

Experience and Representation

1. Qualia, or phenomenal properties of experience, pose a deep and persistent problem for Materialism. Consider the particular throbbing character of a pain in one's toe after stubbing it, or the way a gleaming red fire engine looks in the sunlight. In both cases there is a strong intuition that we are enjoying experiences - conscious, mental states - which can be typed by their qualitative character: the throbbing of the pain or the reddish character of the visual experience. In this sense the qualitative character of an experience, a quale, is a property of the experience. The question then immediately arises: what sort of property is this?

In particular, we might begin by wondering whether qualia, if they are properties of experiences, are intrinsic properties of those experiences. At first blush, they seem to be. But if we accept this, we have a problem. Assuming we are materialists, and assuming that it is our brains that house our experiences, all properties of experiences must be properties of the brain, or of states of the brain. If we claim that qualia are intrinsic properties of experiences, then they must be intrinsic properties of brain states. But which ones? Well, presumably, the best candidates are those that are coinstantiated with qualia. That is, find out through probing the brain which of its neurophysiological properties are instantiated when one is having a throbbing pain in the toe, or a visual sensation of bright red, and then identify these qualia with their neurophysiological correlates. To have a throbbing pain in the toe is just to have these neurons fire in this way.

Once put this starkly the problem is evident. While many philosophers resist this

sort of identity theory because they believe in the possibility of multiple realization - the idea that creatures quite different in their physical constitution could share qualitative experiences - it seems to me that the implausibility of the view goes deeper than that. There is something just utterly arbitrary about the idea that these neurons firing in this way should constitute a toe-throbbing, and these others a reddish visual experience. In fact, it is precisely because of the arbitrary and unmotivated character of the identification of the quale with its neurophysiological correlate that multiple realization itself seems so plausible in the first place.

As soon as one thinks about what really is significant about these particular neurons firing in this particular way it becomes clear that it isn't anything about their intrinsic properties. Rather, their significance derives from their location in a system of neurons, and the significance of their firing as they do derives from the relations that hold between their firing that way and other neurons firing as they do, and ultimately to the way all the neural firings connect to stimuli and responses. After all, what else could it be? Thus the natural way to understand the connection between the neurophysiological properties and the mental properties is just as those who point to multiple realization would have it: the neurophysiological properties *realize* the qualia. On this view, then, qualia are role-properties; roles that are implemented in us by our brain states.

The view that qualia are roles overcomes the problem of the apparently arbitrary connection between an experience's qualitative character and its neurophysiological correlate, since the relation between a role and its implementer is far from arbitrary.

The problem on this side of the coin is that it is very hard to see how qualia can be identified with role-properties. There is a clear intrinsic “feel” to them, and thought experiments involving qualia inversions give expression to this deep intuition. After all, what is supposed to be the role with which I identify the throbbing character of my pain in the toe when I stub it? Certainly it is associated with various dispositions, such as the disposition to hop around holding the toe, to swear loudly, not to pay attention to other matters, etc.. Also, it is clearly a property I am disposed to instantiate when my toe is stubbed, and certainly it maintains various dispositional relations to other mental states I am prone to occupy. Nevertheless, it is a long way from the obvious claim that throbbing pains-in-the-toe bring in their trail all sorts of dispositions to the claim that they just are a particular bundle of dispositions. As I say, there is too strong an intrinsic “feel” to them for such a claim to command intuitive assent.

I am not going to argue further against the role view of qualia, since this is not my target in this paper. I bring it up mostly to motivate another view, one that’s called “intentionalism”, or “representationalism”. On this view, qualia are not to be identified either with functional roles or with their internal, neurophysiological implementers. Rather, qualia are the intentional contents of sensory states. The qualitative character of a sensation is the property that sensation represents its intentional object as having. On this view, then, the sorts of dispositions mentioned above with which a throbbing pain-in-the-toe is normally associated turn out to be, as one would have thought, only contingently related to the sensation itself. On the other hand, the connection between the sensation’s qualitative character and its neurophysiological implementation is also

contingent, since presumably intentional content is possessed only contingently by the representational vehicles bearing that content.

There are of course many versions of intentionalism, and different ways to categorize them. For my purposes the following two divisions are significant: between wide and narrow, and between external and internal. Wide intentionalism is the doctrine that qualitative character is to be identified with wide intentional content, whatever that turns out to be. Obviously, then, narrow intentionalism is the doctrine that it is narrow content with which we identify qualitative character. For the most part I will only be concerned with wide intentionalism, but below I will have a bit to say about the narrow version. Most adherents of wide intentionalism are also externalists, in the sense that the wide contents they attribute to qualitative experiences are the relevant physical properties of the distal objects of perception. But it is possible to be a wide intentionalist and also an internalist, as will be discussed below. Until further notice, however, read “intentionalism” in what follows as wide externalist intentionalism.ⁱ

Now intentionalism, of course, does not vindicate the intuition that qualia are intrinsic properties of experiences. Depending on how one interprets the thesis, they are either relational properties of experiences or intrinsic properties of external objects. So, suppose my mental state, R, is the visual experience I’m having while looking at a red fire engine. On one way of understanding the intentionalist thesis, the qualitative character of R is identical to R’s representing the redness of the fire engine. Thus it is a property of R alright, but not an intrinsic one, since representing what it represents is not an intrinsic property of a symbol. On another way of understanding the intentionalist

thesis, however, the qualitative character of the experience just is the redness of the fire engine. If, for argument's sake, we allow that redness is itself an intrinsic property of the fire engine, then the qualitative character of the experience turns out to be intrinsic after all; however, it also turns out not really to be a property of the experience.ⁱⁱ Either way, then, the intentionalist is committed to some deviation from the intuitive starting point. Nevertheless, the fact that an apparently intrinsic property is involved - the redness of the fire engine - does seem to allow the intentionalist to accommodate intuition to an important extent.

There are two other significant virtues of intentionalism. First, it unifies our conception of the mind. Many philosophers follow Brentano in identifying intentionality as the "mark of the mental". Yet it is a standard feature of most discussions of the mind-body problem to say that there are two fundamental phenomena that define mentality: intentionality and consciousness. It would obviously be more satisfying, and lead to a more thorough understanding of the mind, if it could be shown that consciousness is in some way reducible to intentionality. Identifying qualia with the intentional contents of sensory states would clearly go a long way toward realizing that hoped-for reduction.

Second, there is anyway reason to think that qualitative character and intentional content are not totally separable. As Charles Siewert (1998) has argued convincingly, in at least many cases it is hard to make sense out of the idea that two states could share qualitative character yet differ in their intentional contents.ⁱⁱⁱ For instance, consider what it's like to perceive a rectangular shape. Could it be just like what that's

like, and yet one is representing the shape as circular? That seems almost impossible to conceive. Well, if qualitative character, what it's like, were just a matter of what one's visual experience was representing, then we would have a straightforward explanation for why such a situation was inconceivable.

Despite its virtues, I think intentionalism, at least in its wide externalist form, is untenable. In section 2 I will show why. In the course of doing so I will discuss in some detail a recent defense of the doctrine by Michael Tye. In section 3 I will take up again some of the reasons for thinking some version of intentionalism must be right, and see if any version can escape the argument of section 2 and also meet certain other plausible constraints.

2. Though intentionalism is not subject to certain problems with inversion scenarios that plague functionalism, there are still some thought experiments, and some involving inversions, that undermine the view's plausibility. Intentionalism can escape the inversion objections that involve switching internal roles, since qualitative character, on this view, is independent of a state's internal causal role (at least to a large extent). But so long as we can coherently imagine switching external roles, as it were, intentionalism will conflict with intuition. So, for instance, take Ned Block's () example, "Inverted Earth". On Inverted Earth all the colors are inverted with respect to colors on Earth. We are then asked to imagine that a traveler from Earth, call her Janet, with inverting lenses implanted in her eyes, goes to Inverted Earth to live. To Janet, given the two inversions that cancel each other out, everything would look normal. Yet, the argument goes, since her internal states would covary with the complements of the colors with which

they covaried on Earth, their intentional contents would eventually change. On the intentionalist view, then, Janet's qualia should change too; but this seems absurd, since she notices no change in her qualia.

One serious problem with the example lies in its reliance on the claim that the intentional contents of Janet's visual experiences would shift as a result of her sojourn on Inverted Earth. Of course one can make a case that the contents would shift after enough time, but then again one can also make a case that they wouldn't. In particular, one might appeal to the evolutionary history of Janet's perceptual mechanisms, maintaining that the contents of her visual qualia are fixed by the properties with which they covaried in the conditions under which they were selected. Thus the intentionalist can escape the consequence that Janet's qualia have changed without her knowledge.

The underlying intuition that Block's example is meant to express is that how things seem to us - what it's like for us - cannot be a matter of how things are outside us. This is not to deny that how things are outside us can cause things to seem a certain way; obviously it's normally because I am exposed to light of a certain sort that I seem to see something red. Rather, the point is that how it is with me experientially cannot be constituted by how things are outside my mind. (I say "my mind" here to allow for the possibility that, for all I know, the core realizers of my mental states might extend beyond my body. I doubt it, but that's an empirical question.)

We can avoid the escape hatch in Block's example by imagining that Inverted Earth is also like Twin Earth. It contains molecule-for-molecule duplicates of ourselves, except for one difference. Our twins evolved with inverting lenses. Thus when Twin-

Joe looks at a fire engine, which on Inverted Twin Earth is green, his brain states will be identical to mine when I look at a fire engine, which on Earth is red. In this case it is hard to see how to avoid the consequence that Twin-Joe's visual state has the content that the fire engine is green while mine has the content that the fire engine is red. The intentionalist then has to say that, despite our internal similarity, our experiences have quite different qualitative character.

The problem with this version of the example, however, is that the intentionalist is quite willing to endorse the claim that the two twins are having experiences of different qualitative character. According to the intentionalist, just as we are willing to attribute propositional attitudes with different contents to two molecularly identical twins in the original Twin Earth case, so too should we be willing to attribute sensory experiences with different qualia to the twins in the Inverted Twin Earth case. Why, they ask, should we treat the cases differently?^{iv}

In a recent discussion of this issue Michael Tye (1998), a defender of intentionalism, notes that the standard reply to the intrasubjective version of the Inverted Earth argument, one he himself gave in Tye (1995), isn't adequate. Remember that the intentionalist reply to Block's original case was to say that the intentional contents of Janet's visual experiences wouldn't change as a result of her sojourn on Inverted Earth because intentional content is fixed by reference to the environment in which the relevant perceptual mechanisms are selected. The problem with this reply is that it doesn't deal with the case of Swampman moving to Inverted Earth.

Swampman, as you may remember, is Davidson's (1987) creature who

spontaneously coalesces out of swamp gas into a molecular duplicate of a normal man. It sure seems like such a creature would have qualitative experiences of a determinate type - it would have reddish experiences when looking at a red fire engine - yet, since it/he was not the product of selection, on the view we just presented it appears his visual states would lack intentional content. But if there's no intentional content then there's no qualitative content, and that contradicts our initial assumption.

Initially one might try adopting a split theory: let selectional environment determine content for those creatures having a selectional history and let synchronic causal covariation determine content for those without such a history. This way Swampman gets to have qualitative experiences. However, we then have a problem with Swampman when he travels to Inverted Earth. We can now run Block's original argument on Swampman and we can't say in reply that the intentional contents of his visual states don't change as a result of the move. So now what?

Before turning to the solution Tye proposes, I want to pause to consider more closely the basis of the strong intuitions that drive the objections to (the externalist version of) intentionalism. As I've argued elsewhere (Levine 1998), I think the matter ultimately comes down to the question of self-knowledge. The problem of self-knowledge can be put this way. I know what it's like for me right now as I visually experience the world around me. In particular, I know what it's like for me to see colors. However, if intentionalism is correct, and if the intentional contents of my color experiences are identified with the surface properties of distal objects, then it seems I couldn't really know how it is with me without knowing about these surface properties.

But this seems absurd. Looking back at Block's Inverted Earth argument, we can see a similar appeal to self-knowledge at work. Why, after all, is it supposed to be so absurd to think that Janet's qualia have changed without her noticing? Well, because what it's like for her is precisely the sort of thing to which she has privileged access. A radical change of visual qualia is something she would know about.

As soon as we bring up the question of self-knowledge, however, the intentionalist can reply as follows. According to widely accepted theories of the contents of thoughts and beliefs, these contents are (at least partly) determined by the natures of the external objects with which we causally interact. The point of Twin Earth is to demonstrate that I and Twin-Joe have thoughts with different contents - with different truth conditions - when we think about our respective samples of the local colorless, odorless liquid. Now some have indeed objected that on this theory of content we cannot account for our apparent ability to know the contents of our own thoughts - to know, as one might put it, what we are thinking. The point is this. If one isn't bothered by this consequence for thought, then why does it matter so much for qualia? Don't we have as much good reason to believe that we know what we're thinking as we have to believe that we know what it's like for us to see red?^v

There is of course an extensive literature on the topic of externalism and self-knowledge. For present purposes, I want to focus on one line of argument in particular, due to Burge (1988), to the effect that there is no conflict between an externalist theory of content and self-knowledge. Though he doesn't present it this way, I think Tye's response to Block's Inverted Earth argument can best be seen as an application of the

Burgean strategy.

Actually, one can run the Burgean argument two ways.^{vi} The first goes like this. All states of knowledge, including introspective knowledge, are themselves propositional attitude states. As such, they have canonical linguistic expressions. (If one adheres to a “language of thought” model, then they are literally embodied in functional relations to linguistic tokens.) So, in a typical instance of introspecting one’s own thought, say about water, we would have two thoughts: first, the thought “there’s some water in the glass in front of me”, and second, the thought “I’m now thinking about the water in the glass in front of me”, or something of that sort. The second thought expresses my state of self-knowledge, and it counts as knowledge just in case it is true and also meets whatever reliability and justification conditions are required for knowledge. What’s crucial, however, is that since the expression of my self-knowledge uses the very same term “water” as is used by the thought that is the object of that knowledge, there is no problem from externalist considerations. After all, when introspecting about my water thoughts I don’t think about H₂O and XYZ, at least not under those descriptions, but about water. Whatever “water” refers to in the object thought, it also refers to in the meta-thought, the expression of self-knowledge.

On the second interpretation, which is arguably closer to what Burge had in mind, the relation between the object thought and the introspective thought about it is much closer than identity of vocabulary. Rather, the idea is that the introspective thought actually embeds the object thought - the latter, as a token entity (state, event, or whatever) actually constitutes a proper part of the former. When I’m aware of thinking

about water there aren't two tokens of my representation of water - one in the original thought and the other in the introspective awareness of the original - but rather there is one thought about water and some sort of demonstrative attention mechanism trained on it. That is, I use the very thought token to represent itself as the object of my awareness of it. On this view, where there is only one token of "water" involved in my awareness of myself as thinking about water, there is even less room for slippage, and thus even less room for externalist worries to undermine self-knowledge.

Let's apply this model now to the case of qualia. I'm looking at a red fire engine, and thus having a reddish visual sensation. I'm also currently aware of the reddish character of my visual sensation. The worry is that if intentionalism is correct and my visual sensation's being reddish depends on its representing red, a particular property of the surfaces of physical objects, then I won't be able, from the inside, to know the qualitative character of my experience. The reply then is to invoke either the embedding relation or the same vocabulary relation. My awareness of what it's like for me now as I gaze upon the red fire engine is itself a representational state, and its content is determined in precisely the same way as the content of the original visual experience is. So either it directly includes the original experience as a proper part, or it is expressed by the very same concepts/representations. Either way, whatever determines the intentional content of the experience will determine the intentional content of the introspective awareness. Hence, no room for doubt about what I'm experiencing.

Let's return now to Inverted Earth. Remember, the original Inverted Earth

argument had Janet traveling from Earth to Inverted Earth and then remaining there for the rest of her life. Block claimed that by virtue of her long sojourn on Inverted Earth the intentional contents of her visual states would change to reflect the nomological dependencies in her new environment. The example is supposed to be an embarrassment to the intentionalist because it would entail a change of qualia without her noticing. We then noted that the intentionalist need not accept Block's claim about the switch in content in the first place. However, as Tye argues, this reply doesn't cover the case of Swampman.

But now Tye argues as follows. Why not just say that indeed the qualia have changed, and so have the contents of Janet's introspective judgments? When she now judges that she is having an experience of a reddish sort, that represents what used to be represented by her judgment that she was having an experience of a greenish sort. But she doesn't notice any change, one might argue. She doesn't form an introspective judgment to the effect that her qualia have changed. Well, the intentionalist argues, that's because she misremembers. Block of course stipulates in his example that Janet suffers from no failure of memory. But depending on how one takes the stipulation, it's either question-begging or insufficient to block the reply. If what he means by a failure of memory is some internal malfunction of the mechanisms responsible for preserving information, then indeed he can stipulate that no such malfunction occurs in Janet. Yet there is still a way for her to misremember. If her current states, due to changes in her environment, no longer represent what they used to, then when she purportedly remembers now what it was like to see a fire engine back on Earth, she is mistaken. It's

not through any fault with her memory mechanisms, but it's a mistake nonetheless. To rule out this sort of mistake as well, of course, begs the question against the intentionalist.^{vii}

To be clear about what's being proposed here, let's imagine Janet's situation as vividly as possible. Here she is now on Inverted Earth, looking at a green fire engine, which, because of her inverting lenses, is causing her visual system to respond with the same state (characterized neurophysiologically) as it did when she looked at a red fire engine on Earth, without the lenses. She is now asked whether fire engines look the same to her now as they did when she was a child. She then forms an image of a fire engine she remembers seeing as a child - let's say it was a memorable incident. So, she's got before her mind a current visual impression and this memory image, and she finds them relatively indiscernible (up to differences in shade, etc.). It sure seems as if she's making a judgment about the intrinsic properties of the two experiences. Can the intentionalist resist that characterization of the situation?

What the intentionalist has to say - what Tye, for one, does say - is that the memory image is wrong. That's not the way the fire engine looked to Janet as a child. Given that the memory image now represents the fire engine as green, and when she originally saw the fire engine she correctly perceived it as red, her memory is mistaken. Of course, as Tye argues, it still is the case that this image is an image of the original fire engine, but nevertheless it, as it were, misdescribes it. The moral of the story is clear. If phenomenal changes track intentional changes, then they do so all the way up the chain of introspective awareness.

So, does this reply work? We started out with a strong conviction that it made no sense to attribute a radical change in Janet's qualia without her awareness. But now, under the pressure of the Burgean strategy, as applied by Tye to this case, we see that what it is to notice a change is itself a cognitive state with an intentional content that is subject to the same changes to which the qualia themselves are allegedly subject. Thus our initial conviction is suspect.

In fact, I think there are two cases that we can still press against the intentionalist that cannot be accounted for by the Burge/Tye strategy. First of all, let's reconsider Swampman. According to Tye, we can avoid the counterintuitive consequence that Swampman fails to have any qualitative experiences so long as we adopt a synchronic covariation account of content. Swampman's problem is that he has no history - no selectional history, in particular. But so long as his perceptual states currently maintain lawful relations with distal properties he can count as having experiences.

But now consider not Swampman, but Swampbrain-in-a-vat. That is, imagine that a brain molecularly identical to my brain coalesces out of swamp gas inside a vat. What do we say about the contents of those of its brain states that correspond to my perceptual states? Since it has no sense organs we can't appeal to the synchronic lawful relations between these brain states and the distal properties of objects. In fact, depending on how we decide to hook it up, we can imagine that these states could covary with any number of different distal properties. Now, when we think about a normal human brain in a vat we can solve the problem of content by appeal to selectional history. Swampbrain, however, has neither selectional history nor currently

maintained covariation relations. Thus, it seems, Tye and other intentionalists have to say that Swampbrain lacks experiences.

Is this such a problem? After all, Swampbrain is a pretty outlandish case, so it may seem as if biting the bullet and maintaining that he/it lacks experiences isn't really so bad. In fact, it seems pretty clear that this is the position that Dretske takes. However, there are good reasons for not biting this bullet, and, what's more, it's clear from Tye's remarks that he in particular doesn't want to take this option. Remember that what motivated his new response to the Inverted Earth problem was precisely his refusal to deny Swampman experiences. The very same considerations, it seems, would militate against denying Swampbrain experiences as well.

But leaving aside questions of consistency, there is a very strong argument for granting qualitative experience to Swampbrain. As I sit here typing at my computer, looking out at the room full of furniture, feeling the air around me, hearing various sounds, I know in some sense that it is possible this is all an illusion. That is, after describing the case of Swampbrain, I realize that, for all I know, I am Swampbrain. I'm not claiming I don't know I'm really in a room, typing on a computer. I'm quite happy to acknowledge that claims to knowledge do not need to rule out all possibilities. My point is only that there is a robust sense in which this possibility is not something I can rule out. On the other hand, if someone were to suggest that it doesn't even appear to me as it does, I'm not even having visual and auditory experiences, I couldn't take that possibility seriously at all. There does seem to me to be a significant, and undeniable asymmetry between the two sorts of possibility, and it's one that can't be

accommodated by the (externalist) intentionalist who is willing to go intentional all the way up.^{viii}

There's another way to put the same point, and in this way it connects quite directly with a discussion by McLaughlin & Tye (1998) of the conflict between externalism and self-knowledge with respect to belief. If externalism is right about qualitative character, and if I can know I am having a qualitative experience merely by having it, or, if you like, in an *a priori* manner, then it seems I can know *a priori* that I'm not a Swampbrain. But whatever your favorite theory is about how I can defeat the skeptic, it better not attribute to me *a priori* knowledge that I'm not a brain in a vat.

Now it's interesting to note how McLaughlin & Tye respond to the corresponding challenge about belief, for it seems to me that it conflicts with what Tye wants to say about experience. So, suppose one argues as follows. I can know *a priori* that I have a concept of water. My concept of water is an atomic natural kind concept, and therefore I couldn't have it were there not to exist any water. Therefore, I can know *a priori* that water exists. But obviously I can't know *a priori* that water exists, hence the existence of water can't be a necessary condition for my having an atomic natural kind concept of water.

Their reply to this argument is long and involved, but for our purposes the crucial part is what they say about the epistemic position of Toscar, Oscar's molecular twin on Dry Earth, where there isn't in fact any water, though Toscar is under the illusion there is. They say (McLaughlin & Tye 1998, page 306), "It is consistent with [externalism]...that Toscar has a concept that has exactly the same (narrow) conceptual

role as Oscar's concept of water." What's crucial here is the claim that though Toscar's concept must be different from Oscar's, he still has a concept and it shares a crucial element with Oscar's, namely narrow content. It seems to me that if we adopt the same line in response to the possibility that I'm Swampbrain, we should say this. I can know, from inside, *a priori*, that I'm having an experience, which includes what it's like for me. On this basis alone, I can no more rule out the possibility that I'm Swampbrain than can Toscar rule out that he's on Dry Earth. Of course there is some way of categorizing the experience such that if I'm Swampbrain I'm having an experience of one sort and if I'm who I think I am I'm having an experience of another sort, and this has to do with the (wide) representational content of the experience. But either way, there's something I have *a priori*, or privileged, access to, and that's what it's like for me. That aspect of my experience can't depend on what's out there.

One problem, then, for the Burge-Tye defense against the self-knowledge objection to intentionalism revolves around the case of Swampbrain. In this case the issue is one of the presence or absence of qualitative character. But there is another, perhaps more serious, problem that stems from the possibility of "Frege cases" for qualia. That is, it seems possible that there could be creatures who standardly respond to a certain distal stimulus with different perceptual experiences, and only through empirical investigation come to learn that their phenomenally different perceptual states are tracking the same distal property.

Before describing such a case, let me explain why I think this possibility really undermines the (wide externalist) intentionalist position. In any case of representation,

there are three components in play: the representation itself, the thing (or property) represented, and a mode of presentation. On some views of direct reference the mode of presentation turns out to be quite thin; perhaps in some sense it reduces to the representation itself, together with some causal relation between the representation and what's represented. Still, we can distinguish these three roles in the representational situation.

What's at issue between the intentionalist and the internalist about qualia is where to locate qualitative, or phenomenal character. Is it what's represented, or is it mode of representation, or perhaps even the representation itself? (Either of the latter two options is consistent with an internalist view, so let's just collapse them into the mode of presentation option.) Here's where the "Frege test" comes into play. Can the subject fail to determine a genuine case of identity without suffering a failure of her cognitive mechanisms? If so, then we're dealing with what's represented, not the mode of presentation. But if not, then we're dealing with mode of presentation.

Notice how it works with Venus and water. Can I simultaneously think of Venus in two different ways and not realize that I'm thinking of the same thing? Of course I can. Can I think of water in two different ways and not realize I'm thinking of the same thing? Again, of course I can. Water, the intentional content of my "water"-thoughts and of my "H₂O"-thoughts, is not graspable by me in a way that shields me from mistakes of this sort, even if my internal cognitive mechanisms are functioning perfectly. Obviously the same can be said of my cognitive grasp of Venus.

Suppose, on the other hand, someone were to claim that I could simultaneously

entertain two thoughts involving the same concept, or term, such as “water”, or involving the same mode of presentation of water, and yet not realize - what’s more, be unable to realize from the resources available from within my mind - that they are the same? I don’t want to claim that there isn’t some way of so describing the situation that it makes sense, but I’m not sure what it would be. My bet is it would involve some bizarre malfunction of my cognitive mechanisms. For after all, what individuates a representation, or its mode of presentation, if not the subject’s ability to discern identity or difference? I can be fooled by accidents of interaction with the environment into thinking the Morning Star is distinct from the Evening Star, but I can’t be wrong in the same way that my “Morning Star”-thoughts are distinct from my “Evening Star”-thoughts.

Okay, let’s bring this to bear on the case of qualia. If color qualia are modes of presentation of colors, then it makes perfect sense that there could be circumstances in which I simultaneously entertain two distinct representations of the same color, involving distinct phenomenal properties, and not realize that it is the same color - i.e. the same property of the surfaces of external objects. For example, suppose some creature had eyes on the sides of its head, like fish, so their visual fields did not overlap. Suppose further that the lenses of the two eyes were color-inverted with respect to each other. We can also imagine that the creatures have their heads fixed in one position, so they never look at the same objects with both eyes. On this scenario, they could be simultaneously looking at two surfaces with the very same reflectance properties, but not know it. What’s more, after painstaking scientific research, they might come to

discover this is the case. They might exclaim: “Wow! What looks green to this eye is really the same color as what looks red to that eye!”

However, if we identify the phenomenal characters of the experiences with what’s represented, then we have to say that in this case the two phenomenal characters would be identical, though the creatures judge them to be different. Note, I’m not relying on any heavy-duty doctrine of privileged access that takes mistakes about one’s own experiences to be logically impossible. Perhaps there can be such mistakes. What I’m insisting, however, is that when everything inside is working as it’s supposed to and I judge, after due reflection and attention, that I’m having experiences of two distinct types, then it’s an odd view, to say the least, that claims I would be wrong about this. Of course if qualia are the modes of presentation of the colors, not the colors themselves, then this all makes sense.

What’s crucial about the Frege test is that it deals with type-distinct but simultaneous mental states. The Burge-Tye reply that attributes the same contents to introspective state and to object state works when what’s at issue is the identity of a single state, or a comparison involving memory. But how can it help when I judge of two simultaneous presentations that they are distinct? Am I mistaken or not? The intentionalist about qualia has to say that my introspective judgment that these are experiences of distinct types is mistaken. I don’t really see any way around that.^{ix}

3. If what I’ve argued so far is correct, then it seems as if the intentionalist program, at least in the form it’s usually professed, is in conflict with self-knowledge of qualia. However, I mentioned above that there is a positive argument for intentionalism, one

that seems to pull our intuitions strongly in its direction. I borrow this argument from Byrne (forthcoming), though I've simplified it somewhat.

The idea is this. If intentionalism is correct, then it is impossible for there to be a difference in qualitative content without a difference in intentional content. The anti-intentionalist, on the other hand, claims that it is possible for different qualia to have the same intentional contents. So, consider the following case. I am looking at a red fire engine during the temporal interval t to t_1 . From t until t' , a point in between t and t_1 , my experience has a reddish qualitative character. At t' the qualitative character changes from reddish to greenish, and remains so until I look away at t_1 . Let us imagine that I am attending to my visual experience and clearly notice the change. It seems clear that corresponding to the change in qualia is a change in intentional content. That is, what seems to be happening, from what I can tell visually, is that the fire engine has changed color. Of course I may have good reason to doubt that this really happened, and so I might not believe it changed color. It might be that someone shined colored light on it, or that I've suffered a malfunction in my visual system. Whatever I happen to think about what's happened, however, it's clear that it will appear to me as if the fire engine has changed color; and it's how things appear, not what I think, that counts as the intentional content of my visual state.

Of course the anti-intentionalist can admit this much but still maintain that intersubjectively it is possible for two qualitatively different states to have the same intentional content. This is indeed what the Inverted Twin Earth scenario seems to involve. Byrne, however, argues that if intrasubjectively it's true that there's no

qualitative difference without an intentional difference, then it holds for the intersubjective case as well. He reasons as follows. Suppose it were possible for two subjects to undergo experiences of different phenomenal types but with the same intentional content. Just imagine a third subject who undergoes both of these phenomenally different experiences, but with the same contents. Yet this is just the intrasubjective case we agreed was not possible.

Now I think one can object to Byrne's extension of the intrasubjective case to the intersubjective one on various grounds, though, for reasons I'll make clear presently, I'm not inclined to. One might argue as follows. Visual qualia are the modes of presentation of visually detected properties like color. So within a subject a distinction in modes of presentation will indicate a distinction in what's presented, while the same distinction between subjects need not indicate a difference in what's presented. Different subjects can instantiate different mappings from mode of presentation to what's presented. All that matters, for perception to work properly, is that within a subject there be no ambiguity.

Despite the availability of this response, as I said above, I'm not inclined to avail myself of it. To see why, consider again my case of the inverted-fish-head, which I used to show the applicability of the Frege test to the relation between qualia and colors. According to my argument above, in this case it makes sense to say that the two eyes are representing the same external colors via distinct modes of presentation which the creatures involved can discover empirically. But of course this is precisely the sort of case Byrne has in mind when constructing his argument that intersubjective difference

in qualia entails a difference in intentional content.^x What's going on, then? How can the very same case be used both as support for and a counterexample to intentionalism?

There is in fact a curious intuitive tension present in the fish-head case. On the one hand, it certainly seems to make sense to say that unbeknownst to the subject the two visual experiences are about the same objective color - a surface reflectance property. This appears to cut against intentionalism. Yet, in line with Byrne's argument, it also seems right to say that in some important sense the two experiences are representing the world differently to the subject. The two objectively red surfaces - each seen through a different eye - appear to have different properties to the subject. As we might say, space appears differently filled in on the two sides of the head. This of course seems very much like a difference in intentional content. Can we reconcile these two intuitions?

It's important to note of course that there is nothing like a contradiction here. One can quite consistently maintain that the fish-head case is a counterexample to the position that color qualia are reducible to (or supervenient on) the representation of objective, physical colors, and also maintain that nevertheless color qualia are supervenient on intentional contents. Byrne himself is quite careful to separate the general position of intentionalism from the particular externalist version that identifies the relevant intentional contents with physical properties of distal stimuli. Still, if, contrary to most intentionalists about qualia, we opt not to identify the relevant contents with physical properties of distal stimuli, then how do we characterize the intentional

contents in question?

Of course one might just refuse to take the problem seriously in the first place and insist instead that we treat qualia strictly as modes of presentation, not as what's presented. This, after all, was supposedly the lesson of the fish-head argument. The idea was that since type-distinct qualia can coherently pick out the same distal property, we have reason to identify the qualitative character of a visual experience with the mode of presentation of the object of the experience. If one wants to count this as a version of intentionalism one can: this would be narrow intentionalism, described above.

Having just argued against an externalist version of intentionalism, I readily agree that qualitative character seems to be determined by what is internal to the mind, and so to that extent am friendly to the idea that it is identifiable with narrow content. However, identifying qualitative character with narrow content doesn't really address the argument above. To the extent that one finds it compelling that the two visual fields of the fish-head present different appearances, in the sense of different ways the world appears, this must mean something more than differences in mode of presentation. That is, it seems to involve a difference in the way things are represented as being, not merely a difference in the representations of the way things are. Differences in narrow content are of the second sort, not the first. So again, how can we characterize the intentional content - wide, not narrow - of visual experiences in such a way as to accommodate both the intuition that qualitative character supervenes on it and also that qualitative character supervenes on what is internal to the mind?

One possibility, one that Byrne explicitly mentions, is to treat qualia as

representations of sense data. That is, the green and red that apparently fill the two regions of the visual field on each side of the fish-head are indeed distinct properties, just not the physical properties of the surfaces perceived, but mental properties of a mental visual field. Similarly, when experiencing a bodily sensation such as a pain in one's toe, the qualitative character of the pain does reduce to a representation, but the property represented - the painfulness - is itself an irreducibly mental property. Notice how on this view the Frege test does not intrude, since presumably the same mental property cannot appear under distinct modes so that we can't determine from the first-person point of view that it is the same property. It is precisely supposed to be of the nature of a mental property of a mental object, like the redness of a sense datum, that its identity is transparent to the mind. Indeed, its identity is arguably constituted by how it appears.

There are at least two good reasons for avoiding this route, however. First of all, one wants to avoid the metaphysical extravagance of positing sense data if at all possible. But suppose one didn't share this aversion to dualist metaphysics, or suppose one sought to cash out the metaphysics of sense data in materialist terms somehow, it would still go against the spirit of intentionalism about qualia to identify mental properties/objects with what's represented in perception. For most adherents of intentionalism, the principal motivation behind it is the hope of showing that all mentality is a matter of intentionality, of vindicating Brentano's dictum that intentionality is the mark of the mental. If we now introduce a new unreduced mental property/object to serve as the intentional content (unreduced, that is, to intentional content itself), then we

have merely relocated the non-intentional mental from the realm of qualia to what it is that qualia represent.

Whether or not one has a stake in reducing the mental to the purely intentional - whether or not, therefore, the second reason for avoiding sense data just mentioned is a compelling one, the first reason is quite sufficient. The metaphysics of sense data is to be avoided if possible. Another option then is to adopt an error theory. Perception presents the world as possessing properties it doesn't have. Which properties? Well, properties like those sense data would have if they existed. On this view we get much of the benefit of positing sense data, but without the metaphysical baggage and, for those moved by this consideration, without undermining the identification of the mental with the intentional.^{xi}

I see two problems with the error theory. First, it just seems implausible to attribute systematic error to our perceptual systems. There is no malfunction involved, so what basis is there for attributing contents to our perceptual states that are necessarily non-veridical? But of course if there are strong enough pressures from other quarters this consideration can be overcome. The second, and more serious, reason for opposing the error theory is this. We were pushed to consider adopting any version of intentionalism by noting how differences in qualitative character seemed to entail differences in how things appear, and therefore differences in intentional content. But if we now adopt an error theory, so that these intentional contents are necessarily empty of reference, then we might well wonder what happened to the qualia? Something, after all, really does appear to be reddish. If it's the distal object, fine. If it's

a sense datum, well okay. If it's neither, but we resist the claim that qualitative character supervenes on intentional content, then we can say that the qualitative reddishness is an intrinsic property of the experience, while its intentional content is the objective redness of the distal object. But if we reject this too, then there is literally nothing that is reddish, and this seems very hard to swallow.^{xii}

So what it seems we need is a theory that meets the following desiderata. The properties we are acquainted with in experience are really properties of something; they are not cognitive illusions. These properties of experience are representational; they involve presenting the world to us as being some particular way. The representational contents are not, however, either straightforward physical properties of distal objects, or special mental properties of peculiarly mental objects, like sense data. Is there a theory that meets these constraints?

I can think of only one, and, if I understand him correctly, it's the one defended by Shoemaker (1996). The phenomenal characters of perceptual states do supervene on their intentional contents, but what they represent are, in the first instance, dispositional properties of distal objects. Which disposition? The disposition to cause certain qualitative experiences in us. So, for instance, when I see a red fire engine, the reddishness is presented as a property of the fire engine, but the property it is, is its disposition to cause me to have a certain, well, reddish experience. There appears to be a whiff of circularity here, but in fact it's quite kosher.

What's nice about this view is the way it handles the fish-head case. In that case, remember, the very same physical property, say a spectral reflectance, caused a

reddish sensation in one eye and a greenish one in the other. Since the two eyes evolved that way, there seemed no reason to attribute error to either one. So if the intentional contents of color experiences are the spectral reflectances of distal objects, we either have to say that qualia do not supervene on intentional contents or that the qualia would be the same for both eyes. But since the subjects distinguished between the experiences caused by what they saw through the two eyes, it seemed we had to say that qualia did not supervene on intentional contents.

On the view on offer now, though, we can have our cake and eat it too. The disposition to cause an experience of one type through the right eye is a different property from the disposition to cause an experience of another type through the left eye. These are distinct properties, but not incompatible ones, so the very same objects, with the very same surface reflectances, can have both properties simultaneously. Now above I said that these dispositions are what the perceptual states represent in the “first instance”, but we can add that by representing these dispositions they also represent the categorical grounds of the dispositions in the objects; that is, the surface reflectances themselves. In this extended sense of what the two experiences resulting from the two eyes of the fish-head represent, they represent the same thing. So in one sense the greenish and reddish experiences have the same content - they track the same surface spectral reflectance - and in another they have different contents, since they attribute different dispositions to the two surfaces.

It’s important to emphasize that on the view just outlined, while we can maintain that qualitative character supervenes on intentional content, we do not get the sort of

reduction of the qualitative to the intentional that motivates most intentionalists. We get supervenience because it turns out that there's no difference in qualia without a difference in intentional content. But we do not get the reduction, because the intentional contents themselves involve attributing dispositions that are defined in terms of qualia. Intentionalism of this sort does not bring us any closer to a solution to the mind-body problem.

4. In this paper I have explored the relation between the qualitative character of experience and its intentional content. Most philosophers who defend a version of intentionalism - the doctrine that qualitative character supervenes on intentional content - do so as a way of reducing the problem of qualia to the problem of intentionality. For that reason, they tend to adopt an externalist version of the doctrine, on which the intentional contents on which qualia supervene attribute physical properties to distal objects. I have argued that this version of intentionalism is untenable.

On the other hand, I have admitted that the idea that qualia supervene on intentional contents - when divorced from its externalist interpretation - does have virtue. After surveying various other versions of the doctrine, I argued that the most plausible one is the one defended by Shoemaker (1996), on which the relevant intentional contents attribute dispositional properties to distal objects, but where the dispositions in question essentially involve qualia.^{xiii} This view satisfies various plausibility constraints on a theory of the relation between qualia and intentional contents, but it does not serve the reductive function for which most philosophers of mind hoped to employ intentionalism. So much the worse for reduction.

Joseph Levine

Ohio State University

NOTES

ⁱ Defenders of wide externalist intentionalism include Dretske (1995), Lycan (1996), and Tye (1995) and (1998). Rey (1996) and (1998) defends narrow intentionalism. I'm not sure if anyone defends wide internalist intentionalism, but Byrne (forthcoming) presents it as an option, as will be discussed below.

ⁱⁱ Does anyone actually hold this version of the view? That's a good question. The philosopher who's presentation of representationalism comes closest to expressing it this way is Dretske; he really does seem to want qualia to be properties of objects, not of our mental states. On the other hand, in discussion he has admitted that it's hard to reconcile this version of the view with phenomena like hallucinations. He doesn't want to say that there's no qualitative character if it turns out there's no object. At any rate, the difference between the two versions will not matter for what follows.

ⁱⁱⁱ Siewert goes much further and claims that for any phenomenally conscious state it is impossible for another state to share its phenomenal character but differ with respect to what it represents. I discuss his view in Levine (2001b).

^{iv} Dretske (1995) explicitly makes the argument that a consistent externalist about propositional attitude content should endorse externalism about qualitative character.

^v Again, Dretske (1995) makes just this argument.

^{vi} See Gertler (forthcoming) for an interesting discussion of the various ways of interpreting the Burgean strategy.

^{vii} Block considers this reply in Block (1996), and accuses the respondent of begging the question by assuming representationalism. Tye, however, argues persuasively that it is Block who is begging the question.

^{viii} It's pretty clear that Dretske is willing to allow that I can't know, "from the inside", that I'm really having experiences. This point came out fairly clearly in a recent presentation by Dretske at Notre Dame University (February, 2001), though there he was talking about thought, not experience.

^{ix} Tye himself makes the point that his form of reply to the Inverted Earth case doesn't work for simultaneous presentations, though he doesn't envision this sort of counterexample. In replying to Block's accusation of question-begging, he says the following (Tye 1998, 471):

Of course the above response commits the strong representationalist to supposing that there can be large changes in the phenomenal character of experiences that are inaccessible from the first-person perspective....In my view, the core intuition here is only that within a single context, a single external setting, no unnoticeable changes in the phenomenal character can occur.

Of course he's talking here about changes in a single context, which is not precisely what's going on in my "fish-head" case; but the same reasoning applies.

^x In fact he explicitly employs a fish-head type case to make his point.

^{xi} Byrne (forthcoming) briefly mentions a position like this, but then moves on. Though Rey defends narrow intentionalism, he also combines it with eliminativism, and the resulting, fairly complicated position could be seen as similar to the idea presented here.

^{xii} Though of course eliminativists like Rey (1996) and Dennett (1991) would endorse this consequence. In Levine (2001a) I address eliminativism in some detail.

^{xiii} Needless to say there are subtleties involving Shoemaker's presentation of his view that I glossed over here, so it may not be strictly correct to call what I've presented his position. Also, when it comes to the question of reduction, of course Shoemaker himself does believe that qualia are reducible, but to functional roles and their realizers, not intentional contents. For more on this, see Shoemaker (1984) and Levine (1988).