ in their color vision, but they may have equal claim to see orange as orange. This very simple argument shows that in a reasonable sense of the term “qualia,” there are qualia.

Points similar to the ones mentioned here were made in Block (2003) where I was arguing (as against representationism) that the phenomenal character of experience outruns its representational content. I claimed that two phenomenally slightly different experiences could nonetheless represent the same color (in different phenomenal ways), even the same minimal shade. Much of that paper was concerned with the issue of whether the phenomenally different experiences that represented the same color could nonetheless be said to have different representational contents. However, the point argued for in this paper is much weaker than the earlier argument, as can be seen by noting that even if the slightly phenomenally different experiences are also different in representational content, they are still qualia if those representational contents cannot be expressed in public language. That is, even if my argument against representationism fails, the argument for qualia in the present sense can succeed.

The picture that emerges is that there is a structured space of color qualia that determines the structure of real world colors themselves. Normal people have pretty much the same similarity relations in their structured spaces (as shown by the experiment with the 100 hue test mentioned earlier) but differ in correspondences between that space and colors in the world, probably because of variation in the retina and in pre-retinal structures. Relations among color qualia are determined by the shared structured qualia space, and so those relations can be expressed in terms of the colors themselves (e.g. “yellower”), but there is no objective “location” of color qualia in real world color space because of the differences among people in correspondence between the space and actual colors.¹⁹

The point about qualia being comparable but not expressible in ordinary language applies also to inverted spectra. In the Lefty/Righty shifted spectrum case, Lefty and Righty differ by a rotation of the color circle – slightly towards yellow in Lefty. In the usual ROYGBIV depiction of the color circle, that corresponds to a clockwise rotation. In the inverted spectrum cases to be discussed below, one twin differs from the other by a 180° rotation, which is the same in both directions.

These points will be the basis of my discussion below of the “Frege–Schlick” view.

To summarize: the Putnam two-eye example suggests that qualia can at least sometimes be compared across people. One person’s color experience

NED BLOCK

is yellower than another's or one person's color experience is "complementary" to another's. Does this show that the contents are expressible in public language after all and so are not qualia? No: qualia can be compared in public language without being fully capturable in public language. Even if your experience is yellower than mine, it can nonetheless be the case that there is no color term *F* such that your experience of red things is describable as "looks *F*."

Behavioral indistinguishability

My arguments for qualia on the basis of the shifted and inverted spectrum do not presuppose that the shifted or inverted perceivers are behaviorally indistinguishable from one another. For example, the 51.5%/48.5% split in the population that I mentioned could, for all I know, determine the extent to which warm colors advance (i.e. project forward) or are exciting compared to cool colors. If half of mankind is inverted with respect to the other half, the inversion shows the existence of qualia even if the inverted spectrum is behaviorally detectable – unless one half is thereby seen to be abnormal or unless the behavioral difference undermines the claim of inversion. To see this, note that the little dialectic I rehearsed earlier about who has authority over the word "red" would apply to the two normal groups as well as to two normal individuals. Neither group could pretend to be "the" group for which red things look red. If it is alleged that for members of one group, red things look green, the other group can complain that the situation is relevantly symmetrical, whether or not there are differences in which colors advance and which recede.

Virtually all arguments that inverted spectra are impossible that I have read or heard appeal to one or another sort of asymmetry in color space.²⁰ (I am talking about arguments that spectrum inversion is impossible, not that it makes no sense or is otherwise problematic, as alleged in the arguments by Robert Stalnaker that will be considered later.) For example, Bernard Harrison (1973) argues that there are more colors between red and green going one way around the color circle (via blue) than going the other way (via yellow). A far more significant asymmetry is that the most saturated colors differ in their lightness levels from color to color. If you decrease the lightness of a sample of yellow you get a different color (brown), but the same is not true for green or blue. The most saturated reds are darker than for other primary colors. Further, desaturated red is pink but desaturated green is green, so hue is not symmetrical with regard to saturation.²¹ These are not just differences in categorization, but also differences in similarity. A light and dark blue look much more similar to one another than do light yellow and dark yellow, i. e. brown (Hilbert and Kalderon 2000). Just how significant such asymmetries are is hotly debated. Stephen Palmer (1999) discusses a far larger number of ways in which the color space can be mapped onto itself than had appeared

WITTGENSTEIN AND QUALIA

in the prior literature, some of which may avoid such problems although Justin Broackes (2007) argues otherwise. In addition, as Shoemaker notes, it is not clear why minor asymmetries ought to be significant given that we could imagine a slight variant of human color vision in which those asymmetries are absent but color experience is much the same as ours. But this argument is unsound according to Daniel Dennett (1991). Dennett notes that red is advancing, warm and exciting, whereas green is receding, cool and calming, claiming that such asymmetries in function are what make the colors look different from one another. So, according to Dennett, Shoemaker's imagined race for whom these "minor" asymmetries are ironed out, could not have color experience much like ours. Others (Block 1994) argue that these differences in reaction may not be intrinsic to color experience, their widespread presence in different cultures being explained by environmental regularities that are not necessary at least to the basics of color experience.²²

But this whole set of controversies can be sidestepped – at least if the existence of qualia is the issue. For as I have noted, the argument for qualia that I am talking about depends not whether the inversion or shift is behaviorally undetectable but whether the groups that are inverted or shifted with respect to others are normal. If it turns out – and as I argued, the evidence points in that direction – that a person's spectrum is very often shifted with respect to his neighbor, I don't think anyone should conclude that this shows that defective color vision is widespread.

In the rest of this chapter, I will discuss an argument that an inverted spectrum is possible. I will focus on an intrasubjective inverted spectrum because of the advantage I mentioned earlier. I will be considering an inversion in which the inverted pair are roughly functionally equivalent but not necessarily exactly functionally equivalent. For example, perhaps one member of the pair experiences red things as exciting and advancing whereas the other experiences them as calming and receding. (Earlier versions are in Block 1990, 1994.)

I say that an argument that an inverted spectrum is possible is an argument for qualia. "But doesn't a possible inverted spectrum show at most that qualia are possible rather than actual?" The argument from the possibility of inverted spectra to qualia is this. Suppose an inverted spectrum is possible but non-actual. If we manage to create an inverted spectrum (a procedure that may work is described below), it would be wrong to suppose that we have created qualia where they did not already exist. The fundamental nature of experience does not change because we create a situation that exhibits a fundamental feature of it! So the possibility of an inverted spectrum is sufficient for the existence of qualia.

Inverted spectra

I have made the case for actual shifted spectra. Now I will argue that inverted spectra are possible and perhaps actual. Let us start with a

NED BLOCK

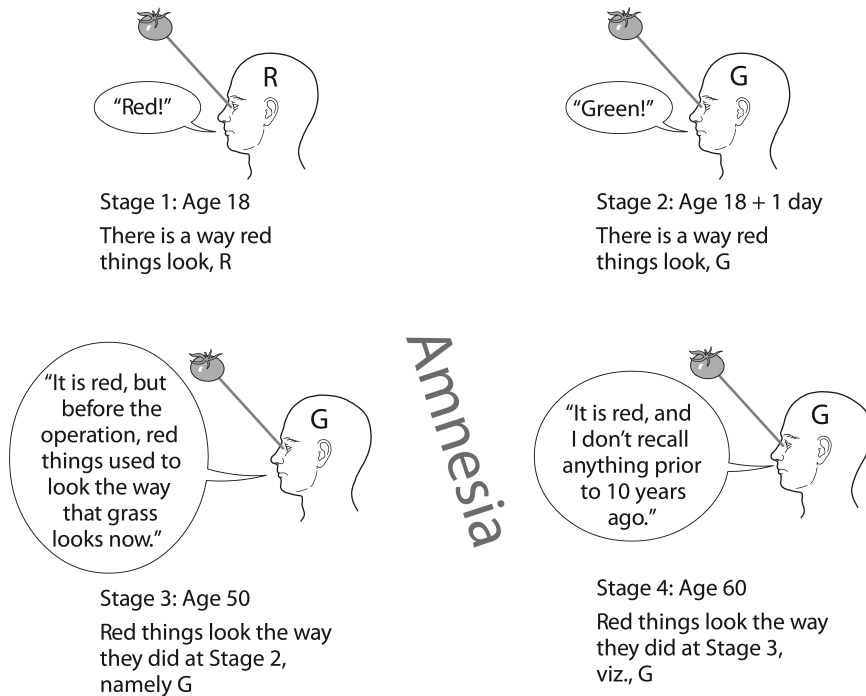


Figure 10.1 Question-begging Inverted spectrum scenario. The captions express the way things look to the subject at each stage in terms of the existence of ways of looking, R and G. R is the way the red tomato looks to the subject at Stage 1. G is the way the red tomato looks to the subject at Stages 2, 3, and 4

question-begging description of the scenario. The point of so doing is to highlight the contrast with another description which is not, I will argue, question-begging. See Figure 10.1.

The captions express the way things look to the subject at each stage in terms of the existence of ways of looking, R and G. R is the way the red tomato looks to the subject at Stage 1. G is the way the red tomato looks to the subject at Stages 2, 3 and 4.

We start at Stage 1 with you at your 18th birthday, at which time you are a normal perceiver. You call red things “red” and red things look red to you. In need of funds, you agree to undergo an experimental color inversion the next day in which a chip is inserted in your optic nerve that changes signals from red things into the signals that would have been produced by green things, and so on for other colors.²³ I have been calling this “crossing the wires in the visual system.” If we were relaxed about what counted as a visual “input,” we could imagine instead virtual reality goggles that do this trick without any surgery (as mentioned earlier).

WITTGENSTEIN AND QUALIA

The operation is a success: at Stage 2, you are disposed to call red things “green” and they look green to you, and so on: you see colored things as having the complementaries of their actual colors. You become famous to the point where you feel that the paparazzi are ruining your life. You change your name, and your appearance with cosmetic surgery and move to another state, trying to blend in. At first, you have to fight your tendency to call red things “green” and your tendency to stop at green lights, to suppose that red tomatoes are unripe and so forth, but after some years have passed, you naturally and spontaneously react to colors in the normal way. You unthinkingly call stop-lights “red” and stop at them. Even though you unthinkingly call stop-lights “red,” at first you think of them as looking green, but your desire to put the operation and its unpleasant sequelae into the past leads you to work to think of stop-lights as looking red. You want to be normal and find it tiresome to be reminded of the operation you had at age 18. By the time you are 30, you spontaneously think of blood as not only being red but also looking red.

How do we know that the colors have not reinverted so that the way things look is the same as at Stage 1? According to the story – which I am claiming is a genuine possibility – you remember very clearly what things used to look like before the operation, and if asked by close relations in private, you say that everything used to look the complementary color of the way it looks now. You can retrieve clear mental images of the look of things at your 18th birthday party, for example. At Stage 3, you are 50, and you haven’t thought about your operation in many years. Someone who recognizes you despite the cosmetic surgery says to you: “Aren’t you the person who underwent the experimental spectrum inversion?” You say: “Oh yes, I haven’t thought about that in many years. But I do clearly recall the look of the ripe tomatoes at my 18th birthday party. They looked to me then the way grass looks to me now, colorwise.” If asked to paint a picture of the way things look to you today, you paint the grass green and the sky blue. If asked to paint a picture of the way things looked to you before the operation thirty-two years earlier, you paint the grass red and the sky yellow (Taylor 1966). Another ten years pass, during which time no one asks you about the operation and during which time you don’t think of the days before 18 or any of the times someone has questioned you about the inversion. At age 60, you develop amnesia for the period up to age 50 (cf. Putnam 1981). Now, and this is Stage 4, you have no memories for the period before your operation, nor do you have any memories for episodes of remembering that period. However, according to the story, the way red things look to you at Stage 4 is the way green things looked to you at Stage 1. If this example were to be used in an argument against functionalism, Stage 1 and Stage 4 have to be functionally equivalent.²⁴ However, if the purpose of the argument is to demonstrate qualia – and that is my purpose here – functional and behavioral equivalence is not required. For example, perhaps at Stage 4, red is no more exciting than green.

NED BLOCK

The way in which this argument is question-begging is summarized in Figure 10.1. The subject at Stage 3 says “I do recall the look of the ripe tomatoes at my 18th birthday party. They looked to me then the way grass looks to me now, colorwise.” But what are these ways and why should we countenance them? And if these suspicious ways are coming in at Stage 3, presumably they must have been involved at the outset, in the argument at Stage 1. The relevant bones of the argument diagrammed in Figure 10.1 amount to this:

- Stage 1: There is a way red things look to the subject, R.
- Stage 2: There is a way red things look to the subject now, G.
- Stage 3: The way red things look to the subject remains G.
- Stage 4: The way red things look to the subject remains G, even though the function of G is similar to (but not necessarily exactly the same as) the function of R in Stage 1. The subject at Stage 4 is a normal perceiver.

The question-begging aspect is the postulation and naming of the ways things look, even at Stage 1. However, ways need not be brought in at the outset. It may seem that in saying that red things look red, I have sneaked in something illicit, maybe something amounting to “ways” for we would not ordinarily say that red things look red. Isn’t that “language going on holiday”? Actually, Wittgenstein is committed to red things looking red at Stage 1 in my view as I will now argue.

Consider Stage 2: red things look green, or as Wittgenstein would put it, at Stage 2, our subject sees red things green. (In Wittgenstein’s version, he says “I see everything red blue today and vice versa.”) But now if asked how the way he sees things now contrasts with the way he saw things yesterday, it is perfectly natural for him to say. “I see everything red as green today, but yesterday I saw everything red as red.” So we are justified in saying that at Stage 1, he sees red things as red. Imagine the subject talking to his doctor where it is better to err on the side of over-explicitness. The doctor is wondering whether there might have been something wrong with the patient even at Stage 1. To be absolutely clear, the patient says that red things look green today but yesterday red things looked red. And that justifies the descriptions given below in Figure 10.2 of Stage 1 and Stage 2.

And once we have gone this far, it is difficult to see how to avoid the descriptions I have given of Stages 3 and 4. If red things look green at Stage 2, then they still look green at Stages 3 and 4. Here is a summary of the essential aspects of the argument.

- Stage 1: Red things look red.
- Stage 2: Red things look green.
- Stage 3: Red things still look green.
- Stage 4: Red things still look green even though the subject calls them “red” and is in other respects a normal perceiver.

WITTGENSTEIN AND QUALIA

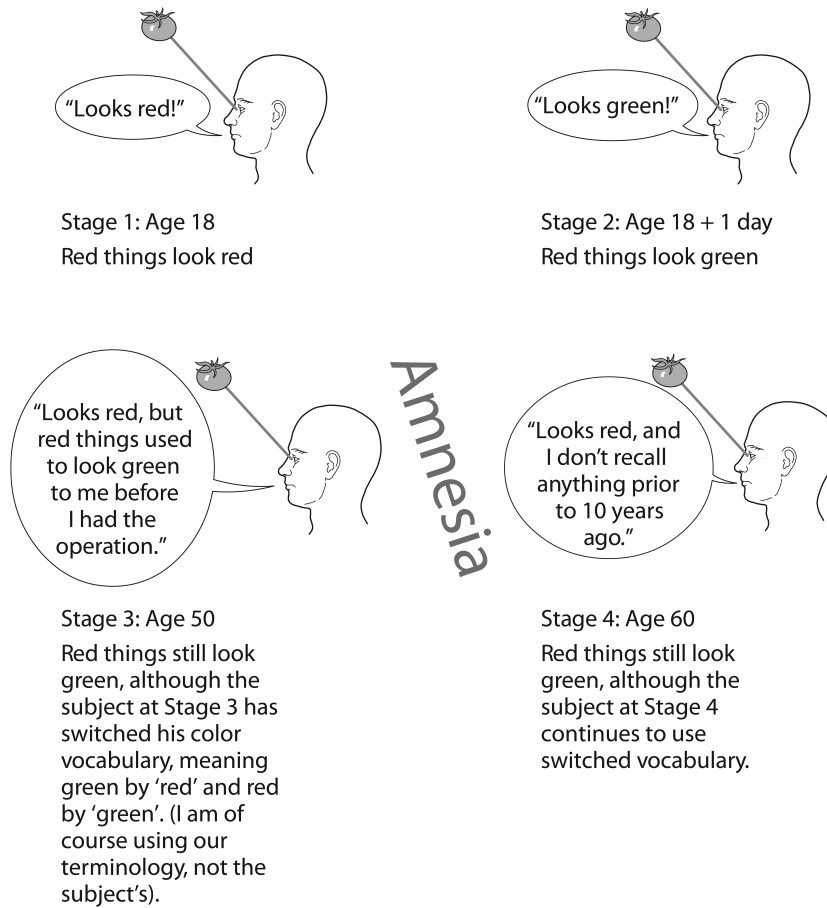


Figure 10.2 Version of the inverted spectrum scenario that is not question-begging, but which violates the Principle of Normality. The same events are depicted as in Figure 10.1, but described without explicit mention of the ways things look or of 'R' and 'G'. The focus is on the colors that things look to have

However, there is a problem with the descriptions in Figure 10.2. At Stage 4, the subject is, I claim, a normal perceiver again. If so, it is false to say that red things look green to him – that would violate the Principle of Normality which says that normal perceivers in normal conditions perceive veridically. And that leads to a further variant of the argument illustrated in Figure 10.3, one that moves back part way to the version given in Figure 10.1.

In this, the final variant, ways are introduced at Stage 4. The idea is that at Stage 4, the subject, who is now a normal perceiver, perceives veridically. White (1986) has suggested that the functionalist should say that once the amnesia hits, the subject's experience instantaneously reinverts. Am I saying

NED BLOCK

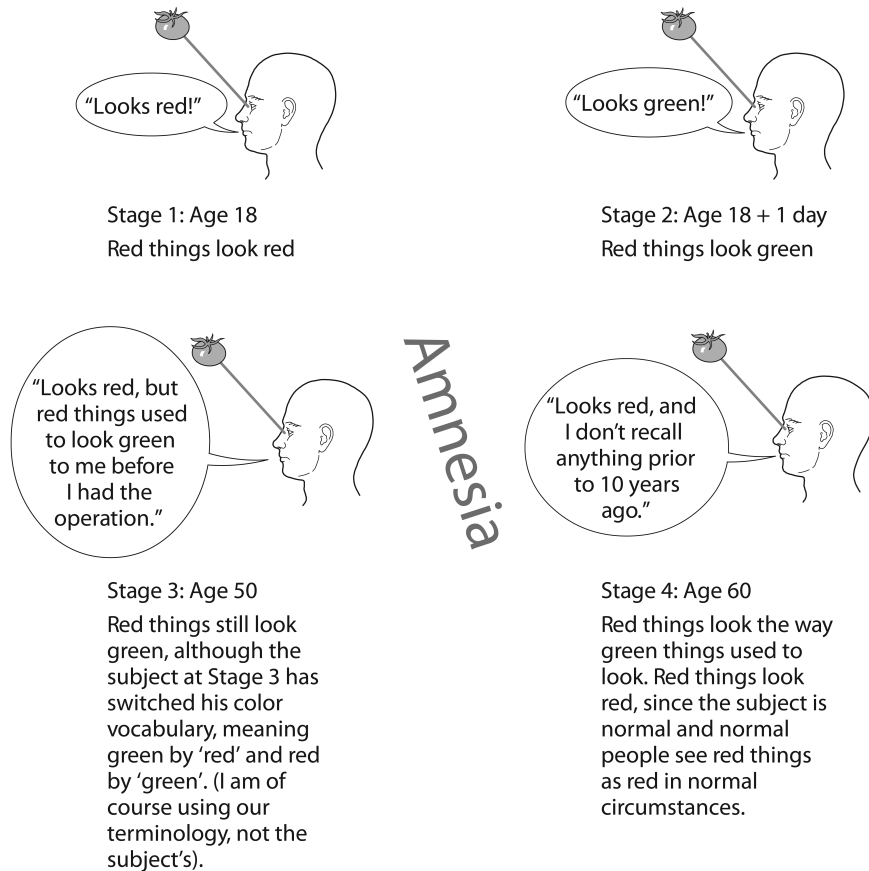


Figure 10.3 Version of the inverted spectrum scenario in which ways are introduced at Stage 4

something similar? Do red things suddenly look red as soon as the subject gets amnesia? No. For an abnormal subject as in Stages 2 and 3, color terms have to be used relative to normality. At Stage 2, the obvious relativity is to Stage 1, just a day earlier, when the subject was normal. But at Stage 3, there is some indeterminacy about how color terms should be used – relative to the normal Stages 1 or 4. Since red things look red at Stages 1 and 4 and since there has been an inversion in between, the most plausible way to accommodate the inversion is in terms of ways. Perhaps ways could be introduced at Stage 3 as well, but in any case Stages 1 and 2 can be described without recourse to ways, so the argument cannot be described as begging the question. The ways at Stage 4 are a conclusion, not a premise. That, in brief, is the argument. In the next section I will elaborate some of the premises.

WITTGENSTEIN AND QUALIA

Premises in the inversion argument for qualia

I have given a way of conceiving of spectrum inversion, but conceiving spectrum inversion does not show that spectrum inversion is possible. It would be agreed by all – even those like Chalmers (2002) who argue that conceivability of the right sort entails possibility – that it is not a straight-forward matter to move from conceivability to possibility. (Gendler and Hawthorne 2002; Hill 1991; Stoljar 2006.) My view is that the most significant reasons for thinking that spectrum inversion is not possible are all reasons that involve the asymmetrical nature of the color solid and thus allege that inversion would have to be behaviorally detectable. But since I am not appealing to a behaviorally undetectable inversion, these arguments do not apply here.

David Cole (1990) appeals to experiments using goggles that invert the retinal image, claiming that these subjects adapt in the sense of experiencing a reinversion of experience. A recent study (Linden, Kallenbach, Heinecke, Singer, and Goebel, 1999) suggests previous reports greatly overplayed the extent to which behavioral adaptation to the goggles involves a reinversion of experience. Cole mentions other experiments which I think are irrelevant to the spectrum inversion case but that I do not have the space here to discuss. (See also Hurley 1998.)²⁵

I said the subject at Stage 4 is normal. But he has undergone an operation in which the wires in his visual system were “crossed.” And what could be more abnormal than that? But wait! “Crossed” is a metaphor. The visual system of the subject at Stage 4 is different from the visual system at Stage 1 as a result of an operation, but why suppose it is Stage 4 that is abnormal rather than Stage 1? An operation can – and often does if successful – restore normality. A hernia is stitched up, a tumor removed, a heart repaired – all cases in which the “after” is more normal than the “before.” If we were to discover that all the rest of us were born with the physiological formation of the visual system that obtains in Stage 4 and that it is Stage 1 that is “peculiar,” we would regard Stage 1 as “crossed.” If we were to find out that our subject at Stage 1 was born with an unusual visual system, we would naturally take the view that the operation “uncrossed” the wires in his visual system, making his visual system normal again. This way of talking makes it look as if the “crossed” visual system must be abnormal, but that is a mistake.

For what would we say if half the population has a visual system like Stage 1 and half like Stage 4? Then we should say that both are normal, that there are two different varieties of normal visual systems. And in that case, both Stage 1 and Stage 4 are normal. The operation has succeeded in changing one normal visual system into another normal visual system. If Stage 1 and 4 are both normal – and possible – there are qualia, since neither our subject at Stage 1 nor our subject at Stage 4 can claim that he uniquely sees red things as red and green things as green.

NED BLOCK

I think the fact that “crossed” visual systems are normal can be shown by example. The example I have in mind was first advanced by Martine Nida-Rümelin for a somewhat different purpose, as I will now explain. (See Nida-Rümelin 1999.) The three kinds of cones in the retina feed to three “opponent-process channels,” as discovered in the nineteenth century by Hering, on the basis of purely behavioral data. The red/green channel operates as follows: If the activation of the R cones is greater than that of the G cones, the red/green channel is excited and the subject sees red (if the other channels are in approximate equilibrium). If the activation of the G cones is greater than the R cones, the red/green channel is inhibited and the subject sees green if the other channels are in approximate equilibrium. The three types of cones send both excitatory and inhibitory signals to three kinds of cells in the structure that is the basis of the next stage of processing beyond the retina (known as the lateral geniculate nucleus). These cells are called “color opponent cells” for obvious reasons.

One form of color blindness occurs when the subject has a genetic defect that substitutes the photo-pigment in the G cones, chlorolabe, for the photo-pigment in the R cones, erythrolabe. With the photo-pigments in the two kinds of cones the same – both chlorolabe – the subject is red/green color blind. It can also happen that another genetic defect substitutes the photo-pigment in the R cones for the photo-pigment in the G cones. In this case, they are both erythrolabe. (See Byrne and Hilbert 2003.) Again, with the photo-pigments in the two cones the same, the subject is red/green color blind. If the two genetic defects occur at once, the subject has the usual G photo-pigment in the R cones and the usual R photo-pigment in the G cones, resulting in normal discrimination. This case – known as pseudo-normal color vision – can be predicted to occur in 14 of 10,000 males, but as far as I know, no such case has ever been noticed. Apparently, the color vision of these people is sufficiently normal so that they do not consult an eye doctor about abnormal color vision.

If the eyes and optic nerves from a pseudo-normal were switched with yours (assuming you are not pseudo-normal), the effect would be – as far as we know – that the red/green opponent cells would be activated in the opposite manner from those in your retina. Red things would look green and green things would look red. The point is that the pseudo-normal is a living example of crossed wires in the visual system.²⁶ What we do not know is whether the development and history of visual experiences of the pseudo-normal yields something different from the imagined transplant. That is, what is controversial is how hard-wired the connection is between the cones and the color-opponent cells and between the color-opponent cells and the rest of the visual system in the brain. If those connections are hard-wired, then we can expect that in pseudo-normals, the red/green channel would be inhibited in input situations in which it is excited in normals and excited when it is inhibited in normals. Then these subjects are red/green inverted with respect to normals such as you and me.

WITTGENSTEIN AND QUALIA

Byrne and Hilbert give 3 arguments against Nida Rümelin:

First, even if pseudo-normal vision actually occurs, its frequency will be very low ... ; thus the possibility of pseudo-normal vision does not show that spectrum inversion might be widespread. Second, there is in any case no reason to suppose that pseudo-normal genes would preserve normal visual pathways: the opponent channels might be switched as well, in which case pseudo-normal subjects would not be red–green spectrally inverted. Third, there is evidence that for the M- and L-cones the development of the retinal circuitry for the red–green opponent channel is insensitive to which pigment the cone contains. In other words, pseudo-normal subjects would just be normal subjects.

(Byrne and Hilbert 2003: 19)

The first point is irrelevant to the issue considered here, where normality rather than frequency is what is in question. On the second point: If Byrne and Hilbert are right that we simply don't know whether the genes for pseudo-normal color vision control the brain channels as well as the eye structures, that is enough for the claim that inversion may be actual and at any rate is possible so far as we know. The third point seems to me to count in the opposite direction of what Byrne and Hilbert say. If the development of the visual system did react differently depending on whether the pigment was erythrolabe or chlorolabe, then that would count against the possibility that the cones are hard-wired into the opponent-process channels. The force of the pseudo-normal configuration depends on pigment-insensitive connections between cones and channels.

Nida-Rümelin and others have used this case in an argument for the possibility of intersubjective inversion. I agree with that argument, but I am using pseudo-normals for a less controversial purpose: as an example of crossed wires in the visual system that are normal. It is a case of crossed wires because if a pseudo-normal's eyes were switched with yours, you (and the pseudo-normal) would experience color inversion as in Wittgenstein's original case of the person who wakes up one morning finding that red things look green and conversely. Pseudo-normals who live with their eyes from birth (and before) might not be inverted with respect to the majority, but if their eyes are switched with yours, the effect will be to stimulate the R–G channel in the opposite from normal way.

So pseudo-normals have crossed wires. But are they normal? How can that be when pseudo-normal color vision is the combination of two genetic defects? Chlorolabe and erythrolabe must be more or less equally suited to the two kinds of cones as attested by the fact that none of the more than 4 million pseudo-normals in the world have come to light – as far as I have

NED BLOCK

been able to determine. A pseudo-normal person can equally say that normals are the result of two genetic defects: the cases are symmetrical. Pseudo-normals are less frequent, but I hope no one will take that to be decisive. If pseudo-normals were the more numerous, no doubt they would be in control of the terminology.

The example of pseudo-normals shows I think that intersubjective inversion may be actual and is for all we know possible. But that is not my main point in introducing it. Recall that we are discussing intrasubjective inversion and the issue at hand is whether the visual apparatus of Stage 4 should be regarded as abnormal solely on the ground of difference from Stage 1. I offer the case of pseudo-normals as normal visual systems that differ in a way that could be like the difference that I am hypothesizing exists between Stages 1 and 4, namely “crossed.”

The argument for qualia is supposed to be mainly based on the possibility of an intrasubjective inverted spectrum. I did detour through the pseudo-normals as an example of intersubjective inversion but also as an example of a kind of “crossed” visual system that is normal. And the argument can make do with the latter. I talked about others who have visual systems like that in Stage 1 and others who have visual systems like that in Stage 4, but without appealing to intersubjective inverted spectrum. The point is that there is no reason why both visual systems cannot be normal and so no reason why our intrasubjective inversion subject has to be abnormal at either the beginning or end of the process.

Sydney Shoemaker and Daniel Taylor (Byrne and Hilbert 2003; Shoemaker 1982; Taylor 1966) have argued from the possibility of intrasubjective inversion to the possibility of intersubjective inversion. They argue that if Fred undergoes intrasubjective spectrum inversion, then the color experience of others must be radically different from Fred’s either before his inversion or after it (or both). My argument is different since my argument does not appeal essentially to intersubjective inversion. An opponent of the Shoemaker–Taylor intra/inter argument²⁷ might acknowledge that crossed wires in the visual system make for color inversion in the intrasubjective case – as is shown by the testimony of subjects who undergo the procedure – while resisting the claim as applied to the intersubjective case. The idea would be that congenital differences of that sort between different groups of normal perceivers are just alternative realizations of the same phenomenology of experience. I think that this refusal to accept intersubjective inversion can be defeated on its own terms, but that is not my main purpose here.

Objection: “Although the result of the operation is a normal visual system, our subject at Stage 4 is abnormal because of the operation which produced it.” Not so, because an operation need not produce anything abnormal. An operation can correct an abnormal defect, yielding normality.

WITTGENSTEIN AND QUALIA

“But the process leading to Stage 4 includes both the operation and the adaptation, and doesn’t that make Stage 4 abnormal?” No: adaptation to an operation is often an essential part of the normalizing process. For example, there is a congenital “heart inversion” condition called transposition of the great arteries in which the outputs from the ventricles are reversed. In the normal heart, the pulmonary artery rises from the right ventricle, taking deoxygenated “blue” blood to the lungs (which is then returned to the left side of the heart via pulmonary veins), whereas the aorta rises from the left ventricle pumping “red” blood to oxygenate the body. In the heart inversion condition, the aorta rises from the right ventricle, pumping “blue” blood to the body, whereas the pulmonary artery arises from the left ventricle, circulating “red” blood unnecessarily to the lungs. The effect is that the left ventricle does very little work, quickly losing muscle mass as a result. If an arterial switch operation is done too late (more than about a month after birth), the left ventricle will fail and the child will die, but if the operation is done early enough, the left ventricle can build up mass after the operation via the process of making new heart cells, an ability that the heart loses. The point is that the operation does not immediately restore normality. First, there is a period of adaptation in which the left ventricle regains the lost cells. But the need for a period of adaptation after the arterial switch operation does not make the child’s heart forever abnormal.²⁸

Taylor & Francis

Memory

Wittgenstein raised the question of how we could distinguish inversion from misremembering, raising the issue of “whether the things stored up may not constantly change their nature” (Shoemaker 1996a: 204). And Dennett (1996a; 1993) has emphasized the issue of the unreliability of memory in arguing against inversion scenarios. Figure 10.4 shows a schematized Stage 1 at the top (that is the topmost of the 4 schematic depictions). Seeing a ripe tomato causes the subject to experience red which causes him to say “Red!,” indicating a normally functioning experience as of red. I symbolize a normal experience of red with “R” as in Figure 12.1. I am assuming, contrary to the intervening argument, and purely for mnemonic purposes, that the inverted subject can be seen as abnormal with an experience symbolized with “G.” (Less mnemonic symbols that don’t make this mistake would be “X” and “Y.”) The second portion shows the subject at Stage 2 where seeing the red tomato causes an experience as of green, which causes the subject to say “Green!,” indicating the complementary function. The bottom half of the diagram indicates two different versions of the situation at Stages 3 and 4. According to the possibility that I have been emphasizing (“phenomenal realism”), at Stages 3 and 4, red things look green. The subject says “Red!” because there has been a terminological inversion that cancels out the color perception inversion. But memory

NED BLOCK

skeptics such as Dennett wonder how we can rule out the case schematized in the last line in which the compensating inversion occurs earlier than I imagined and the subject misreports his experience, remembering falsely what red used to look like.

Dennett's claim is that the two possibilities represented in the two scenarios at the bottom half of Figure 10.4 cannot be distinguished, concluding that they are not really different. (He is explicit about advocating "first person operationalism.") However, he gives no reason to doubt that they can be distinguished empirically, as indicated in Figure 10.5.

The bottom two hypotheses from Figure 10.4 elaborated present implementations of the bottom two scenarios of Figure 10.4. In the top case in Figure 10.5, the possibility that I have been emphasizing is elaborated. The subject at Stage 3 sees a ripe tomato and contrasts the way it looks color-wise with the way ripe tomatoes used to look. The machinery of this comparison involves a comparator device that compares the neural implementation of the experience with the neural implementation of the memory of previous tomato experiences. Of course, this story is physicalistic, but no more so I think than the *Tuesday New York Times* (see note 12). A Wittgensteinian who wishes to deny this level of physicalism would have to adopt a revisionary theoretical perspective, something that a Wittgensteinian should not do.

In the skeptical scenario at the bottom of Figure 10.5, the memory is itself inverted, as is the experience, so when the subject compares them, the subject says the current experience is not what he remembers. The report is

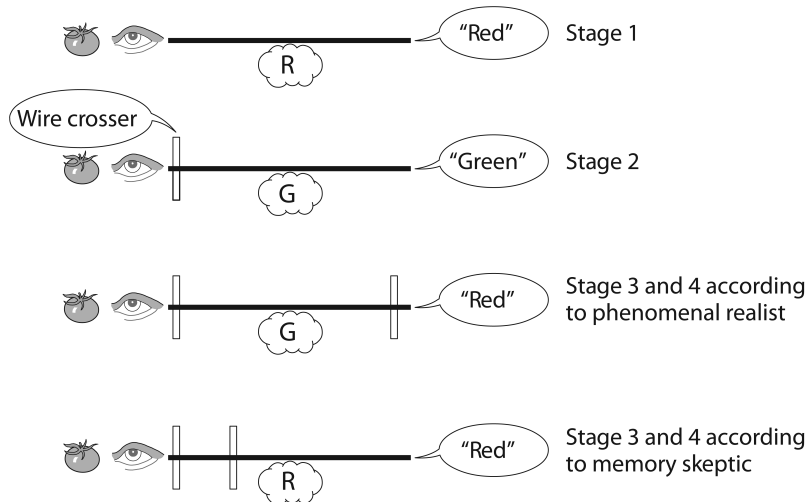


Figure 10.4 The phenomenal realist contrasted with the memory skeptic

WITTGENSTEIN AND QUALIA

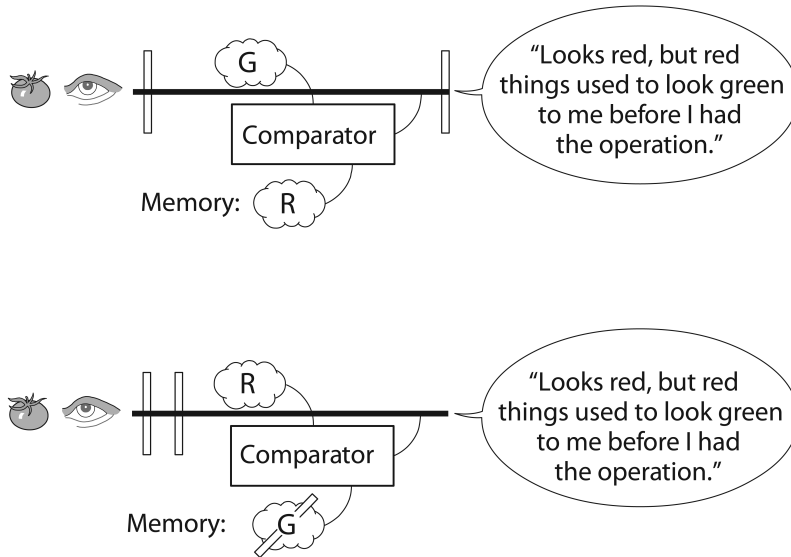


Figure 10.5 The bottom two hypotheses from Figure 10.4 elaborated

the same but the internal reality is not. Dennett wonders how these two hypotheses could possibly be distinguished. Since they cannot be distinguished, he thinks there is no real difference. My response is that there is no reason to doubt that the normal methods of science could distinguish them. For example, perhaps it could be shown that the memory representation does not change in the relevant period. Dennett might object that one can tell whether the memory representation in the brain changes only if one can distinguish the memory representation from the physical basis of the quale itself, whereas his point is that these cannot be distinguished. But since memory and experience are to some extent distinct, they must be distinguishable at the level of the brain. The point of the detail illustrated in Figure 10.5 is that once one makes the hypothesis concrete, the claim that the normal methods of science cannot possibly resolve such issues begins to look like mere skepticism.

The point is further illustrated by a different form of the skeptical hypothesis illustrated at the bottom of Figure 10.6. In this form of the skeptical hypothesis, the memory representation is actually the same as the perceptual representation, but because of an inversion involved in the comparison process, the subject thinks otherwise. Again, why should anyone suppose that the normal methods of science could not find the difference between this case, the one at the bottom of Figure 10.5, and the phenomenal realist hypothesis at the top of both figures?

NED BLOCK

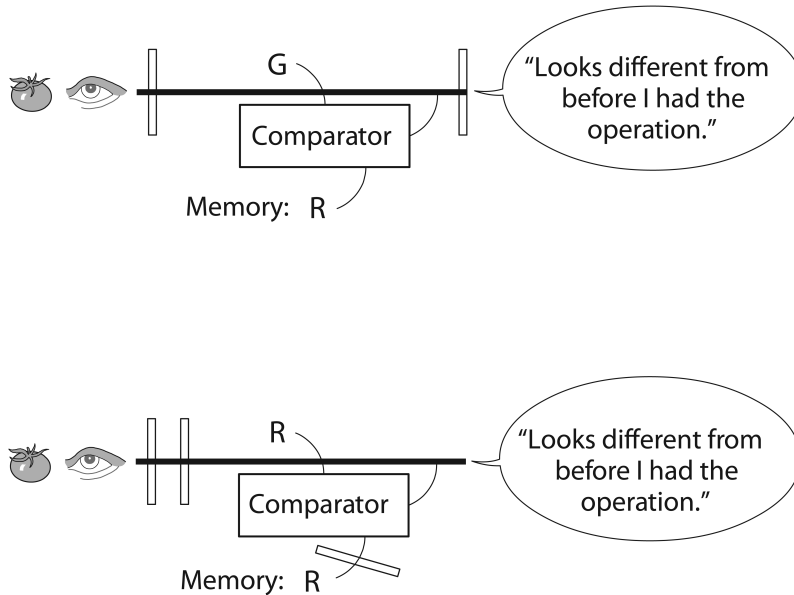


Figure 10.6 The top figure is the same as in Figure 10.5. The bottom figure illustrates a different sort of memory skeptic

Some functionalists hold that the identity of experience and relations among experiences are constitutively tied to memory of those experiences. For example, Sydney Shoemaker (Shoemaker 1966b: 147) says:

The functional role of a quale must surely include the ways in which the instantiation at one time combines with instantiations of the same or different qualia at later times to produce certain effects ... This means that the total realization of a quale will have to include the memory mechanisms by which qualia have the appropriate "downstream effects".

Functionalists often argue that in crucial cases it is arbitrary whether one says that experiences are qualitatively the same or that there is a memory illusion. Shoemaker and Robert Stalnaker (Dennett 1988; Shoemaker 1996a; Stalnaker 2006) have had an interesting debate in part about this issue, but more generally about whether the kind of combination of functionalism and physicalism endorsed by Shoemaker precludes an inverted spectrum. I am not a functionalist (Block and Fodor 1972; Stalnaker 1999, 2006), and so I don't accept the premise of Shoemaker's argument, but this is not the place to go into that issue.

WITTGENSTEIN AND QUALIA

The Frege–Schlick view

One interesting response to inverted spectrum arguments for qualia involves the doctrine that Shoemaker (1975, 1981; 1982, 1984) calls the “Frege–Schlick” view, and the best defense of that view to date is Robert Stalnaker’s (1999, 2006). Shoemaker’s version of the Frege–Schlick view is that relations of qualitative similarity and difference only apply intrasubjectively and not intersubjectively. On this view, there can be a genuine issue about whether my experience of red is qualitatively the same or different from my experience of green now or last week but not whether my experience of red is qualitatively the same or different from your experience of red. On that version of the Frege–Schlick view, qualitative relations do not supervene on any conglomeration of naturalistic properties, and not surprisingly, Shoemaker’s 1982, 1996a and 2006 discussions of the issue reject it.²⁹

I have already given a strong argument for rejecting it, the example of Lefty and Righty mentioned earlier and below.

Stalnaker’s version of the Frege–Schlick view is in a way much less radical because what his version rejects “is only a notion of qualia that is both independent of representational content, and also comparable across persons” (2006: 391). He is happy to acknowledge interpersonal qualitative comparisons so long as they can be cashed out in representational/functional terms. In a way though, Stalnaker’s version is more radical than Shoemaker’s since he rejects intra-personal comparisons of experience over time as well as inter-personal comparisons of experience – if the qualities of the experience cannot be cashed out representationally/functionally.

I said that skepticism about interpersonal comparisons of experience is incompatible with a generally physicalistic world-view. To see this, suppose there are two people who are physical (and functional) duplicates looking at the same object, say the moon, from the same visual angle. Suppose also that they are identical in history of stimulation. Despite identity in everything about them and their environment, current and historical, that could reasonably be supposed to be relevant, the form of skepticism under discussion insists that there is still no issue of fact as to whether their experiences of the moon are the same or totally different. Whether qualia are independent of representational content or not, those duplicates can reasonably be supposed to have highly similar experiences.

Hilary Putnam argued in response to a presentation of this paper at the Putnam at 80 conference (see note 1) that a single sufficient condition – perfect physical identity – does not establish a matter of fact for any other case or give us any reason to think there is any kind of a physically determined metric of similarity. Right, but there is no reason to doubt that normal procedures could not discover the determinants of many comparisons between people.

As I mentioned earlier, Putnam himself has slightly different color vision in his two eyes. Consider Lefty and Righty mentioned earlier, one of whom

NED BLOCK

has copies of Putnam's left eye (inserted at birth), the other of whom has copies of Putnam's right eye. If Putnam sees things yellower through the left eye, we have every reason to think that Lefty will see things yellower than Righty, especially if we can pinpoint whatever it is about the retina or pre-retinal structures in Putnam's two eyes that very likely are responsible for the difference. Further, examination of your eyes and mine could provide evidence that, other things being equal, your color vision is yellower than mine. One can imagine many cases in which one would be reasonably confident about a comparison between people, independently of any representational or functional theory or any other philosophical theory of the nature of experience. In short, we have mundane reason to believe certain interpersonal comparisons of experience. Wholesale rejection of interpersonal comparisons is incompatible with scientifically informed common sense. Stalnaker's view conjoined with this fact entails that some sort of representationism is true. The possibility of an inverted spectrum is incompatible with leading versions of representationism, so not surprisingly, Stalnaker argues against the inverted spectrum.

Stalnaker's argument is based in large part on two analogies, of which I will discuss only one. It depends on a relational theory of space. Consider an analog of an inverted spectrum for space. We can make sense of the idea that this table is moved 3 feet in the direction defined by the arrow from here towards say the Andromeda Nebula. And that this chair is also moved in that direction. But if everything moved in that direction, would the result be a universe in which everything has moved 3 feet from where those things are now? Of course not. What it is for something to move 3 feet in that direction is for it to move 3 feet relative to everything else. If object A moves 3 feet to the left, then when object B moves 3 feet to the left the "everything else" is slightly different from what it was in the case of A's move, so in the end the relations that define position are the same as when we started. The suggestion is that in the intrasubjective inverted spectrum scenario, there is similarly no fact of change except relative to other color experiences. If they all shift, then there is no change at all.

If one does not believe that the qualitative character of experience is entirely determined by relations to other qualitative characters of other experiences, then the analogy will look wrong. The analogy is certainly at variance with some aspects of common sense. For example, I think common sense leaves it open whether a color-blind person can have exactly the same experience of, say, a black-and-white drawing as non-color-blind people have, despite the fact that the color-blind person's "color space" is different. But the most important point is that this is not an issue to be settled a priori or by common sense. The nature of experience – and I would say the same of the nature of space – is to be settled by the best scientific picture.

Shoemaker (2006) makes another point about the analogy. On the relational theory of space that Shoemaker attributes to Stalnaker, diachronic

WITTGENSTEIN AND QUALIA

distance is dependent on synchronic distance. Something at t_2 has moved from where it was at t_1 if its synchronic distances from other things are systematically different in the appropriate ways at t_2 from its synchronic distances from those things at t_1 . But there is no comparable dependence of diachronic qualitative relations on synchronic ones. Shoemaker's example (p. 22) is that if two colored lights are flashed, one after the other, I can compare the experiences of them without comparing either of them with the kinesthetic experiences, smells, sounds and aches and pains that accompanied either one of them. Some functionalists might suppose that there is a dependence on synchronic relations even if I don't have to be aware of it in order to judge the relation between the two lights, but if that is the objection to Shoemaker's point, then once again the defense of Stalnaker's analogy appeals to the conclusion.

Stalnaker (2006) replies that the relational view of space that he defends is not the one that Shoemaker attributes to him. On Stalnaker's version of the relational view of space, diachronic distance is not dependent on synchronic distance. Both kinds of distances are a matter of relations in a coordinate system. Actual objects are used to fix the reference of the coordinate system but not to define it in any stronger sense. On Stalnaker's version of the relational theory, distances are real, but locations are conventional. So, once everything has moved 3 feet in the same direction, all relative distances are the same as at the beginning, and there is no overall motion with respect to any location.

Here is a different analogy. Students of Galileo founded the Accademia del Cimento in Florence in the seventeenth century. Members of the Accademia pioneered the study of heat phenomena, making the first thermometers, an example of which is found in Figure 10.7. These thermometers were marked with gradations visible in the picture, but the gradations were not carefully spaced and there was no attempt made to relate the gradations at one end of the scale to those of the other end or of one thermometer with those of another. At the beginning of their investigations, the experimenters did not know whether an intra-thermometer difference of 5 degrees at one end was in any objective way equivalent to an intra-thermometer difference of 5 degrees at the other end or whether there was any inter-thermometer correspondence between 5 degrees as measured by one thermometer and 5 degrees as measured by another. A big advance was made when thermometers and their scales were calibrated to specific freezing and boiling points such as the freezing and boiling point of water. Now there was a meaning to inter-thermometer comparisons, but it was still unknown whether an analog of Stalnaker's theory of space applied. Perhaps there were facts about the temperature difference between the boiling and freezing point of water, but no "locational" facts about either point. That is, perhaps the only facts of temperature were distances between certain temperature points such as freezing and boiling points for

NED BLOCK

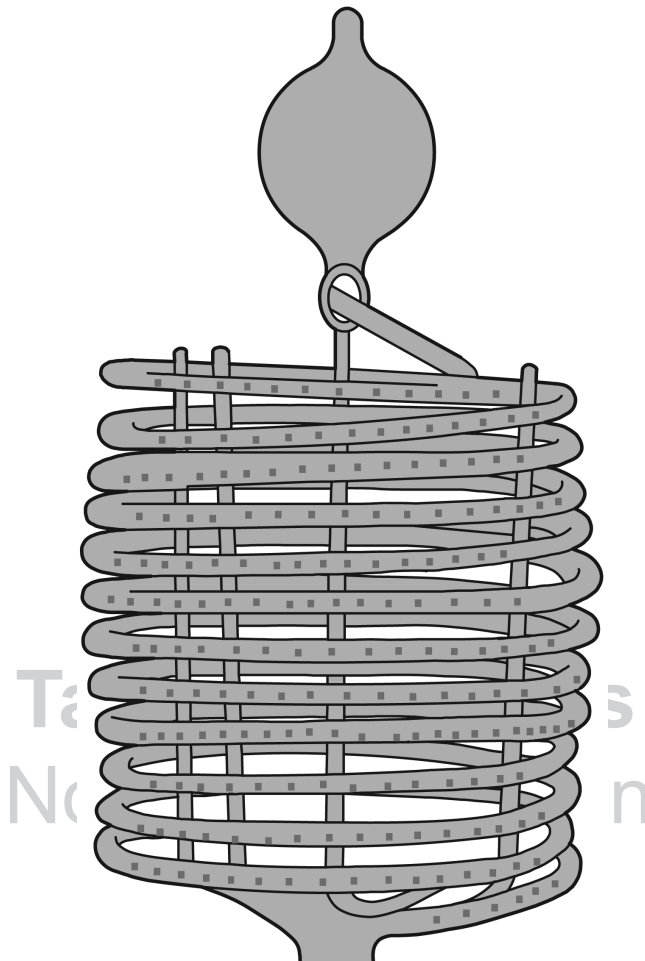


Figure 10.7 Early thermometer

various substances. But as it happened, that turned out false when it was discovered that temperature is mean molecular kinetic energy and hence that there is an absolute zero point at which kinetic energy is zero. That is a “locational” fact that would correspond in the spatial arena to the discovery of a kind of “absolute” space.

Which is the better analogy for phenomenal experience, Stalnaker’s space analogy or the temperature analogy? I don’t see how anyone could think this is an a priori question. There are two accounts, a functionalist account that roughly fits Stalnaker’s view and a physicalist account that roughly fits my view (Stalnaker 2006). I think that there is some preliminary

WITTGENSTEIN AND QUALIA

evidence that favors the physicalist view – some of which I have mentioned here – but there is no way to be sure now.

Conclusion

I started with a shifted spectrum argument for qualia and then moved to an inverted spectrum argument. A simple version of the shifted spectrum argument is that there are well-known features of the retina, lens and other parts of the eye that can make color vision in one eye different from the other, say yellower in one eye. And there can be differences between the eyes of two people of the same sort, so we can have reason to think that I see colors as yellower than you do. Human color vision differs markedly from person to person, but people do arrange chips in more or less the same order, differing in where they draw the boundaries. But since training on boundaries and typical examples of colors has some effect, you and I may both see a chip as focal orange, despite the fact that I see it as yellower. So there must be ways of seeing orange, and those ways are qualia. A key feature of that argument and my inverted spectrum argument was that the versions of the shifted and inverted spectrum needed for qualia do not require behavioral indistinguishability of the shifted and inverted pairs – a point that disarms many objections. The upshot is that Wittgenstein is right to say “This is a very serious situation.”

Notes

- 1 This is a somewhat revised version of an article by the same title that appeared in *Philosophical Perspectives* 21, Philosophy of Mind, 2007. An earlier version was delivered at the “Putnam at 80” conference in Dublin, Ireland, March 14–16, 2007. The chapter was written originally for a conference on the work of Paul Horwich in Pécs, Hungary, May 15–16, 2006. I am grateful to Horwich and Putnam for their responses and to Pierre Jacob, the commentator on my paper at the Putnam conference for his comments and to the audiences there, at the University of Oslo in November, 2006, the University of California at Santa Cruz in April, 2007, and at the Ecole Normale Supérieure in May, 2007. I am also grateful to Paul Horwich, John Morrison and Stephen White for comments on an earlier draft.
- 2 Alex Byrne (2006) and Tim Crane (2000) discuss uses of the term “qualia,” and the term’s history.
- 3 Representationism is the view that the phenomenal character of an experience supervenes on its “referential” representational content, where “referential” is meant to abstract from modes of presentation. (The contents assumed by representationists are not object-involving.) In favor: Byrne 2001; Clark 2000; Crane 2001; Dretske 1995; Harman 1990; Lycan 1996; Thau 2002; Tye 1995. Opposed: Block 2003; Burge 2003; Nickel 2007. Harder to classify: Carruthers 2000; Chalmers 2004; Davies 1997; Kriegel 2002; Levine 2003; Pautz 2005; Shoemaker 1994a.
- 4 Functionalism (Armstrong 1968; Lewis 1970; Putnam 1967) says that pain (for example) can be characterized in terms of its causal relations to sensory inputs, behavioral outputs and to other mental states.

NED BLOCK

- 5 Useful discussions of the inverted spectrum that won't be discussed other than in this note are to be found in Chalmers 1996; Cohen 2001; Cole 1990; Kirk 1982; Lycan 1973, 1996; Rey 1992, 1998; Tye 1995, 2000.
- 6 Wittgenstein's notes have been re-edited by David G. Stern (Wittgenstein 1993) who says that Rush Rhees left out about half the material (pp. 200–201). Since Stern gives the reader much more information about the text, I am using his version. This passage appears on pp. 230–31 of Stern's version and p. 284 of Rhees's version. The square brackets indicate changes in the text by Stern. The bracketed "red" in the first line is accompanied by a footnote saying that the text reads "green" but that this was probably a slip. The 3 dots at the end of the second paragraph are in the text and do not indicate anything left out by me. The "[[284]]" indicates a page number in Wittgenstein's notes themselves.
- 7 Stern says "The text reads 'green,' but this is because Wittgenstein had changed an earlier use of 'green' to 'blue' but failed to make the change here" (p. 231).
- 8 In his introductory notes Rhees (1968: 274) notes:

All that is printed here is a collection of rough notes or memoranda which Wittgenstein made for his own use. He would never have published them – he would not even have had them typed – without revising and rearranging them. Certainly he would have revised the language.

- 9 Recall that the first passage quoted above seemed to involve a shift from the example of red/green to red/blue inversion. I will treat the first passage as if it concerned red and green.
- 10 Michael Tye (2002) has responded to me on this point by an account of the difference as one between an experience that ascribes indeterminacy and one that indeterminately ascribes. This distinction certainly makes sense for verbal representations, but bringing it in to solve this problem begs the question by assuming that phenomenal representations are like verbal representations in allowing for this kind of distinction. This issue has been discussed in varying forms in Block 1983; Dennett 1968; Peacocke 1993; Tye 1995.
- 11 On my view, there are two such epistemic problems. One is the "Hard Problem" (Chalmers 1996): we don't understand why the neural basis of a given experience is the neural basis of that experience rather than another or none. The second problem is the "harder" problem, a version of the problem of other minds. On a naturalistic picture of consciousness, there is no problem about the consciousness of other humans, but when we ask about robots that don't overlap with us in what we take to be the physical basis of our conscious states, there is a problem of how to think about the question of whether they have conscious states (Block 2002).
- 12 More specifically, I am assuming that the phenomenal character of experience supervenes on physical states of the body and, in a concession that makes the view I am assuming a very weak form of physicalism, the physical environment surrounding the body and even the past physical states of that environment. Of course some will balk at even such a weak form of physicalism, but it would be very interesting if it turned out to be incompatible with a Wittgensteinian point of view that supposedly prescind from "theorizing," given that this minimal physicalism is deeply engrained in the day-to-day practice (and not just the theology) of the psychology and neuroscience of perception and in my view in common-sense ways of thinking about perception. (See Nida-Rümelin 1999.) I don't think an approach to the mind that stresses "leaving everything as it is" can lightly require giving up such practices.

WITTGENSTEIN AND QUALIA

- 13 It should also be said that although some kinds of visual experience are localized in the brain, notably experience as of a face or as of motion, the same cannot be said for color. Color appears to be represented at a number of occipital sites, mainly in V4 and V8.
- 14 Note that it is clear from the first passage I quoted that for the kind of inverted spectrum that Wittgenstein is willing to accept, he accepts *both* intrasubjective and intersubjective versions of the innocuous scenario. The individual he describes sees everything differently today than yesterday – that’s the intrasubjective version. And he also endorses the intersubjective version in saying “I think we should under these or similar circumstances be inclined to say that he saw red what we saw blue.” So it does not appear to be the intrasubjective/ intersubjective difference that is at issue between the innocuous and dangerous scenarios.
- 15 I have emphasized the importance of normality and expressibility in ordinary language in arguments for qualia in “Sexism, Racism, Ageism and the Nature of Consciousness” (Block 1999), in which it is argued that males and females, although equally normal, see colors that they classify as red slightly differently. Likewise for old/young and races that differ in pigmentation. (Versions of these arguments transformed for the somewhat different purposes of this chapter will be briefly summarized in the section below on shifted spectra.) And, as mentioned earlier, Horwich (forthcoming) also argues that the difference between an innocuous inversion and a dangerous one is expressibility in public language by normal perceivers.
- 16 In Block (1990) I inadvisedly called violations of the Principle of Normality the fallacy of intentionalizing qualia.
- 17

Suppose everyone had a box with something in it: we call it a ‘beetle’. No one can look into anyone else’s box, and everyone says he knows what a beetle is only by looking at *his* beetle. – Here it would be quite possible for everyone to have something different in his box. One might even imagine such a thing constantly changing. – But suppose the word ‘beetle’ had a use in these people’s language? – If so it would not be used as the name of a thing. The thing in the box has no place in the language game at all; not even as a something: for the box might even be empty. No, one can ‘divide through’ by the thing in the box; it cancels out, whatever it is.

(Wittgenstein 1958: §293)
- 18 The Fregean view has some affinities with that of Paul Churchland. Churchland concedes the conceivability of an inverted spectrum, allowing that different perceivers might have qualitatively different experiences which all have the functional role characteristic of the experience of red.

Such intrinsic qualia merely serve as salient features that permit the quick introspective identification of sensations, as black-on-orange stripes serve as a salient feature for the quick visual identification of tigers. But specific qualia are not essential to the type-identity of mental states, any more than black-orange stripes are essential to the type-identity of tigers.

(Churchland (1984: 40; see also Churchland & Churchland 1981; Clark, 1985)

What I don’t understand about Churchland’s view is how he can maintain functionalism despite allowing that qualia escape the functionalist net. In the

NED BLOCK

- end, I suspect that his view is really a mixed functionalist/physicalist view similar to that of Shoemaker (1982): functionalism supplies conditions of similarity and difference of qualia, but physicalism supplies the natures of specific qualia.
- 19 One complication that I have not mentioned is that the structure of qualia space contains four privileged planes that correspond to the “unique” colors, e.g. reds that are not at all yellowish or bluish. See Pautz (2005).
 - 20 One argument against inverted spectra that is not based on an appeal to asymmetries is given in Byrne & Hilbert (1997). However, that argument is directed towards the conjunction of the possibility of an inverted spectrum together with the thesis that there is a failure of match between the phenomenal character of a sensation and its intentional content. It is not an argument that purports to show that inverted spectra are impossible.
 - 21 For more details see Broakes (2007); Byrne (2006); Churchland (2002); Hardin (1997); Hilbert and Kalderon (2000). A slightly more accurate description of the point about red and green is that desaturated near-red is pink but a desaturated near-green is near-green rather than green.
 - 22 One point on the side of an undetectable inverted spectrum is that the undetectability can be relative to culture and environment. That is, what is required for an undetectable inverted spectrum is that there is an environment – including cultural environment – in which there are two individuals (or stages of a single individual) for whom there is no possibility of detecting an inversion. If those two subjects are undetectably inverted with respect to one another, then they cannot express the difference in public language, even if in a different environment and culture they would have been detectably inverted. For example, consider a people who live underground in circumstances of little light, never seeing the sky and in which there are few stable colors – other than black and white. We would also have to imagine that the culture has a prohibition against seeing the unclothed human body, one’s own or that of others. The point is that the environment can be restricted in ways that restrict the application of color asymmetries that would show up in less restricted circumstances.
 - 23 I speak as if there is a single kind of inversion but actually there are many different types of “inversion” in the sense of different ways to systematically map the color solid onto itself without disturbing relations among points in the solid. See Palmer (1999).
 - 24 The spectrum inversion argument against functionalism that I have given in other places is a straightforward extensional argument. The functional state that functionalists must associate with an experience can occur without the experience, and conversely. Of course, it is always open to the functionalist to complicate the functional role so as to avoid the counterexamples. But so doing will in effect bring in physicalist ideas in an ad hoc manner, and so the resulting view will be an ad hoc combination of functionalism and physicalism. Patricia and Paul Churchland (1981) argue that although spectrum inversion is clearly possible, nonetheless functionalism wins. Their reason is that a functionalist natural kind will be more natural than a qualia-centered natural kind, the latter being unnatural from the point of view of neuroscience. However, they do not take on the issue of whether the ad hoc changes I just mentioned will push the functionalist kind below the qualia-centered kind in naturalness.
 - 25 One example: subjects do adapt – experientially – to goggles that are colored with one color on one side and another on the other side. But in that case, there is a line in the middle of the visual field where the colors meet and that is the kind of defect that the visual system is built to adapt to.

WITTGENSTEIN AND QUALIA

- 26 Let me remind the reader that I am assuming a weak form of physicalism in which phenomenal character is supervenient on the physical body and its surrounding current and past environment. See note 12.
- 27 Called the “intra-inner” argument in Shoemaker (1975).
- 28 I am grateful to Dr. Peter Lang, Department of Pediatric Cardiology at Children’s Hospital, Boston, for filling me in on this issue. The description in the text concerns “dextro” or “complete” transposition of the great arteries. See the Wikipedia entry at http://en.wikipedia.org/wiki/Dextro-transposition_of_the_great_arteries
- 29 Shoemaker’s 1984 discussion takes it much more seriously.

Bibliography

- Armstrong, D. M. (1968) *A Materialist Theory of the Mind*, London: Routledge and Kegan Paul.
- Bacon, J., K. Campbell, and L. Reinhardt (eds.) (1993) (eds.) *Ontology, Causality and Mind: Essays in Honour of D. M. Armstrong*, Cambridge: Cambridge University Press.
- Block, N. (1978) “Troubles with Functionalism,” *Minnesota Studies in the Philosophy of Science* 9, 261–325.
- (1980) “Are Absent Qualia Impossible?,” *Philosophical Review* 89, 257–74.
- (1983) “Mental Pictures and Cognitive Science,” *Philosophical Review* 93, 499–542.
- (1990) “Inverted Earth,” in J. Tomberlin (ed.), *Philosophical Perspectives* 4. Atascadero, CA: Ridgeview, 53–79.
- (1994) “Qualia,” in S. Guttenplan (ed.), *A Companion to Philosophy of Mind*, Oxford: Oxford University Press, 514–20.
- (1995) “On a Confusion about a Function of Consciousness,” *Behavioral and Brain Sciences* 18(2), 227–47.
- (1996) “Mental Paint and Mental Latex,” *Philosophical Issues* 7, 19–49.
- (1999) “Sexism, Racism, Ageism and the Nature of Consciousness,” *Philosophical Topics*, 26(1 and 2), 39–70.
- (2002) “The Harder Problem of Consciousness,” *Journal of Philosophy* 99(8), 1–35.
- (2003) “Mental Paint,” in M. Hahn and B. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, Cambridge, MA: MIT Press, 165–200.
- Block, N., and J. A. Fodor (1972) “What psychological states are not,” *Philosophical Review*, 81, 159–81.
- Boghossian, P., and J. D. Velleman (1989) “Color as a Secondary Quality,” *Mind* 98, 81–103.
- (1991) “Physicalist Theories of Color,” *Philosophical Review*, 100, 67–106.
- Broackes, J. (2007) “Black and White and the Inverted Spectrum,” *The Philosophical Quarterly*, 57(227), 161–75.
- Burge, T. (2003) “Qualia and Intentional Content: Reply to Block,” in M. Hahn and B. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, Cambridge, MA: MIT Press, 405–16.
- Byrne, A. (2001) “Intentionalism Defended,” *The Philosophical Review*, 110, 199–240.
- (2006) “Inverted Qualia,” in E. N. Zalta (ed.), *Stanford Encyclopedia of Philosophy* (Vol. Winter 2006 Edition). Stanford, CA: The Metaphysics Research Lab of the Center for the Study of Language and Information
- Byrne, A., and D. R. Hilbert (1997) “Colors and reflectances,” in A. Byrne and D. R. Hilbert (eds.), *Readings on Color*, Volume 1: The Philosophy of Color: MIT Press.
- (2003) “Color realism and Color Science,” *Behavioral and Brain Sciences*, 26, 3–21.

NED BLOCK

- Carruthers, P. (2000) *Phenomenal Consciousness: A Naturalistic Theory*, Cambridge: Cambridge University Press.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- (2002) “Does Conceivability Entail Possibility,” in T. Gendler and J. Hawthorne (eds.), *Conceivability and Possibility*, Oxford: Oxford University Press, 145–200.
- Chalmers, D. (2004) “The Representational Character of Experience,” in B. Leiter, *The Future for Philosophy*, Oxford: Oxford University Press.
- Churchland, P. (2002) “Brain-Wise: Studies in Neurophilosophy,” Cambridge, MA: MIT Press.
- Churchland, P. M. (1984) *Matter and Consciousness*, Cambridge, MA: MIT Press.
- Churchland, P. M., and P. S. Churchland (1981) “Functionalism, Qualia and Intentionality,” *Philosophical Topics*, 12, 121–32.
- Clark, A. (1985) “Spectrum Inversion and the Color Solid,” *Southern Journal of Philosophy*, 23(4), 431–43.
- (2000) *A Theory of Sentience*, Oxford: Oxford University Press.
- Cohen, J. (2001) “Color, Content and Fred: On a Proposed Reductio of the Inverted Spectrum Hypothesis,” *Philosophical Studies*, 103, 121–44.
- Cole, D. (1990) “Functionalism and Inverted Spectra,” *Synthese*, 82, 207–22.
- Crane, T. (2000) “The origins of Qualia,” in T. Crane and S. Patterson (eds.), *History of the Mind–Body Problem*, New York: Routledge, 169–94.
- (2001) *Elements of Mind*, Oxford: Oxford University Press.
- Davies, M. (1997) “Externalism and Experience,” in N. Block, O. Flanagan and G. Güzeldere, *The Nature of Consciousness: Philosophical Debates*, Cambridge MA: MIT Press, 309–27.
- Dennett, D. C. (1968). “The nature of images and the introspective trap,” *Content and Consciousness*, London: Routledge.
- (1988) “Quining Qualia” in A. Marcel and E. Bisiach (eds.), *Consciousness in Contemporary Science*, Oxford: Oxford University Press, 381–414.
- (1991) *Consciousness Explained*, Boston: Little Brown.
- Dretske, F. (1995) *Naturalizing the Mind*, Cambridge, MA: MIT Press.
- Farrell, B. (1950). “Experience,” *Mind*, 59, 170–98.
- Field, H. (1973). “Theory Change and the Indeterminacy of Reference,” *Journal of Philosophy*, 70(14), 462–81.
- Frege, G. (1884/1953) *The Foundations of Arithmetic*, Oxford: Blackwell.
- Gendler, T., and J. Hawthorne (2002) *Conceivability and Possibility*, Oxford: Oxford University Press.
- Hardin, C. L. (1997). “Reinverting the Spectrum,” in A. Byrne and D. R. Hilbert (eds.), *Readings on Color*, Volume 1: *The Philosophy of Color*, Cambridge, MA: MIT Press, 289–301.
- Harman, G. (1990) “The Intrinsic Quality of Experience,” *Philosophical Perspectives*, 4, 31–52.
- Harrison, B. (1973) *Form and Content*, Oxford: Blackwell.
- Hilbert, D. R., and M. E. Kalderon (2000) “Color and the Inverted Spectrum,” in S. Davis (ed.), *Color Perception: Philosophical, Psychological, Artistic and Computational Perspectives*, Oxford: Oxford University Press, 187–214.
- Hill, C. S. (1991) *Sensations: A Defense of Type Materialism*, Cambridge: Cambridge University Press.

WITTGENSTEIN AND QUALIA

- Horwich, P. (in progress) "The 'Mystery' of Consciousness," a chapter of an unpublished draft of a book on Wittgenstein.
- Hurley, S. (1998) *Consciousness in Action*, Cambridge, MA: Harvard University Press.
- Kind, A. (2003) "What's so Transparent about Transparency," *Philosophical Studies*, 115(3), 225–44.
- Kirk, R. (1982). "Goodbye to Transposed Qualia," *Proceedings of the Aristotelian Society*, 82, 33–44.
- Kriegel, U. (2002) "Phenomenal Content," *Erkenntnis*, 57: 175–98.
- Levine, J. (1991). "Cool Red," *Philosophical Psychology*, 4, 27–40.
- Lewis, D. (1970) "How to Define Theoretical Terms," *Journal of Philosophy*, 67, 427–45.
- Linden, D. E. J., U. Kallenbach, A. Heinecke, W. Singer, and R. Goebel (1999) "The myth of Upright Vision. A psychophysical and functional imaging study of adaptation to inverting spectacles," *Perception* 28(4), 469–81.
- Lycan, W. G. (1973) "Inverted Spectrum," *Ratio*, 15, 315–19.
- (1996) *Consciousness and Experience*, Cambridge, MA: MIT Press.
- Moore, G. E. (1903) "The Refutation of Idealism," in *Philosophical Studies* (pp. 1–30), Totowa, NJ: Littlefield, Adams & co. (1965), p. 25.
- Nagel, T. (1974). "What Is it Like to Be a Bat?" *The Philosophical Review*, 83(4), 435–50.
- Neitz, J., and G. Jacobs (1986) "Polymorphisms of long-wavelength cones in normal human color vision," *Nature*, 323, 623–25.
- Neitz, M., and J. Neitz (1998) "Molecular genetics and the biological basis of color vision," in W. G. Backhaus and R. Kliegle and J. S. Werner (eds.), *Color Vision: Perspectives from Different Disciplines*, Berlin: De Gruyter, 101–19.
- Nickel, B. (2007) "Against Intentionalism," *Philosophical Studies*, 136(3), 279–304.
- Nida-Rümelin, M. (1999) "Intrinsic Phenomenal Properties in Color Vision Science: A Reply to Peter Ross," *Consciousness and Cognition*, 8, 571–74.
- Palmer, S. (1999) "Color, Consciousness and the Isomorphism Constraint," *Behavioral and Brain Sciences*, 22, 923–43.
- Pautz, A. (2005) "Sensory Awareness is not a Wide Physical Relation," *Nous*, 40, 205–40.
- Peacocke, C. (1993) "Review of M. Tye, *The Imagery Debate*," *Philosophy of Science*, 60, 675–77.
- Putnam, H. (1967) "The nature of Mental States," in W. Capitan and D. Merrill (eds.), *Art, Mind, and Religion*, Pittsburgh: Pittsburgh University Press, 37–48.
- (1981) *Reason, Truth and History*, New York: Cambridge University Press.
- (1999) *The Threefold Cord: Mind, Body and World*, New York: Columbia University Press.
- Rey, G. (1992) "Sensational Sentences Switched," *Philosophical Studies*, 68, 289–319.
- (1998) "A Narrow Representationalist Account of Qualitative Experience," *Philosophical Perspectives*, 12, 435–58.
- Rhees, R. (1968). Notes for Lectures on "Private Experience" and "Sense Data": Note on the Text, *The Philosophical Review*, 77(3), 271–75.
- Shoemaker, S. (1975) "Functionalism and Qualia," *Philosophical Studies*, 27, 291–315.
- (1981) "Absent Qualia Are Impossible: A Reply to Block," *Philosophical Review*, 90: 581–99.
- (1982) "The Inverted Spectrum," *The Journal of Philosophy*, 79, 357–81.
- (1984) "Postscript to 'The Inverted Spectrum'," in S. Shoemaker (ed.), *Identity, Cause and Mind*, Cambridge: Cambridge University Press, 351–57.
- (1994a) "Phenomenal Character," *Nous*, 28, 21–38.

NED BLOCK

- (1994b) “Self-knowledge and ‘inner sense’,” *Philosophy and Phenomenological Research*, 54, 249–314.
- (1996a) “Intrasubjective/intersubjective,” in S. Shoemaker (ed.), *The First-person Perspective and other Essays*, Cambridge: Cambridge University Press, 141–54.
- (1996b) *The First-Person Perspective and Other Essays*, Cambridge, UK: Cambridge University Press.
- (2006) “The Frege–Schlick View,” in J. Thomson (ed.), *Content and Modality*, Oxford: Oxford University Press, 18–33.
- Stalnaker, R. (1999) “Comparing Qualia Across Persons,” *Philosophical Topics*, 26, 385–405.
- (2006) Shoemaker, “The Frege–Schlick View,” in J. Thomson (ed.), *Content and Modality*, Oxford: Oxford University Press, 254–57.
- Stoljar, D. (2006) *Ignorance and Imagination*, Oxford: Oxford University Press.
- Taylor, D. (1966) “The Incommunicability of Content,” *Mind*, 75, 527–41.
- Thau, M. (2002) *Consciousness and Cognition*, Oxford: Oxford University Press.
- Tye, M. (1995) *Ten Problems of Consciousness*, Cambridge, MA: MIT Press.
- (2000). *Consciousness, Color, and Content*, Cambridge, MA: MIT Press.
- (2002) “Representationalism and the Transparency of Experience,” *Nous*, 36(1), 137–51.
- White, S. (1986) “Curse of the Qualia,” *Synthese*, 68, 333–68.
- Wittgenstein, L. (1958). *Philosophical Investigations*, 2nd edn. Oxford: Blackwell.
- (1968) Notes for Lectures on “Private Experience” and “Sense Data,” *The Philosophical Review*, 77(3), 275–320.
- (1993) Notes for Lectures on “Private Experience” and “Sense Data,” in J. Klagge and A. Normann (eds.), *Ludwig Wittgenstein: Philosophical Occasions 1912–1951*, Indianapolis: Hackett Publishing Company, 202–89.

Not for distribution