

A Puzzle About Perception

Andy Egan and James John

1. The Problem

The following theses form an inconsistent triad.

REPRESENTATIONISM: The phenomenal properties of a perceptual experience are identical to (some of) the experience's representational properties.

PHENOMENAL INTERNALISM: The phenomenal properties of a perceptual experience supervene on the intrinsic properties of the experience's subject.

STRONG EXTERNALISM: None of the representational properties of a perceptual experience is fixed by the intrinsic properties of the experience's subject.

The fact that these three theses are jointly inconsistent is one of the emerging problems in the recent literature on the philosophy of perception and consciousness. It's a *problem* because the theses are all quite attractive. Our aim here is to make the problem explicit and survey the options for resolving it.

First, we need to explain the claims in more detail.

2. REPRESENTATIONISM

The *representational content* of a perceptual experience is the way the experience represents its subject's environment as being. We will take these contents to be propositions that either are or determine possible worlds truth conditions. If the subject's

environment is the way the content represents it as being, the content is true and the experience is veridical. If the environment isn't the way the content represents it as being, the content is false and the experience is non-veridical.

Whenever a subject enjoys an experience with representational content, we can say that the subject's experience has a *representational property*. For example, if a subject's experience has the content that there is a lemon on a tabletop, then her experience has the representational property of representing a lemon on a tabletop.

The phenomenal character of a perceptual experience is the subjective or qualitative feel of the experience, "what it's like" to have the experience. Whenever a subject enjoys an experience with phenomenal character, we can say that the subject's experience has a *phenomenal property*. For example, non-spectrum-inverted subjects' experiences of lemons on tabletops (in good lighting conditions, etc.) have a distinctive phenomenal property, the one Hilbert and Kalderon mark by calling the experience 'yellow-feeling'. [Does this go back to Byrne & Hilbert?]

REPRESENTATIONISM is a claim about the phenomenal character of perceptual experience. It says that the phenomenal properties of a perceptual experience are identical with some subset of the experience's representational properties.

The main reason for believing REPRESENTATIONISM is that it promises to make the world safe for physicalism and/or functionalism by furnishing a representational theory of phenomenal character. There is a rough consensus that we are at least well on the way to having a purely physico-functionalist theory of mental representation. Phenomenal character, though, has always managed to elude

philosophers who have sought to account for it in such terms. If REPRESENTATIONISM can be made out, we will have shown that phenomenal character can be accounted for in physico-functionalist terms after all. In short, we will have reduced the problem of “what it’s like”, often always thought to be intractable, to the problem of “aboutness”, which most philosophers believe we are on our way to solving.

3. *PHENOMENAL INTERNALISM*

PHENOMENAL INTERNALISM is the claim that phenomenal properties are narrow; that is, that they supervene exclusively on “skin-in” properties. More precisely: for any two subjects S1 and S2 having perceptual experiences, if S1 and S2 are intrinsic duplicates, their experiences have the same phenomenal character. Why believe this?

First, most of us have relatively robust swampman and brain-in-a-vat intuitions that support PHENOMENAL INTERNALISM. The feeling is strong that, should a molecule-for-molecule duplicate of me miraculously form in a swamp, the duplicate would enjoy (perceptual?) experiences with phenomenal characters that match mine. Likewise for a brain-in-vat that is a perfect intrinsic duplicate of my brain.

Second, we get some evidence about where we should look for the supervenience base for phenomenal properties from facts about where we should focus our energies if we want to *change* people’s phenomenal properties. Surely we must look to the brain or some such bodily system to make changes in phenomenal character. This is the thought that grounds our swampman and brain-in-a-vat intuitions. How could the environment effect a phenomenal change in a subject without effecting a brain/body change? To think

that it could would be to think that the environment can alter the phenomenal character of experience by causally bypassing the physical mechanisms that make experience possible. It seems extremely plausible that, if you want to change the phenomenal character of someone's experience, you have to mess with their brain. *Just* messing with their environment (in a way that doesn't induce any changes inside their head) won't do the trick. In general, if the B-properties are part of the supervenience base for the A-properties, which A-properties a thing has will be sensitive to changes in its B-properties. (If, for example, paint-distribution properties form part of the supervenience base for beauty, we should be able, at least sometimes, to change whether a painting is beautiful by changing how its paint is distributed.) To the extent that a things A-properties are insensitive to changes in its B-properties, we should be suspicious of claims that the B-properties form part of the supervenience base for the A-properties. Since a thing's phenomenal properties are insensitive to changes in its extrinsic properties, we should be suspicious of the claim that extrinsic properties form a part of the supervenience base for phenomenal properties. (We don't claim that this insensitivity *entails* that phenomenal properties are narrow. More on this later.)

4. *STRONG EXTERNALISM*

The claim we're concerned with here is that *no* content is narrow—that is, that none of the content of a subject's experiences are fixed (entirely) by the intrinsic, or 'skin-in' properties of the subject. More precisely: for any subject S having an experience with a given content, it is possible for there to be an intrinsic duplicate of S

whose experience does not have that content.

There is an extensive literature motivating externalism about content. It's worth noting, though, the difference between the weaker claim that *some* content is broad from the stronger claim discussed here, that *all* content is broad. Arguments that focus on particular kinds of content, like those involving natural kinds, won't be enough to establish the strong claim.

Swampman cases might do the trick, though. It's quite plausible (at least once we've started buying into a broadly causal theory of content) that Swampman's experiences don't have *any content at all*, since Swampman isn't hooked up to his environment in the right way. If this is right, then no content can be narrow, since Swampman's experiences don't have any content in common with his molecule-for-molecule duplicate with a normal history and origin.

Another motivation (perhaps the one that's behind the swampman intuitions) is that representation is to be explained in broadly causal terms—the representational properties of states depend on their causal connections to various things in, and features of, the subject's environment. Take away all the causal connections, and you take away all the content. There is no aspect of the content that can't be changed or removed by changing the way that the subject is embedded in their environment.

5. *Why is the triad inconsistent?*

Suppose that we accept REPRESENTATIONISM and STRONG EXTERNALISM. Then we'll be committed to the possibility that two people could be in

the same brain state, but their experiences have different contents, and therefore different phenomenal properties. So we'll be forced to deny PHENOMENAL INTERNALISM.

Suppose that we accept REPRESENTATIONISM and PHENOMENAL INTERNALISM. Then, we'll need to say that (necessarily) the experiences of people who are intrinsic duplicates, since they have the same phenomenal properties, have at least some of their representational content in common. But then this shared content will be narrow, and we'll have to deny STRONG EXTERNALISM.

Suppose that we accept PHENOMENAL INTERNALISM and STRONG EXTERNALISM. Then we'll be committed to the possibility of two people, molecule-for-molecule duplicates, whose experiences have exactly the same phenomenal properties (because they're intrinsic duplicates), but *no* representational properties in common (because they're embedded in their environments in radically different ways). This is, obviously, incompatible with REPRESENTATIONISM. If phenomenal properties just *are* content properties, then if two people's experiences have exactly the same phenomenal properties, they must also have at least some representational properties in common.

6. So which claim should be rejected?

Should We Give Up PHENOMENAL INTERNALISM?

A philosopher who rejected PHENOMENAL INTERNALISM would hold that phenomenal properties are wide; that is, that they fail to supervene on exclusively skin-in properties. We think that this move is highly problematic.

For one thing, neuroscience seems to offer good grounds for retaining PHENOMENAL INTERNALISM. Indistinguishable veridical and hallucinatory perceptual experiences seem to have the same phenomenal character. For most such pairs of experiences, neuroscience identifies the same proximal physiological cause for both. The inference to the best explanation, then, is that phenomenal character supervenes on skin-in properties.

Of course, some philosophers might suggest that the phenomenology of perceptual experience tells against PHENOMENAL INTERNALISM. We doubt this. But even if there is an apparent tension between the phenomenology of experience and PHENOMENAL INTERNALISM, it is important to keep in mind that we go out on a huge empirical limb if we take the appearance of such a tension at face value and reject the claim. Appearances can be deceiving in all sorts of ways.

Should we give up STRONG EXTERNALISM?

We noted above that STRONG EXTERNALISM is stronger than the conclusion of Putnam's and Burge's famous arguments for externalism. This makes the following solution to our problem look attractive: Deny STRONG EXTERNALISM by saying that while of course the standard arguments show that *some* representational properties are broad, not *all* of them are, and If we give up STRONG EXTERNALISM, we can say that the phenomenal properties of an experience are identical to the *narrow* representational properties of the experience. Then we can , and keep both REPRESENTATIONISM and PHENOMENAL INTERNALISM. There are two related

difficulties with this move. a couple of problems with this move.

First, the picture of representational content as bound up with information-carrying, and information-carrying as being bound up with causal history, that motivates STRONG EXTERNALISM, is really very attractive. One way to give up STRONG EXTERNALISM, would be to have to give up this picture for at least *some* of the contents of experience. But it's not clear what other story is available. States of subjects don't come by their contents by magic, and it doesn't seem like any description of a subject in purely *intrinsic* terms can entail that they're in a state with any particular content. It really does seem as if, by varying the subject's environment appropriately, we can change, or eliminate, *any* aspect of the content of the subject's experience. At least, it seems this way if we're only thinking about how the experience represents the external world, and not how it represents the subject herself. In the case of representational properties which have got to do with how the subject of the experience is represented, narrowness seems much more plausible, at least partly because it's compatible with a causal/informational theory of representation. So the sorts of representational properties that look like good candidates for narrowness are properties that specify how the experience represents the subject of the experience as being. If this is right, there's a rather sharp limit on what sorts of representational properties will be good candidates to be the phenomenal properties. This leads to our second difficulty:

Second, even if we deny this, and say that there are some representational features of experience that would remain constant under any change in the subject's connections to their environment, we still need a non-mysterious story about which features these are,

and just *how* they can be fixed entirely by the skin-in properties of subjects. The need for such a non-mysterious story is particularly urgent if the reason for giving up STRONG EXTERNALISM was to keep REPRESENTATIONISM in order to pave the way for a materialistically acceptable account of phenomenal character. It's not at all clear, though, how such a story would go. Certainly there isn't any such story on the table right now.

In order for the denial of STRONG EXTERNALISM to dissolve the puzzle, there also needs to be *enough* narrow content, and of the right kind, to make it plausible that the phenomenal properties of experience supervene on representational properties having to do with *that* kind of content. One thing that we'll have to deny, since it's pretty clear that *color* content isn't narrow, is that the distinctive phenomenology of color experience is determined by *color* content. The difficulty here is that it's not enough just to have *some* narrow content – the narrow representational properties that we have available need to track the phenomenal properties of experiences in the right way.

These are obstacles to solving the problem by denying STRONG EXTERNALISM. We don't claim that they're insuperable. One of us is quite drawn to this strategy, though we are both concerned about its long-term prospects. While denying STRONG EXTERNALISM looks better to us than denying PHENOMENAL INTERNALISM, it is certainly not smooth sailing.

Should We Give Up REPRESENTATIONISM?

A philosopher who rejected REPRESENTATIONISM would most likely hold

that, while experiences have representational content, their phenomenal properties are not identical to any representational properties. Instead, phenomenal properties are (or are determined by) non-representational qualia, intrinsic properties of experience that are neither identical with nor determined by representational properties.

The problem with this move is that it means giving up the theoretical benefits of reducing the problem of what it's like to the problem of aboutness. Indeed, there is a prima facie worry about whether a position that appeals to qualia is compatible with physicalism.

But perhaps it doesn't matter. We assumed that to reject REPRESENTATIONISM would be to take up a view on which there is no substantial connection between an experience's phenomenal properties and its representational properties. Such a view would burden qualia with all of the phenomenal labor. But there is another way to reject REPRESENTATIONISM, one which does posit a substantial connection between phenomenal and representational properties: INTENTIONALISM.

INTENTIONALISM says that phenomenal properties supervene on representational properties without being identical with any of them. More precisely, it says that, for all possible subjects and across all sense modalities, if two perceptual experiences have the same representational content, then they have the same phenomenal character. INTENTIONALISM still holds out the promise of a representational theory of phenomenal character. Most importantly, though, it seems to avoid the inconsistent triad since it is a supervenience claim, not an identity claim. If phenomenal properties supervene on, but are not identical to, representational properties, then it could easily be

that several *different* representational properties could underlie the same phenomenal property. (Which would avoid the problem about having to say that the experiences of intrinsic duplicates in sufficiently different environments would, since they would have different representational properties, also have different phenomenal properties.) Can the philosopher attracted to REPRESENTATIONISM but concerned to avoid inconsistency hang his hopes on INTENTIONALISM?

We are skeptical. The problem with INTENTIONALISM is that it looks as if it doesn't escape the inconsistent triad after all. Consider the following case.

Ernie is a normal, non-spectrum-inverted perceiver. Vert is his spectrum-inverted twin. As it happens, due to their different causal histories, Ernie and Vert are such that when Ernie is in overall brain state B his visual experience represents greenness and when Vert is in B his visual experience represents redness. This is precisely the sort of situation allowed for by STRONG EXTERNALISM. Ernie, let's suppose, is in B and enjoying a green-representing experience, and Vert is in B and enjoying a red-representing experience. So far, no problem. Ernie and Vert are in the same brain state, so their experiences must have the same phenomenal properties. Even though their experiences have different *representational* properties, the Intentionalist is unperturbed, because she doesn't *identify* the phenomenal with the representational properties. So she is not (yet) forced to say that Ernie's and Vert's experiences differ phenomenally.

But now consider a third subject, Grover. Grover is a normal, non-spectrum-inverted perceiver like us. By INTENTIONALISM, both Grover's and Ernie's green-representing experiences will be "greenish"—they will have some phenomenal property

G—and both Grover’s and Vert’s red-representing experiences will have a “reddish” phenomenal property R. And we know from our own experience (since Grover is a perceiver like us) that $G \neq R$. Now the Intentionalist must give up her PHENOMENAL INTERNALISM, since now we have it that Ernie’s experience is G while Vert’s is R, *even though they are in the same brain state*. This, obviously, contradicts PHENOMENAL INTERNALISM. Therefore, there is an inconsistent triad composed of PHENOMENAL INTERNALISM, STRONG EXTERNALISM, and INTENTIONALISM.

What can be done? Note that this case works only if the version of INTENTIONALISM in question is one that holds across all possible subjects. A version of the claim restricted to experiences within the conscious life of a single subject would avoid the inconsistency.

Unfortunately, such views aren’t very satisfying as attempts to support a representational theory of phenomenal character. The problem can be brought out by reflecting on *intramodal* INTENTIONALISM. Intramodal INTENTIONALISM says that, within a sense modality, phenomenal properties supervene on representational properties. This leaves the phenomenal differences between the various sense modalities unexplained in representational terms. It is consistent with intramodal INTENTIONALISM, after all, that for each distinct sense modality there is a special intrinsic quality—a quale—that all and only experiences within that modality possess. This quale would go unexplained in representational terms.

The problem for intrasubjective INTENTIONALISM should now be clear. It is

forced to leave possible phenomenal differences between subjects unexplained in representational terms. It is consistent with intrasubjective INTENTIONALISM that for each subject there is a special quale that all and only experiences had by that subject possess.

Thus, the price of going Intentionalist is that one will probably have to abandon the rationale for such a theory in the first place: the hope of a thoroughgoing representational theory of phenomenal character.

7. Conclusion

Our aim has been to make as explicit as possible what we see as the main issue underlying much of the recent literature on consciousness and perception. We can't have all of REPRESENTATIONISM, PHENOMENAL INTERNALISM, and STRONG EXTERNALISM. In fact, PHENOMENAL INTERNALISM and STRONG EXTERNALISM are jointly incompatible with even the weaker claim of INTENTIONALISM.

There quite a bit of room to maneuver in deciding which one to give up, but all of the maneuvers have their costs. STRONG EXTERNALISM is well-motivated by the classic arguments for externalism. It's quite a strong externalist thesis, however, and maybe it will turn out that the arguments for externalism only mandate a weaker version. PHENOMENAL INTERNALISM has been questioned, but it really does seem overwhelmingly plausible. Finally, we could reject INTENTIONALISM, and allow for

phenomenal properties that do not supervene on the representational. This would mean giving up an attractive physicalist strategy, but it wouldn't require us to give up physicalism.