

What Constitutes Phenomenal Character?

Murat Aydede
University of British Columbia
Department of Philosophy
murat.aydede@ubc.ca

Abstract. Reductive strong representationalists accept the Common Kind Thesis about subjectively indistinguishable sensory hallucinations, illusions, and veridical experiences. I show that this doesn't jibe well with their declared phenomenal externalism and argue that there is no sense in which the phenomenal character of sensory experiences is constituted by the sensible properties *represented* by these experiences, as representationalists claim. First, I argue that, given general representationalist principles, no *instances* of a sensible property constitute the phenomenal character of the sensory experience that represents them. Second, I argue that, with two very plausible assumptions in place, no sensible property *qua universal* can constitute the phenomenal character of experiences either. At the end, I offer an alternative picture.

Suppose Sam is intently looking at a blue and round ball (call the ball, Tom) in front of her against a roughly uniform neutral background in good day light. Let's say that Sam is having a veridical visual experience VE1 as of something being blue and round.

Let B be the property complex, being blue and round:

$$B = \lambda x (x \text{ is blue} \ \& \ x \text{ is round})$$

B is a type, a universal, and is instantiated by Tom. Call this particular instance of B , $b1$. Sam is seeing $b1$ — the instantiation of B (by Tom).

Sam is intently and carefully looking at Tom for about 5 seconds. It is natural to say that he is aware of $b1$. This is a direct *de re* awareness, if anything is. Sam's visual experience seems to put Sam directly in contact with $b1$ (and with Tom, of course). This experience has an immediate seemingly world-disclosing presentational character, which has a certain phenomenological profile that we may call its *phenomenal character* — there is something it's like to undergo this particular experience which seems to immediately present $b1$ to Sam. What is the relation of VE1's phenomenal character to $b1$? Is this relation merely

causal, or rather is it constitutive (partly or fully¹)? If the answer is the latter, I'll say that Sam's experience is instance-involving.

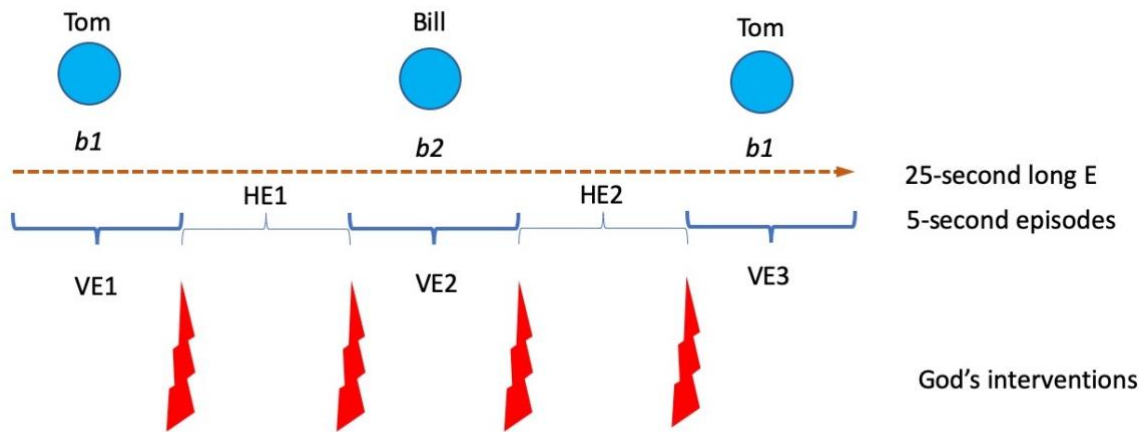
Reductive strong representationalism is meant to be a view committed to a form of phenomenal externalism, according to which the phenomenal character of sensory experiences is constituted by the character of (non-conceptually or sensorially) represented sensible properties. On this view, physical duplicates being in the same state may differ in the phenomenal character of their respective experiences — if these somehow sensorially represent different sensible properties. According to representationalists,² the phenomenal character of a sensory experience doesn't supervene on the narrow physical constitution of the experiencing subject. Thus, the represented sensible properties are constitutive of phenomenal character. So, one would expect that a representationalist of this sort would answer the above question by saying that *b1* constitutes the phenomenal character of Sam's veridical visual experience — indeed they often say that the phenomenal character is *identical* to the represented content or feature. First, I will argue that this externalist claim about property *instances* cannot be true given what representationalists have to say about hallucinations. Second, I will show that sensible properties *qua universal* cannot constitute the phenomenal character either. My overall conclusion will be that phenomenal externalism is false, and that if representationalism entails such an externalism, it too is false. In conclusion, I'll offer an alternative picture that is (weakly) representationalist but internalist and naturalist.

1 The role of property instances

Suppose, at the end of the five seconds, God intervenes and takes over the causal route stimulating Sam's brain in such a way that Sam doesn't notice anything when God removes Tom. It's a smooth transition. Sam is now having a subjectively indistinguishable hallucinatory experience, HE1, as of a blue and round ball in front of her. We can extend the thought experiment. Another five seconds pass and God puts a qualitatively identical but numerically distinct ball (call it, Bill) back in where Tom had been when Sam was looking at it and lets Bill take over the causal operation on Sam: Sam is now having a VE2 with another instantiation of *B*, *b2*. Another five seconds pass and God intervenes again in the same way, smoothly removes Bill while maintaining the neural activity in Sam associated with *B*. Sam is now having another hallucinatory experience, HE2. Finally, we can suppose that after another 5 seconds, God puts Tom back where it was twenty seconds ago and lets the causal stimulation be controlled by Tom again. Sam is now having another veridical experience, VE3, which makes her aware of *b1*. Sam has no clue about what is going on. (The following diagram may help.)

¹ Metaphysical constitution may be partial or whole. I'll omit this qualification from now on except when it matters — see below.

² From now on, when I talk about "representationalists" without qualification, I'll have in mind *reductive strong representationalists* in mind such as Dretske (1995) and Tye (2013, 2014), among others.



By stipulation, the phenomenal characters of these smoothly connected experiential episodes (VE1, HE1, VE2, HE2, VE3) are subjectively indistinguishable. Indeed, throughout 25 seconds, Sam falsely but justifiably believed that she was looking at a blue and round ball that remained identical. The Common Kind theorists in philosophy of perception think that the subjective indistinguishability in such cases is to be explained by the presence of positive phenomenology: VE1, HE1, VE2, HE2, VE3 all have the same phenomenal character. These five episodes share a common fundamental phenomenological or experiential kind. Representationalists accept the Common Kind Thesis.

According to representationalists, the phenomenal character of sensory experiences is exhaustively a matter of what sensible properties are represented in the experience, whether or not the experience is veridical. In our example, all the five sensory episodes represent *B* as instantiated, and it is this fact that determines the identity of the phenomenal character of Sam's experience during the 25 seconds she was intently looking at the "ball." This entire experience — call it *E* — is an experience that remains phenomenally identical throughout 25 seconds, where *b1*, *b2* are the instances causally related to *E* during the first, third and fifth 5-second periods (therefore making VE1, VE2, VE3 accurate), while *E* has no actual objects or instances during the second and fourth periods, which makes HE1 and HE2 hallucinatory. Thus, if *E* has the same phenomenal character throughout, this phenomenal character cannot constitutively involve *b1* and *b2*. So, if the veridicality of VE1 is what partly makes for Sam's awareness of *b1*, the phenomenal character of this awareness (VE1) cannot be constituted by *b1* — similarly with VE2 and VE3. VE1 is not in this way instance-involving. The relation of VE1 to *b1* is *only causal*. So, the phenomenal identity of *E* is not instance-involving at all. In fact, given the way the thought experiment is set up with "external" physical objects and mind-independent sensible properties, we can generalize: for Common Kind theorists who believe that subjectively indistinguishable

veridical, illusory and hallucinatory experiences share a common positive phenomenal core, even in cases where the experiences are veridical, the sensible property instances the subjects are aware of never constitute the phenomenal character of these experiences.

According to representationalists, what metaphysically fixes the identity of the phenomenology of *E* is this: *E* represents *B* (as instantiated)³. *E* is veridical when it is sustained by an appropriate causal/informational link to phenomenologically irrelevant instances of *B*, and non-veridical otherwise.

Generalizing, the situation is the same with *all* sensory experiences: their phenomenology is never constituted by the instances of the sensory properties or property-complexes they veridically represent (when they do). Veridical sensory experiences, according to representationalists, are not only not object-involving, but are also not instance-involving.

This result may come as a surprise to some.⁴ For phenomenal externalism seems to demand that veridical sensory experiences are instance-involving. But recall the ease with which the representationalists claim that Sam is sensorially aware, *de re*, of a (locally uninstantiated) universal, *B*, while having HE1 and HE2.⁵ This sounds mysterious and puzzling, but I'll assume that all they mean with this is that in hallucinatory experiences like HE1 and HE2 there is, in an obvious sense, still sensory representation: Sam's hallucinatory experiences still represent a (locally) uninstantiated sensible property complex, namely *B*, a universal — it just *misrepresents it as instantiated*.⁶ It is this fact, according to representationalists, that determines the phenomenal character of the hallucinations — no property instances are ever involved. But given the Common Kind Thesis, this phenomenal character is also the *very same character* of the veridical episodes. Thus, even the veridical episodes don't involve instances of sensible properties as the constitutive determinants of the experiences' phenomenal character.

A slightly different way to express the main point is this: whatever the phenomenal character of any sensory experience involves, it involves it *essentially*. But, given what representationalists say about hallucinatory experiences, property-instances are not *essentially* involved in the constitution of phenomenal character of any experiences. Hence, no sensory experiences (veridical or not) are ever instance-involving.

³ Sensory experiences as of a sensible property *F* are always awareness of *F*-instances when veridical. The non-existence of *F*-instances is what makes *F*-experiences illusory or hallucinatory, i.e., non-veridical. So, *pace* Tye (2014), Sainsbury (2019) and Gottlieb & Rezaei (2021), sensory representation of sensible properties (when deployed in perception rather than, say, imagination, mode) has always assertoric force, thus accuracy conditions.

⁴ For instance, Pitt (2017) and Gow (2018) seem to write with the assumption that for representationalists property instances are phenomenology-constituting in veridical experiences.

⁵ See, for instance, Tye (2015: 485, 2013: 51–52), Dretske (1999: 107).

⁶ In my view there cannot be a sensory awareness *of* a universal uninstantiated. The function of sensory awareness is to detect property instances in the physical (including, bodily) environment — see next section. There is no such thing, properly speaking, as detection of properties *qua* universals uninstantiated. I simply take this talk of sensory awareness of universals as expressing (PC) — see below.

2 Phenomenal character as the property of ‘representing P’

But then we seem to have a puzzle. If the phenomenal character of an experience of a sensible property *P* is never constituted by the instances of *P*, in what sense is it constituted by the property *P* (*qua* universal)? Indeed, what does it mean to say that the phenomenal character is constituted by *P* but not by its instances? For surely, as pointed out at the start, phenomenal externalism requires that sensible properties themselves are constitutive of the phenomenal character of sensory experiences that represent them. Representationalists keep telling us that it is the represented properties that constitute the phenomenal character. If it is not their instances, what is it for the sensible properties *qua* universals to constitute sensory phenomenology? It seems to me that the only plausible thing to say at this juncture is this:

(PC) The phenomenal character of a token sensory experience, *s*, as of a sensible property *P* at time *t* is constituted by the fact that *s* possesses the intentional property of sensorially *representing P* (as instantiated) at *t* — whether or not *s* is veridical.

One might think that the sense in which such a view is phenomenal externalist is that *P* is a property that can be instantiated only by external physical (mind-independent) objects. But there is more to this claim as we will see in a moment. So, no instances of *P* are ever constitutive of sensory experiences as of *P*. All that is needed for a sensory state, *s*, to have the relevant phenomenal character at a time is that *s* be representing *P* (as instantiated) at that time — that is, *s* have the property at *t* of *representing P*. Veridicality, but not the phenomenal character, of *s* comes with the causation of *s* by an appropriately related *instance* of *P*. We may even say that *s* is what makes the subject sensorially *aware of* the relevant instance of *P* when *s* is veridical, i.e., appropriately caused. But the phenomenal character of *s* is not constituted by the relevant *P*-instance. Representationalists may point out that in veridical cases the subject is aware of the properties themselves *as well as their instances*. But again, the point is that the *phenomenal character* of such sensory awarenesses is solely due to the relevant states’ *representing* the sensible properties *qua* universals — not due to the particular instances that are merely causally involved in the awareness. This much seems clear given what representationalists say about hallucination and their acceptance of the Common Core thesis.

Note that, if I’m right so far, the phenomenal character of a sensory experience as of *P*, is a property of the experience: it is the property of sensorially *representing P*. Having *this* property is what metaphysically constitutes the phenomenal character of a sensory experience as of *P*. In whatever sense we have introspective access to the phenomenal character of such an experience, it is to this property (*representing P*) that we have access, not just to the property, *P*.

Representationalists sometimes also say things like this:

(PC!) The phenomenal character of a token sensory experience, *s*, as of a sensible property *P* at *t* is constituted wholly by *P* (the represented universal) — or, *is* just *P*!

It is completely obscure how to make sense of such claims. Suppose the sensible property in question is an instance (*b*₁₆) of a particular shade of blue (*B*₁₆) that Sam is aware of during VE1. According to representationalists, this property, *B*₁₆, is a physical surface property, say, a certain set of surface spectral reflectances, *SSR*_{*B*₁₆}. Whether or not Sam is hallucinating during *E*, representationalists claim that the phenomenal character of Sam's (colour) experience remains identical. But if *SSR*_{*B*₁₆} *is* the phenomenal character of *E*, it would of course be completely unsurprising that this physical property (*qua* universal) has been self-identical and remaining identical — whatever that means. But of course! Nobody would take this claim to be making a philosophically controversial or even interesting point. Therefore, when representationalists make claims of this sort (PC!), we will interpret them as meaning (PC).

Our next task, then, is to understand what sorts of facts constitute a state's representing *P* (even when *P* is not locally instantiated)? But it would be useful to summarize our discussion so far and draw some lessons before we do that. The most important point to keep in mind is that the phenomenal character of even normal veridical experiences of sensible properties *P* is not constituted by the instances of *P*. The instances are causally/informationally implicated in generating these experiences, and therefore, in this *causal* sense, they determine what experiences with what phenomenal character to be tokened. But the phenomenal character itself is metaphysically constituted by a property that doesn't involve the instances of *P*. The phenomenal character of an experience as of *P* is metaphysically constituted by the experience's having the property of *representing P*, according to representationalists. The instantiation of this intentional property (representing *P*) by a sensory state doesn't metaphysically require the simultaneous existence of any *P*-instances anywhere. Given that Sam is a *bona fide* member of human species, her internal physical/functional constitution, for instance, is metaphysically sufficient to instantiate sensory states representing *B*, whether or not there are any instances of it around. To this extent, then, Sam's internal constitution is metaphysically sufficient for her to have experiences with the phenomenal character that is here identified with *representing B*. Without any further externalist account of what the possession of this intentional property comes to, we don't yet have a phenomenal externalist position.

3 Naturalistic psychosemantics for 'representing *P*'

The project now is to understand how externalism may arise out of the representationalist account of intentional facts. For reductive or naturalist representationalists, the intentional facts (i.e., the possession of the property of *representing P* by sensory states) concern some

combination of facts about causal co-variation, indication, teleological function, tracking, etc. For our purposes, a simplified Dretskean version will suffice (Dretske 1995: 14ff):⁷

- (a) The sensory state token *s* has the phenomenal character it has in virtue of the fact that it *systemically represents P*.
- (b) *s* systemically represents *P* in virtue of the fact that it's a token of a state type *S* whose function is to indicate (track, carry information about) instances of *P*.

Thus:

- (c) *s* has the phenomenal character it has in virtue of the fact that it's a token of a state type *S* whose function is to indicate (track, carry information about) instances of *P*.⁸

The indication function here is entirely causal/nomological with a certain historical selection condition (depending on how one understands the notion of function involved). Given that the indication function isn't sufficient to make *P*-instances *constitutive* of phenomenal character of *S* tokens *now*, we can ask: is there any reason to think that tokens of *S* have had their phenomenal character constituted by *P*-instances in the evolutionary history during which the state type *S* was selected because its tokens regularly indicated instances of *P*? The answer clearly is *No*. Information transmission works, roughly, by there being a lawful causal correlation between instantiations of two properties, *P* and *S*: when the channel conditions are right, *P*-instances causally determine *S*-tokens. That, then, constitutes *S*-token's indicating a *P*-instance. If any *S*-token throughout the selection process, happened to have phenomenal character, this character wasn't constituted by the *P*-instance that caused it in the circumstance. Recall, as per (c), phenomenal character is constituted only by a state-token's belonging to state type whose function is to indicate *P*-instances. This type wasn't there, to begin with, in the selection stage.

According to representationalists, then, external objects and sensible property instances never metaphysically constitute the phenomenal character of sensory experiences.⁹ Rather the phenomenal character is constituted by what sorts of state types get to be causally tokened. What is constitutive for the token experiences to have the phenomenal character

⁷ Tye gives the following formula: "a sensory state is about a property, *P*, just in case the state is of a type that is Normally tokened if and only if *P* is tokened and because *P* is tokened." (2014, fn.20)

⁸ More is required here for the emergence of phenomenal character such as the fact that the state type needs to have certain format constraints (e.g., non-conceptual, imagistic, analog, etc.). In particular, the sensory representation types, the *S*'s, need to belong to a sensory *system*, **S**, whose state types are systematically inter-defined according to a multi-dimensional discriminability space. And this whole system needs to be coupled to a certain kind of cognitive architecture with conceptual and conative states that extract information for further processing and behavior. I will ignore these sorts of complication and assume that whatever else is needed is in place. I'll sometimes call these 'background' conditions for the emergence of phenomenal character.

⁹ Compare the robust phenomenal externalism of disjunctivist naïve realists. The phenomenal character of sensory experiences, in the good cases, is metaphysically constituted by physical objects and the mind-independent properties instantiated by these objects (plus perspectives, etc., perhaps). See, among others, Campbell (2002), Martin (2004), Brewer (2011).

they do is that they belong to a *state type* whose tokens are under the nomic control of property instances that they track under Normal conditions. In other words, the phenomenal character of token experiences is inherited from the *type* they belong to, not from the property instances these tokens purport to indicate. This state type is a functional type whose tokens purport to indicate and are the realizers of the experiences that represent sensible properties — thus constituting their phenomenal character.

Let me clarify a point about the *causal* determination of phenomenal character. There is of course a clear sense in which the phenomenal character of Sam's experience VE1 was determined by *what* he saw, namely an instance of blue. I argued that this determination was *causal* rather than constitutive. Causal determination of this sort is more like the causal selection of a sensory state from among a system (\mathcal{S}) of states already possessing different phenomenal characters — as per (c) above. For instance, if the ball Sam saw were red, instead of blue, VE1 would have a different phenomenal character. Not because the particular instance of red Sam saw would constitute the phenomenal character of Sam's visual experience, but rather by causally activating a token of a different sensory state type in Sam's \mathcal{S} that has the function of indicating instances of red.

This kind of causal determination is not relevant to phenomenal externalism that representationalists usually have in mind. What they need to defend is externalism of the constitutive kind. Above we've determined that the most plausible version of the claim that phenomenal character is constituted by "external" universals is given by (PC). And (c) is one way to cash out (PC) in completely naturalistic terms — in terms of Dretskean indication functions.¹⁰ Does it deliver what is needed? I will argue in what follows that it doesn't.

4 The alleged phenomenal externalism

Representationalists typically argue for their case in the following way. VE1 is a brain state, a certain activation of a set of neurons in the relevant circuitries implementing the quality spaces in color and shape detection. For ease of exposition, let's just concentrate on color and ignore the shape. Let \mathcal{S} be Sam's color visual system whose different state types, S_i , implement the relevant neural activations in her visual pathways and cortex — these activations corresponding to registering different colors. In particular, let S_{16} be the state type belonging to \mathcal{S} that has the systemic function of indicating instances of Blue₁₆ (B_{16} , the universal). We can say, then, the token state, s_{16} , is the realizer of Sam's color experience e (say, during the first 5 seconds) indicating b_{16} , the instance of B_{16} had by

¹⁰ I won't bother to try out other naturalist proposals like Tye's or Millikan's. Differences in these proposals won't make a difference in my argument in what follows. Also, although I'm very sympathetic to a Dretskean psychosemantics (see Aydede & Güzeldere 2005), I won't assume here that these sorts of proposals can naturalize perceptual intentionality.

Tom.¹¹ A representationalist can say that even if the phenomenal character of e is not instance-involving (hence not constituted by b_{16}), it does involve B_{16} — it systemically represents B_{16} in virtue of having the function to indicate instances of B_{16} . So, B_{16} is what partially but essentially individuates S_{16} of which s_{16} is a token.

S_{16} , the realizer of the experience type of which e is a token, is a state type whose historically relevant tokens got selected because they have indicated B_{16} -instances. Although these indication relations have all consisted of particular causal interactions between the tokens of B_{16} and S_{16} , the result was that S_{16} acquired the function of indicating B_{16} -instances, thus the power of representing B_{16} (as instantiated) — veridically or not.¹² The individuation of S_{16} (in fact the whole \mathcal{S}) thus essentially adverts to the historical and causal interactions with B_{16} through its instances. This is what systemic representation comes to. e 's veridically representing b_{16}/B_{16} is therefore an essentially relational property of e . Indeed, E 's systemically representing B_{16} (thus its having the same phenomenal character during 25 seconds) is a relational property of E . A representationalist would then conclude: change the relation, thus the type-identity of the token state, as per (c) above, you change the phenomenal character of E , thus, of the experience type of which E is a token. You've got your phenomenal externalism of the constitutive kind.

Before I start arguing against this, let me say a few things about Sam's \mathcal{S} : Well, it is Sam's color perception system. So, it doesn't have any historically relevant tokens that contributed to its own selection and passing its blueprint to Sam's descendants. Rather \mathcal{S} belongs to a type of system \mathbf{S} that is phylogenetically fixed for the human species. The only known way of phylogenetic development of sensory systems is at the biological level — at the level of the mechanics of biological inheritance (involving DNA replication and expression). At this level, \mathbf{S} has a fairly robust neurophysiological description whose “system-level analysis”, as engineers call it, can be given at the neurofunctional level. So, if representing a sensible property like B_{16} is a relation, it is a relation with two relata: B_{16} and S_{16} *qua a neurofunctional state type belonging \mathbf{S}* . Therefore, the state types of Sam's \mathcal{S} acquire their relational character by being of the same neurofunctional system type as \mathbf{S} — or whatever the descriptive level required by the transmission of a phylogenetic trait may be. \mathbf{S} had had millions of more tokens after it'd acquired its function — let's idealized away all the messy variations in this phylogenetic process (we don't have any good account of when the acquisition process is considered over or why it cannot change later). In almost all these cases, the internal constitution of people with \mathbf{S} who are in S_{16} suffices for them to have a sensory experience with the attendant relevant phenomenal character —

¹¹ This is partial realization given that VE1 involves representing other properties. But ignore this for the moment. We'll concentrate on the simpler color case.

¹² It's highly unlikely that it is the individual state types, independently of others, that acquired the function of indicating specific sensible properties. Rather, it is the system type \mathbf{S} as a whole, whose states are interdependent, that acquired the function of indicating a *range* of sensible properties within a certain stimulus domain.

whether or not they are veridical. This is not argument against relational individuation yet, but it's important to keep in mind the robust neurofunctional character of the system.

5 “Shifted” phenomenal character

Scenario 1. Now consider Mary, who is a contemporary of Sam and roughly of the same age. Both are considered to have “normal” color vision. But Mary’s colour phenomenal space, although the same as Sam’s, responds to a systematically shifted color (hue) spectrum. For instance, the tokens of Mary’s S_{16} are under the nomic control of instances of B_5 — a slightly but noticeably less reddish shade of blue. So, Mary’s tokens of S_{16} regularly indicate instances of B_5 — not B_{16} .¹³ What is the phenomenal character of Mary’s S_{16} states? Are they of the same kind as those of Sam’s? For a representationalist, the answer depends on whether they are both *representing* the same color property or not. For instance, it may be that Mary is systematically misrepresenting the colors that she sees — she may be systematically misrepresenting an instance of B_5 as B_{16} (and similarly for the rest of the shifted spectrum). This could be for a variety of reasons. For instance, if, due to a genetic fault, the pigments in her cones have slightly different compositions so their response curves are slightly different, or maybe her eye lenses are slightly more yellowish from birth, etc. If this were so, even though the states of her **S** have the same indication function (thus the same representational contents) as that of Sam’s, they regularly would fail to perform their function successfully resulting in systematic misrepresentation. In other words, we may think of Mary’s **S** not fulfilling its function in the way it was selected for. Mary’s **S** may be an anomaly. In such a scenario, Sam’s and Mary’s experiences realized by S_{16} would have the same phenomenal character despite their systematically seeing different shades of colour (Mary seeing instances of B_5 and Sam seeing instances of B_{16}). Of course, it is very likely that, given how widespread the shifted spectrum cases actually are among the normally color sighted people, both Sam and Mary may be systematically misrepresenting colors all the time. No problem so far.

Scenario 2. But let’s modify the example slightly. Let’s assume that Mary’s color vision isn’t an isolated case. Rather it’s the *function* of Mary’s S_{16} states to indicate instances of B_5 and similarly it is the function of her visual system to respond to the light spectrum in this “shifted” way. So, we are assuming that Mary’s visual system came to be where it is now due to an evolutionary process that selected for it. Now there is a question about whether to count Mary’s visual system as of the **S** kind. The issue here concerns how narrowly or broadly we should individuate **S**. Visual color processing starts with photons hitting the cones and its later stages involve whatever neural circuitry (including opponent processes running through LGN and various parts visual cortex) implements the final discrimination behavior that underlies the color quality space — sometimes known as the

¹³ Not only that, but almost all her hue circle is shifted slightly compared to Sam’s stimuli giving rise to same color experiences. We don’t need to assume that the degree of shift is even or thoroughly systematic. There are plenty of actual cases like that.

three-dimensional color solid. The processing in the cones as well as the retinal and early post-retinal processing may be manipulated without massive differences resulting in the implementation mechanisms of the color quality space. There is no reason to think that among the normally sighted but shifted color spectrum cases people have different color quality spaces. I will just stipulate that **S** be individuated without including these very early processes. It is an empirically plausible assumption that most people with “shifted color qualia” share the same color quality space implemented in more or less neurofunctionally type-identical neural structures. If we individuate **S** this way, then our assumption about Mary amounts to the assumption that Mary belongs to a group of community whose **S**-state types have evolved to acquire a different indication function due to some differences in their environment and in their early (pre-LGN) neural processing.¹⁴ So, for instance, while Mary’s **S**(S_{16}) has the function to indicate instances of B_5 , Sam’s **S**(S_{16}) has the function to indicate instances of B_{16} . The result is that the two tokens of the same neurofunctionally identified state type **S**(S_{16}) sensorially represent different colors for Mary and Sam.

Now we have reached the kind of phenomenal externalism that representationalists have in mind — the constitutive kind. In this last scenario, Sam and Mary share a neurofunctionally identified **S** whose type-identical states, S_i , represent different colors in Sam and Mary. When Sam looks at the ball, being in **S**(S_{16}) she is veridically seeing an instance of B_{16} . When Mary looks at another ball that is blue₅, being also in **S**(S_{16}), she is veridically seeing an instance of that color (B_5). Representationalism delivers the result that the phenomenal character of Sam’s and Mary’s experiences is of different kinds because they represent different color properties (*qua* universals). Thus, despite their neurofunctional type-identity, the phenomenal character of Sam’s and Mary’s experiences is constituted by different color universals — different sensible properties that are instantiated only by “external” mind-independent entities. We can even think of Sam and Mary sharing all their narrow internal constitution relevant for conscious color processing. Representationalists claim that Sam’s and Mary’s experiences despite being realized by the same internal physical state — **S**(S_{16}) — have different color phenomenology because they represent different colors. This claim, we have already seen, doesn’t entail that the instances of these colors, b_5 and b_{16} , are metaphysically relevant to the constitution of the phenomenal character of their respective experiences. The role the color instances play is merely the causal generation of their respective **S**(S_{16})-tokens that nevertheless differ in their phenomenal character. Similarly, when Sam and Mary look at the same blue₁₆ ball they have experiences with the same phenomenal character despite the fact that they occupy different states of **S**.

¹⁴ In order for this work, we’ll probably need to assume that the communities Sam and Mary belong to have been relatively isolated from each other throughout the evolutionary process — or at least they haven’t mixed their lineages much. For the thought experiment to work, all that is needed is the metaphysical possibility that the states of **S** may have acquired distinct indication functions (indicating slightly different spectra) in two different phylogenetic lineages.

Note the difference in the phenomenal character of Mary's experience in the two scenarios when she is looking at the blue₅ ball. In both cases she is in the same physical/functional state — $\mathbf{S}(S_{16})$. But the phenomenal character of her experience in the first scenario (when she misrepresents a b_5 as B_{16}) is different from the phenomenal character of her experience in the second scenario when she veridically represents b_5 as B_5 . This is striking. In fact, with some minor adjustments, we can conceive of the two scenarios as involving the same physical *personal* history of Mary with identical internal constitution. Depending on how we conceive of Mary's *evolutionary* history (selection history of her ancestors), therefore, we get *different* phenomenal characters in the counterfactual scenarios despite the type-identity of Mary's internal constitution in both.

If you are one of those who think that the phenomenal character of sensory experiences is episodic, categorical, and essentially belongs to the *here-and-now*, you are likely to balk at the representationalist conclusion that Mary's experience differs in its phenomenal character from Sam's, or that Mary's experiences in the two scenarios have different phenomenal character — especially given that no property instances are ever constitutive of phenomenal character. But we can do better than just having an incredulous look. Let's examine how they will behave when we start testing their color vision under carefully setup laboratory conditions. Suppose we ask Sam to pick a color chip from a bunch of others that she sees as a unique blue (a shade of blue not at all reddish and not at all greenish). We then invite Mary to look at the same chip under the same viewing conditions. We know that when Mary visually experiences the chip, she occupies a state of \mathbf{S} different than the state Sam occupies. Representationalists would have us believe that Sam and Mary looking at this chip would represent the color of the chip accurately and therefore would have the same phenomenal character. But it would border on the absurd if we refuse to take Mary's report about her phenomenology seriously when she says that this chip looks slightly greenish blue, not unique blue. Thus, contrary to the prediction by representationalists, Sam's and Mary's experiences have different phenomenal character despite their both representing the color of the same chip correctly.¹⁵ But then phenomenal externalism of the kind representationalists have had in mind is refuted. The kind of relational individuation of \mathbf{S} 's states does not assign the correct phenomenal character to them. It looks like the neurofunctional type-identity of $\mathbf{S}(S_i)$ will trump the relational individuation any time the two schemes come apart.

The structure of the above argument is simple. Two actual "normal" perceivers, Sam and Mary, one having slightly shifted color spectrum relative to the other, are looking at the same chip that instantiates a color property, C . Their color experiences represent the color of the chip *accurately*. If the phenomenal character of their respective color experiences were constituted by the color properties (qua universals) they represent, then the

¹⁵ Pautz (2006) offers an argument very similar to the above but he doesn't cast it in terms of actual shifted spectrum cases. My assumptions for the actual cases seem less expensive: all I need are two hard-to-deny assumptions — see below.

phenomenal character of their experiences would be constituted by *C*. So representationalism implies that their experiences have identical phenomenal character. But this implication is falsified by the fact that the phenomenal character of their experiences differs as revealed by Mary's report. We arrive at this conclusion by making two very plausible assumptions. First, in many actual shifted spectrum cases, like in Sam and Mary, the color quality space is roughly the same and implemented by same or similar (more centrally located) neural structures — there is good empirical evidence for this assumption. Second, it is metaphysically possible for people with shifted spectrum to have slightly different evolutionary histories so that whatever the naturalistic psychosemantics are needed for errorless representation, they are in place in Sam and Mary.

As far as I can tell, the only venue for reductive strong representationalists to resist this conclusion is to insist on the impossibility of the second scenario involving Mary, saying that there can at most be one evolutionary development of human color system that sets the correctness conditions of color experiences. There are plenty of actual individuals living among us with shifted color spectrum.¹⁶ The representationalist ought to claim that none of their kinds (except one, perhaps) could have acquired their shifted color vision through a relatively independent evolutionary process that selected for it. But the modal strength of this claim seems empirical, not metaphysical. Representationalists need to establish this claim as a metaphysical necessity. I don't see how that can be done. In fact, for all we know, it wouldn't be too surprising if it turns out that this claim is in fact *empirically* false. The claim that some people could have acquired their shifted color spectrum through a relatively independent evolutionary process that selected for it is clearly *nomologically possible*, although for contingent factors it may be empirically very unlikely. We just don't know.

I conclude that a Dretskean psychosemantics doesn't deliver the kind of phenomenal externalism for which the view has been advertised. As I said, I'll generalize this conclusion, without argument, to all extant naturalistic proposals (versions of informational and/or teleological psychosemantics) about what it is for a sensory state to *represent* a sensible property. This is because, it seems to me, whatever naturalistic conditions are required for sensory states to represent sensible properties, they can be met in such a way that not only metaphysically but also nomologically allows for errorless representations of a single shade of color with demonstrably different phenomenal characters, or for there being the same phenomenal character correctly representing different "shifted" colors.

6 Conclusion and an alternative internalist picture

So, we still don't have phenomenal externalism. This conclusion shouldn't be all that surprising. It is difficult to fathom a philosophical account of sensory perception that accepts the Common Kind Thesis and offers a truly phenomenal externalist position. Non-

¹⁶ Hardin (1993: 79–80), Kuehni (2004).

reductive representationalism has been uniformly phenomenal internalist. It would have been somewhat perplexing if reductive representationalism of the Dretskean sort had turned out to be phenomenal externalist. If you have sympathies for phenomenal externalism you should look at the naïve realist or disjunctivist camp — although I would not hold my breath for their ability to successfully deal with shifted spectrum cases either. For my money, the overall conclusion to draw is that phenomenal externalism is just false. If reductive strong representationalism entails phenomenal externalism, then, it too is false. In fact, once it is realized that, for representationalists, instances of sensible properties we are sensorially aware of play no constitutive role (as opposed to a causal role) in determining the phenomenal character of our sensory awareness, the job of finding a constitutive role for a sensible property (*qua* universal, in terms of sensorially representing it) becomes somewhat obscure, and as they say, “academic.” But a naturalistic story about how this intentional property (sensorially ‘representing *P*’) is acquired doesn’t deliver a constitutive role for the universal either: Sam and Mary are related to the same shade (universal) when they look at the same chip and accurately represent its color but their experiences have different phenomenal character. Once the role of property instances is reduced to causal but not constitutive determination of phenomenal character, all the intuitions start crying out for an internal contribution to the metaphysical determination of color phenomenology. A naturalistic psychosemantics, as we have seen, doesn’t change this at all.

As an alternative, I offer the following picture, which is naturalist, intentionalist, but phenomenal internalist. Let’s treat **S** as before having internally interdependent state types, S_i , purporting to indicate instances of most determinate color shades along the axes of the color quality space. **S** is a genetically transmitted and neurofunctionally specifiable system with an informational function. I will just say that the particular states of **S** — $\mathbf{S}(S_i)$ — all *purport to indicate* instances of colors. This *general* fact (if it’s fact) may be necessary for *any* particular state to have some phenomenal character (with the background conditions in place). But what *particular* character they each will have may be (at least partly) internally determined at the level of engineering. And what instances of particular color shades each will purport to indicate may vary in different people having tokens of the same **S** — whatever empirical accommodations are required to explain the widespread phenomenon of “shifted color qualia.” We can think of S_i as sensory color predicates belonging to a system of representations (**S**). In other words, the states of **S** may be taken as syntactically structured representational vehicles whose semantic values are assigned according to local laws and whatever naturalistic psychosemantics is in place. They would still have the job of indicating/representing colors, yes, but without this fact metaphysically determining the *particular* phenomenal characters each may have. But one can still maintain that sensory systems having an indication function for a range of magnitudes for sensible properties is a necessary condition for sensory phenomenology to arise. In the older jargon, in other words, one may allow for the possibility of inverted or shifted qualia without *thereby* allowing for the possibility of absent qualia. Such a view needs to elaborate what it is about informational functions and the way they are imbedded in a larger, richer, and more

complex information processing architecture that allow them to reductively explain phenomenal character.¹⁷ Representationalists also typically accuse phenomenal internalists of violating perceptual transparency and of having implausible views about introspection. But these are relatively separate worries and have satisfactory resolutions anyway.^{18,19}

References

- Aydede, Murat, and Güven Güzeldere (2005). “Cognitive Architecture, Concepts, and Introspection: An Information-Theoretic Solution to the Problem of Phenomenal Consciousness.” *Noûs* 39 (2): 197–255.
- Aydede, Murat (2019). “Is the Experience of Pain Transparent? Introspecting Phenomenal Qualities.” *Synthese* 196 (2): 677–708.
- Aydede, Murat (2020). “What Is a Pain in a Body Part?” *Canadian Journal of Philosophy* 50 (2): 143–58.
- Brewer, B. (2011). *Perception and Its Objects*. Oxford University Press.
- Campbell, J. (2002). *Reference and Consciousness*. Oxford University Press.
- Dretske, F. (1995) *Naturalizing the Mind*. MIT Press.
- Dretske, F. (1999). “The Mind's Awareness of Itself.” *Philosophical Studies* 95 (1): 103–24.
- Hardin, C. L. (1993). *Color for Philosophers* (Expanded Edition). Indianapolis: Hackett.
- Gottlieb, Joseph, and Ali Rezaei (2021). “When Nothing Looks Blue.” *Synthese*, May. Springer Netherlands, 1–9.
- Gow, Laura (2018). “Why Externalist Representationalism Is a Form of Disjunctivism.” *Ratio* 31: 35–49.
- Kuehni, R. G. (2004). “Variability in Unique Hue Selection: A Surprising Phenomenon.” *COLOR Research and Application* 29 (2): 158–62.
- Martin, M. G. F. (2004). “The limits of self-awareness.” *Philosophical Studies* 120: 37–89.
- Pautz, A. (2006). “Sensory Awareness Is Not a Wide Physical Relation.” *Noûs* 40 (2): 205–240.

¹⁷ See Aydede & Güzeldere (2005) for a comprehensive attempt.

¹⁸ See Aydede (2019, 2020) for internalist explanations of transparency and introspection.

¹⁹ This paper grew out of a small discussion group discussing Pitt (2017) and Gottlieb & Rezaei (2021). I'd like to thank the authors for providing such a stimulating material as well as for their valuable comments on an earlier version. I'm also grateful to Dom Alford-Duguid, Jonathan Cohen, Matt Fulkerson, and Laura Gow for their comments and questions.

- Pitt, David (2017). “The Paraphenomenal Hypothesis.” *Analysis* 77 (4): 735–41.
- Sainsbury, Mark (2019). “Loar on Lemons.” In *Sensations, Thoughts, Language Essays in Honour of Brian Loar*, edited by Arthur Sullivan, Oxford UP.
- Tye, Michael (2013). “Transparency, Qualia Realism and Representationalism.” *Philosophical Studies* 170: 39–57.
- Tye, Michael (2014). “What Is the Content of a Hallucinatory Experience?” In *Does Perception Have Content?* Edited by Berit Brogaard, 1–22. Oxford UP.
- Tye, Michael (2015). “Yes, Phenomenal Character Really Is Out There in the World.” *Philosophy and Phenomenological Research* 91 (2): 483–488.