

Experiences Are Representations

An Empirical Argument

ADAM PAUTZ

To a first approximation, *representationalism* is the hypothesis that experiences are representational states akin to beliefs. I will sketch an argument for representationalism on the basis of an inference to the best explanation. I will not attempt to show that it is superior to every alternative. Instead I will focus on its main rivals—namely, “naïve realism” and the “inner state view”.

My plan is as follows. I will start with contemporary naïve realism, defended by John Campbell, Bill Fish, and Michael Martin, among others. I will argue that it violates *internal dependence*: the empirically determined role of the *internal processing* of the brain in shaping phenomenal character. This will bring us to the inner state view, defended by Ned Block, Brian McLaughlin, and David Papineau, among others. I will argue that, while it accommodates “internal-dependence”, it fails to accommodate the essential “externally directedness” of experience. Finally, I will argue that only *representationalists* can adequately explain both of these features of experience. The upshot is a largely empirical case for representationalism.

At the end, I briefly address what I consider to be the most profound objection to representationalism. Representationalism faces an overlooked modal puzzle about what experiences are possible.

1. Naïve Realism Violates Internal-Dependence: An Empirical Refutation

I will introduce “naïve realism” by way of what I consider to be the best argument for it (Pautz 2010, pp. 283–295): To illustrate, suppose you see a blue ball (Figure 2.1). The most basic philosophical question about experience is this: what



Figure 2.1 A blue-looking ball

grounds the qualitative character of experience? Now there is something to this thought: it is obvious without argument that in such a case the character of your experience is *grounded in* nothing but your being acquainted with the bluish color and round shape of some concrete item. If so, then there are a couple of live options as to what this item is. It may be a mental image (a “sense datum”) created by the brain (the “sense datum theory”), or it may just be the ball itself. Since there are well-known problems with the first option, we have a case for accepting the second. The result is naïve realism: the character of your experience is grounded in your being acquainted with the color and shape of the physical object.

In general, naïve realists hold that, even before sentient creatures evolved, external items possessed multiple objective (mind-independent) sensible properties: color properties, smell properties, loudness levels, and so on. Objects also possessed viewpoint relative but objective shapes, like *being elliptical from here*. The role of brain is not to construct experience. Rather the brain “opens the window shutter” to reveal objective properties of the items in the world. That is, when the brain responds to these objective sensible properties in the biologically normal way, this *enables* the mind to “reach out” and become acquainted with them. This long causal process is the *supervenience-base* of worldly acquaintance. In such veridical cases, the qualitative character of your experience is fully grounded in what external states you are acquainted with. The naïve realist John Campbell (Campbell and Cassam 2014, p. 27) also mentions your “point of view” as a factor, but then he says (p. 28) that this too is just a matter of which external states in the scene you are acquainted with (together perhaps with your own location in space).

Naïve realism is externalist. The brain configures qualitative character only to a very limited extent: only to the extent that it selects what objective external states we get to be acquainted with. For instance, pigeons have different color experiences than humans, only because their different sensory systems enable them to be acquainted with different objective colors (constituted by UV light). As Campbell (2010, p. 206) puts it:

[Naïve realism holds] that qualitative properties are in fact characteristics of the world we observe; our experiences have the qualitative characters

that they do in virtue of the fact that they are relations to those aspects of the world. So looking for the qualitative character of experience in the nature of a brain state is looking for it in the wrong place; we have to be looking rather at the [properties] of the objects experienced.

So much for what naïve realism is. Is it right? I think that there are serious problems with the simple argument for it as just outlined, even if I consider it to be the best argument for naïve realism (Pautz 2010, pp. 295–297). I also think that there are much stronger arguments *against* naïve realism. Many argue that naïve realists cannot adequately explain illusion and hallucination. Here I will set this issue aside and develop a new line of argument: even in “normal” cases, naïve realism is inadequate. Psychophysics has shown that, even in normal cases, qualitative similarity is very poorly correlated with external physical similarity. At the same time, neuroscience has shown that *neural* similarity is the only accurate predictor of qualitative similarity. In short, the typical situation is that there is “good internal correlation” even while there is “bad external correlation”. Naïve realists like Campbell, Fish, and Martin neglect the scientific facts. They have it the wrong way around. Looking for the basis of qualitative character in the external world is looking for it in the wrong place; we have to be looking rather at the brain.

For example, Figure 2.2 shows reflectances typical of purple-looking grapes, a blue-looking ball, and a green-looking leaf.

By any natural measure, it is not the case that the reflectance of the ball objectively resembles the reflectance of the grapes more than the reflectance of the patch of grass. (In fact, if anything, the opposite is true.) Nevertheless, the blue appearance of the ball resembles the purple appearance of the grapes much more than the green appearance of the leaf. So there is “bad external correlation”.

Since the explanation of similarity in color appearance is not to be found in the external world, there is reason to think it must reside in the brain. And

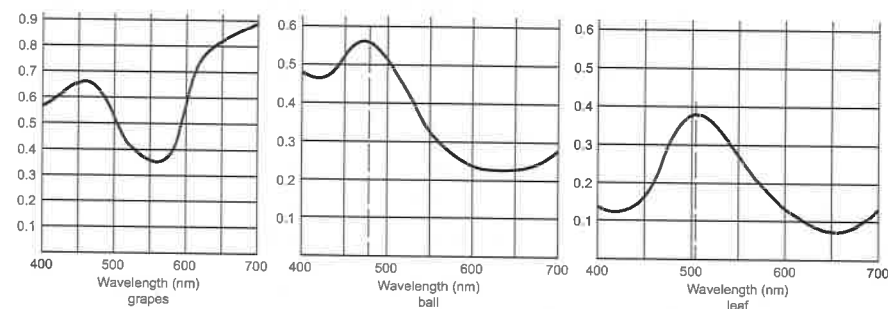


Figure 2.2 Reflectances of grapes, a blue-looking ball, and a leaf. From MacAdam (1985)

this is exactly what recent neuroscience suggests. Neuroscience demonstrates “good internal correlation”. As Brouwer and Heeger write,

The visual system encodes color by means of a distributed [neural] representation [in area V4] . . . similar colors evoke similar patterns of [neural] activity, and neural representations of color [in V4] can be characterized by low-dimensional “neural color spaces” in which the positions of [experienced] colors capture similarities between corresponding patterns of activity.

(2013, p. 15454)

So, your *internal V4 neural representation* of the blue-looking ball resembles your *V4 neural representation* of the purple-looking grapes more than your *V4 neural representation* of the green-looking leaf. This is the only available explanation of the resemblance-order among your color experiences.

Now let’s take a parallel example involving smell. Consider the chemical properties in Figure 2.3 below.

It is not the case that the middle chemical-type, *R*-limonene, resembles citral more than *R*-carvone. Rather, it resembles *R*-carvone more closely. Nevertheless, the perceived *smell quality* of *R*-limonene resembles the perceived smell quality of citral much more than the perceived smell quality of *R*-carvone. *R*-limonene and citral smell *different*, but their apparent smell qualities can both be described as “citrus-like”. That is why I call them citrus smell “1” and “2”. By contrast, *R*-carvone smells “minty”. This is another case of “bad external correlation”.

At the same time, neuroscience demonstrates “good internal correlation”. Howard and co-workers (2009) found that “spatially distributed ensemble activity in human posterior piriform cortex (PPC) coincides with perceptual ratings of odor quality, such that odorants with more (or less) similar fMRI patterns were perceived as more (or less) alike”. For instance, they found that your PPC neural representation of *R*-limonene resembles your PPC neural representation of citral more than your PPC neural representation of *R*-carvone, in perfect agreement with the character of your smell experiences. As Margot (2009) says, “Because the chemical structure of the odors in [the citrus] odor category are very different, this is strong support for the idea that the PPC codes odor quality rather than structural and chemical similarity”.

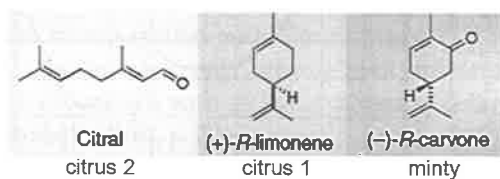


Figure 2.3 The first two chemicals smell citrus-like (“citrus 2” and “citrus 1”); the third smells minty. From Margot (2009)

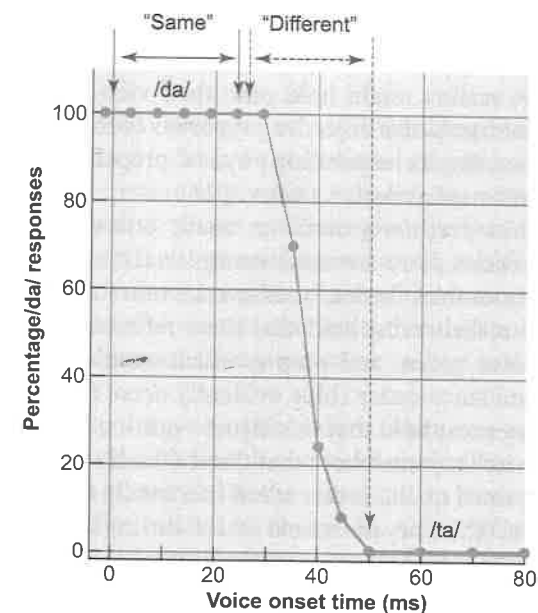


Figure 2.4 From Goldstein (2009)

Finally, consider auditory experience. As Figure 2.4 illustrates, if you continuously increase the “voice onset time” of a speech signal (the time between opening the lips and the onset of vocal fold vibration), then suddenly at 30 ms there will be a big, categorical change in the audible quality, from /da/ to /ta/.

This categorical change in the perceived sensible property *corresponds to no categorical change in the objective stimulus*. It corresponds only to a categorical change in your *neural representation* in the brain (Chang et al. 2010).

In addition, even under normal conditions, perceived loudness is related in an enormously complex, non-linear fashion to a number of objective physical properties, including intensity, frequency, and “critical bands”. By contrast, it is related in a simple fashion to the total neural activity produced by a sound, according to standard models (Moore 2003).

Given bad external correlation and good internal correlation, naïve realists’ externalist approach fails. To show this, I offer two arguments.

First, the *argument from irregular grounding*. Naïve realists hold that sensible properties (color properties, smell and taste properties, loudness levels) are brain-independent properties of physical items. What, in their view, is the relationship between these objective sensible properties of these items, and the underlying ordinary physical properties: reflectances, chemical properties, and acoustic properties?

Bad external correlation (illustrated in Figures 2.3 and 2.4) means that sensible properties and the underlying physical properties fall into different

resemblance-orders. This rules out the view that sensible properties are *identical with* the underlying physical properties (Pautz 2016).

However, naïve realists might hold onto their view by claiming that sensible properties are *irreducible* objective properties of things that are *grounded in*, but *distinct from*, the corresponding physical properties. In that case, their resemblance-orders needn't match (Allen 2015).

But this requires *irregular grounding*: totally unsystematic and arbitrary grounding connections. Here are some examples. (I) As Figure 2.2 shows, the ball's reflectance resembles the leaf's reflectance more than the grape's reflectance. Still, naïve realists must hold that these reflectances ground objective colors—namely *blue*, *green*, and *purple*, which stand in a totally *different*, *autonomous* resemblance-order (blue evidently *doesn't* resemble green more than purple). They must hold that this is just a quirk of reality with no explanation. (II) Likewise they must hold that the chemical structures in Figure 2.3 ground objective smell qualities that stand in a totally different resemblance-order than they do. (III) They must hold that if varying the voice onset time for a speech signal continuously, then at precisely 30 ms there is a big, discontinuous “jump” in its irreducible, objective audible quality, from /da/ to /ta/. Since they take this audible quality to be *independent* of our neural response, they have *no explanation* of why the jump takes place precisely there. (IV) Finally, since they hold that loudness is objective and entirely independent of neural response, they have no explanation of why the loudness of a simple tone doubles when there is a tenfold (10db) increase in its intensity, rather than (say) when there is a twenty-fold increase.

The broadly internalist approaches we will look at later—the “inner state view” (§2) and “brain-based representationalism” (§3)—*deny* the claim of naïve realism that the sensible properties (sensible colors, smell properties, audible properties) of items are explained by those items' objective physical properties alone. Rather, according to them, qualitative facts are always (at least partially) grounded in our *neural responses* to external items. We know that our neural responses, unlike the physical properties of the items we perceive, nicely line up with qualitative character (“good internal correlation”). So the grounding connections posited by internalist theories are much more *systematic and regular* than the ones required by naïve realism. Since it is a truism that we should prefer more systematic theories, we should prefer an internalist approach to naïve realism. Again, naïve realists like Campbell are just wrong in holding that to explain qualitative character we should ignore the brain and just look at the objective properties of the things we perceive.

The naïve realist Bill Brewer suggested a response to me.¹ There are, he says, other cases of irregular grounding. For instance, in their non-aesthetic properties, painting *a* might resemble painting *b* less than painting *c*, even though *a* and *b* are both beautiful, whereas *c* is ugly. Or, again, high-fiving someone and slapping someone are physically quite similar, but they are morally very

different (example due to Brian Cutter). If we have to accept some cases of irregular grounding, why is it so bad if naïve realists excessively multiply them with the likes of (I)–(IV)?

But, first, I dispute that these are cases of irregular grounding: if in these cases the grounding-bases include our emotional responses, then similarities and differences in the grounded properties do line up with similarities and differences in the grounding properties. Second, even if we must accept *some* cases of arbitrary, unsystematic grounding in nature, this doesn't undermine my basic point: we surely should *prefer* theories with more systematic grounding connections. So we should prefer internalist theories to naïve realism.

My second empirical argument against naïve realism is the *argument from hypothetical cases*.

First, a preliminary point—on naïve realism, when you perceive a ball or a smell, you are acquainted with its color or smell quality, but not its electronic charge. Why? The only answer is that the visual system is *causally responsive* to its color and smell quality, but not its charge. So, naïve realists must hold that your acquaintance with simple property-instances is grounded in your bearing some *causal* relation, *R*, to those property-instances.

Now consider an example involving smell. Recall that objective chemical similarity badly fails to predict qualitative similarity among smells; instead, it is only *neural similarity* that corresponds with smell similarity. So, for instance, if you smell the chemical types shown in Figure 2.3, then your PPC neural representation of *R*-limonene will resemble your PPC neural representation of the citrus-smelling citral much more than your PPC neural representation of the mint-smelling *R*-carvone. Your smell experiences are perfectly in line with this: *R*-limonene produces in you an experience of a citrus smell quality, one resembling the citrus smell of citral much more than the mint smell of *R*-carvone.

Imagine now that you have a “twin” on a “Twin Earth”. We make two stipulations. *First*, your twin belongs to a human-like species that is just like our own but for one thing: because of differences in their evolutionary history, their postreceptor wiring for smell differs from our own. Because of this, your twin's PPC neural representation of *R*-limonene resembles his PPC neural representation of the *minty-smelling R*-carvone more than his PPC neural representation of the citrus-smelling citral—the opposite of what we find in a normal human like yourself. Consequently, while you sort *R*-limonene with the citrus-smelling citral, your twin sorts it with the minty-smelling *R*-carvone. *Second*, we stipulate that, when you and your twin smell these odorants, you both bear complex causal relation *R* to exactly the same complex olfactory properties of those odorants.

Given these physical differences, what is the correct verdict on how your twin's smell experiences would compare to your own? Previously, we noted that the empirical fact that neural similarity is the *only* good predictor of smell similarity and that objective chemical similarity is a lousy predictor

of smell similarity. So, given this empirical fact, the only reasonable verdict is that, whereas *R*-limonene smells *citrus-like* to you, it smells *minty* to your twin. This also fits with their sorting differences. The case is a hypothetical “twin earth case”, but this verdict about the case is supported by empirical facts, not our “intuitions”.

By contrast, naïve realism implies that you and your twin have exactly the same smell experiences. True, you two undergo totally different PPC neural states, falling into different resemblance-orders and leading to different behaviors. But, we stipulated that, in doing so, you both bear causal relation *R* to the very same complex olfactory properties of those odorants. Therefore, on naïve realism, your and your twin’s neural states, although they fall into different resemblance-orders and lead to different behaviors, “enable” you to be acquainted with the very same objective olfactory states. On naïve realism, this means that there can be no qualitative differences between your and your twin’s experiences, despite all the neural and behavioral differences telling against this verdict

Here is a similar case. You and your twin view the grapes, ball, and leaf with reflectances shown in Figure 2.2. Again, we make two stipulations. *First*, your twin is just like you but for one thing: because of naturally evolved differences in his species’ postreceptoral neural wiring, your twin’s V4 internal neural representation of the ball is more like his V4 neural representation of the green-looking leaf than his V4 neural representation of the purple-looking grapes—the opposite of the actual situation with you. Consequently, whereas you sort the ball with the grapes, your twin sorts the ball with the leaf. *Second*, you and your twin nevertheless bear the acquaintance-grounding causal relation *R* to the very same objective chromatic properties of the ball, the grapes, and the leaf (the ones that, on naïve realism, are their colors). This case is illustrated in Figure 2.5.

This case is certainly possible. Compare how a mercury thermometer and a thermoelectric thermometer have different internal states but detect the same temperatures and temperature-differences.²

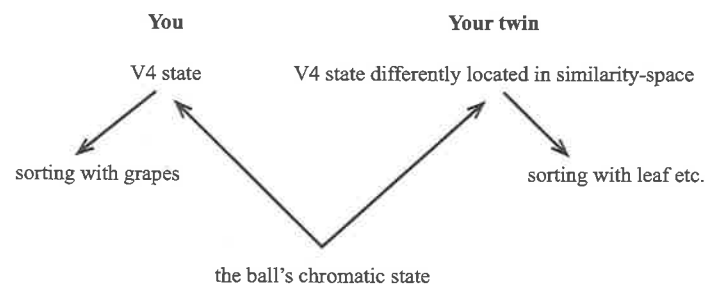


Figure 2.5 A hypothetical case

We previously noted the empirical fact that neural similarity is the only good predictor of similarity in color experiences. So, given the empirical facts, the only reasonable verdict is that you and your twin have different experiences. Whereas to you the ball looks *bluish*, to your twin it looks *greenish*. But, again, naïve realism delivers the empirically implausible verdict that they have exactly the same color experiences, despite the radical neural and behavioral differences between them, since it implies that they are exactly alike in what objective states they are acquainted with. Further, what Campbell (Campbell and Cassam 2014, p. 28) calls your “point of view” on the ball is the same: you both perceive the ball from the same spatial location, and you are acquainted with exactly the same subset of states involving the ball.³

These examples show that, besides requiring “irregular grounding” (my first argument), naïve realism requires giant flukes. Unlike an internalist view, naïve realism implies that similarity in the “enabling” neural representations *has nothing to do with* qualitative similarity in our experiences—just as similarity among words (e.g., “Ned” and “Fred”) has nothing to do with similarity among the people and things they are about. The cases involving you and your twin show this. The similarity relations holding among your twin’s neural states are *totally different* from those holding among yours; nevertheless, as we saw, on naïve realism, this *makes no difference* to character of his experiences—there is *no* qualitative difference between his experiences and yours. So in him there is a mismatch between neural similarity and qualitative similarity. Therefore, naïve realism has the absurd implication that it is just a fluke that in normal humans like yourself there is quite generally a nice agreement between neural similarity and qualitative similarity (“good internal correlation”)—somewhat as it would be a fluke if it turned out that similar-looking names (e.g., “Ned” and “Fred”) always named similar-looking individuals.

Again, Bill Brewer has suggested a response.⁴ On this response, the externalist approach of naïve realism is correct after all. So, you and your twin do have the same experiences *despite* the radical neural and behavioral differences—then what about “good internal correlation”, which seems to show that the neural differences should make for qualitative differences? Brewer’s response is that we would expect good internal correlation *even if* naïve realism’s externalist theory of qualitative character is correct.

To illustrate, look at citral and *R*-limonene in Figure 2.3. These chemical structures are totally different by any objective measure. Nevertheless, as we saw, naïve realists must say that even before we evolved they grounded similar objective citrus-like smell qualities (“irregular grounding”). According to Brewer, we would expect that we should have evolved so that the neural states that enable us to be acquainted with these similar objective qualities are themselves similar. The neural similarities are patterned by the real similarities in the objective smell qualities.

However, this response fails for a simple reason. Even if naïve realists are right that the chemically very *different* structures citral and *R*-limonene ground very *similar* objective smell qualities (despite my objection against such “irregular grounding”), this is no reason to expect that we should evolve to detect them with similar neural states. The similarities among a creature’s neural states simply aren’t determined by similarities in these alleged objective, autonomous qualities. Rather, they are determined by physical similarities and differences (e.g., similarity in chemical structure or in reflectance), by how these interact with the receptor systems, and by postreceptor writing further downstream, which is itself dependent on the species’ unique ecology together with the vagaries of the evolutionary process.

So the argument from hypothetical cases stands. Given their radically externalist approach, naïve realists cannot explain “good internal correlation” and must consider it a giant fluke. And they must accept the empirically implausible view that, in the cases I described, you and your twin have the same experiences *despite* the vast neural and behavioral differences. This is unacceptable. In short, naïve realism violates *internal-dependence*. We need a more internalist view.

2. The Inner State View Violates External Directedness

I will now consider the simplest internalist view of experience: the *inner state view* recently defended by Ned Block, Brian McLaughlin, and David Papineau. Naïve realism holds, roughly, that qualitative character is grounded in relations to the environment. The inner state view is the polar opposite. It holds that having an experience with a certain character is necessarily identical with having some intrinsic physical-functional property, such as a complex distributed neural pattern in the visual system or the olfactory system. All aspects of experience can be reductively explained in internal neuro-functional terms. That exhausts the essence of experience.

This view evidently accommodates internal-dependence. Still, I think we should reject the inner state view. For experience is also essentially “externally directed”. It is essentially *spatial*.

Consider, for instance, the visual experience of the blue ball in Figure 2.1. Call this specific type of experience “the ball experience”. The following is obvious:

External Directedness. Necessarily, if an individual has the ball experience, then that individual has an experience *as of* a bluish and *round* thing in a certain *viewer-relative place*, even in hallucination cases where there is no such thing. So any such individual has an experience that fully “matches the world” only if a bluish and *round* thing is present in some kind of space.

This is pretheoretical; it doesn’t presuppose any controversial “representationalist” ideas (Pautz 2010, pp. 269–270).

Here is a related point, one emphasized by Bertrand Russell in *The Problems of Philosophy* (chap. 10):

Thought-Grounding Role: Necessarily, if you have the general capacity for thought, and if you have the ball experience, then this *explains* your having the additional capacity to have thoughts (e. g. *something is that way*) that are true only if something has the property *being round*. Further, if you have the ball experience together with other such experiences, then you are *in* a position to know certain truths about *shape and color properties*, for instance their patterns of similarity. These truths lack nominalistically acceptable paraphrases: they are *essentially* about general properties.

(Pautz 2010, p. 279)

In what follows, I will use “external directedness” to cover both of these two claims.

To see why the inner state view violates external directedness, consider a *brain in the void*, or BIV for short. This BIV popped into existence out of the blue in an otherwise empty universe. BIV only has the neural machinery required to have the neuro-computational state, which, on the inner state view is necessarily identical with the ball experience. Call that neuro-computational state “*N*”. Let’s also stipulate that it bears no interesting physical relation to the property *being round*. This property is not instantiated *in* BIV. And BIV doesn’t (and can’t) track the instantiation of this property in the outside world. Since BIV bears no interesting *physical* relation to *being round*, inner state theorists must hold it bears *no interesting relation at all to being round* (for they accept an austere physicalism).

It follows that, if the inner state view is true, BIV cannot be said to have an experience that *matches the world* only if something is round—that is, has the property *being round*. It also lacks the capacity to *predicate* this property of anything in thought. So, in this scenario, the neural pattern *N*—and hence, on the inner state view, the ball experience—doesn’t count as an experience as of a round thing and doesn’t ground the capacity to have roundness-involving beliefs or knowledge. So the inner state view violates external directedness.

In short, considered in itself, the inner neural-computational state *N*—and hence, on the inner state view, *the ball experience*—is no more essentially related to the property of being round than it is essentially related to (say) the property of being a pink elephant. In this way, it is like the orthographic sequence “r-o-u-n-d”, or a headache that was in fact caused by getting hit on the head by a round ball. This result is just clearly wrong. Normal visual experiences are *essentially spatial*; there is some *essential* connection between visual

experiences and spatial properties and relations, properties that needn't be instantiated inside the head when we have those experiences.⁵

We need a theory of experience that accommodates “external directedness” as well as “internal-dependence”.

3. Experience as Representation: How Experience Might Be Externally Directed and Internally Dependent

I am now ready to explain how a “representational” approach to experience might accommodate *both* external directedness and internal-dependence. It is a middle way between the extremes of naïve realism and the inner state view. I will focus on the ball experience, but afterward I will generalize the argument.

Take external directedness first. The simplest explanation is there is a single mental relation *R* such that (I) having the ball experience *essentially* involves standing in *R* to the property *being bluish and being round*, (II) standing in *R* to properties grounds one's ability to predicate those properties in thought and to know truths about those properties (Russell 1912), and (III) bearing relation *R* to some properties implies its *seeming* to you that they are instantiated out there in the world.

This hypothesis explains external directedness. That is, it explains the fact that, *necessarily*, if you have the ball experience, then you have an experience that matches the world only if something has *these properties* (rather than some other properties), and you *thereby* have the capacity to predicate *these properties* to things in thought and to learn truths about them. If having the blue ball experience didn't involve standing in such a relation to these specific properties, what else could explain these facts?

Now the question arises, what is the *nature* of the relation with these features?

Sense datum theorists like Russell (1912) believed in such a relation to color and shape properties and called it “acquaintance”. They held that what it is to bear such a relation to a property like *being bluish and round* is to be acquainted with the *instantiation* of that property by “sense datum” created by the brain. This view accommodates internal-dependence. It also accommodates external directedness, even though sense data are brain-created. For it holds that sense data literally have spatial properties and relations and occupy a kind of space.

But the sense datum view faces problems—for instance, problems indeterminate of experiences and experiences of impossible scenes (Pautz 2010, pp. 280–281). The best view is that in non-veridical cases the property *being bluish and round* is not instantiated by *anything* in your vicinity at all. Therefore, we can add a thesis to our proto-theory: (IV) you can bear *R* to some properties even though they are not instantiated by anything in your vicinity. This makes the theory broadly representationalist.

Let us call the hypothesized relation *R* with features (I)–(IV) “the perceptual predication relation” or the “phenomenal representation relation”.⁶ The inner state view is inconsistent with the thesis here that having the ball experience essentially involves perceptual predicating the spatial property *being round*, for it holds that the complete essence of this experience can be completely given in purely internal, neural-computational terms.

Now so far our minimal representationalist theory is actually *consistent with* a form of naïve realism. For naïve realists might hold that, in all cases in which we have the ball experience, we phenomenally represent the property *being bluish and round*, but add that *in veridical cases* this is somehow “grounded in” our being acquainted with the instantiation of this property by a real blue sphere (Pautz 2010, pp. 295–296). Then phenomenal representation is externally determined.

It is only when we add the empirical fact of internal dependence that we get a view that is inconsistent with naïve realism. For, given internal dependence, we must add a clause to our theory: (IV) *what sensible properties (sensible colors, smell qualities, audible qualities) we phenomenally represent are entirely fixed by the intrinsic properties of the subject*. For instance, going back the hypothetical case discussed in §1, when you and your twin view the ball, then, due to your different V4 processing, you two phenomenally represent distinct sensible colors (*being bluish* and *being greenish*) as co-instantiated with *being round*. This is a “brain-based” representational theory.

Now this may seem puzzling. It seems to us that sensible properties are instantiated in *external* space. How can this depend on *internal* processing? Elsewhere I have addressed this puzzle (Pautz 2016). I think that the solution involves recognizing that sensible properties are not real, objective properties of external things; rather, they are “projected” properties, or “virtual properties”. (In fact, because of relativity theory and quantum mechanics, Chalmers [2012, p. 333] extends this kind of neo-Galilean view to the *spatial properties and relations* that we phenomenally represent.) This view is consistent with representationalism but not naïve realism, since naïve realism requires that the sensible properties are real, objective properties of external items.

The argument is general. Different experiences involve the seeming-presence of different properties and ground the ability to think about and know about different properties. Russell (1912, chap. X) was right to emphasize this point, but he was wrong to think that the properties must be instantiated by things we perceive (“sense data”). The best view is that having different experiences necessitates *perceptual predicating* different property-complexes, where the relevant property-complexes *needn't be instantiated*. So, for instance, having a certain auditory experience might entail perceptual predicating the property-complex $\lambda x \lambda y$ (x is high-pitched and to the left, and y is low-pitched and to the right), and having a certain smell experience might entail perceptual predicating λx (x is citrus-like and in the region before me).

We have arrived at a necessary covariance between experience and phenomenal predication. This leaves open what the “direction of explanation” is. There are three options. (I) Having an experience with a certain character is distinct from, but essentially grounds, standing in the *phenomenal predication relation* a certain complex property, an abstract object. (Analogy: the concrete fact that John and Mary are in the room grounds the fact that the people in the room stand in the *numbering* relation to the number two, an abstract object.) (II) Perceptual predicating a certain property-complex is distinct from, and grounds, having an experience with a certain character. (III) Having an experience with a certain character *just is* perceptual predicating a certain property-complex, allowing no room for an explanatory, grounding connection between them.

I think options (I) and (II) raise puzzles. If experience properties and representational properties are *distinct*, then what explains the necessary connection between them? And what is the non-representational nature of experience that grounds its representational nature? I favor option (III) because it is simple and avoids such questions. This is a strong “identity” form of representationalism (Pautz 2010, Chalmers 2013, Tye 2014).⁷

4. A Neglected Problem for Representationalists: The Laws of Appearance

Finally, having sketched an argument for representationalism, I turn to what I consider to be its most difficult and interesting problem—a problem that has been almost entirely neglected. Roughly, the problem is that if experiences are just representations, then why can’t you have experiences that represent certain *extremely bizarre* scenarios? This requires explanation.

On representationalism, experience essentially involves the phenomenal predication of *complex properties*. So experiences can be associated with *propositional contents*. For instance, if you have an experience in which you perceptually predicate the complex property $\lambda x \lambda y$ (x is red, and y is green, and x is to the left of y), then you phenomenally represent the proposition $\exists x \exists y$ (x is green, and y is red, and x is to the left of y). If your experience is non-hallucinatory, you also phenomenally represent a singular proposition attributing this complex property to some specific viewed objects.

Now the start of our puzzle is that *not every content corresponds to a metaphysically possible experience*. In particular, at least *some* of the following modal claims about the limits of experiences are true:

- I. It’s impossible to have an experience with the content *something is pure blue and also greenish blue*. So-called pluralists about color (Pautz 2016) *believe* such contents, but it’s impossible to have an *experience* with such a content. (Here and in what follows, by “impossible”, I mean *metaphysically impossible*.)

- II. It’s impossible to have an experience with the content *something is spherical and cubicle*.⁸
- III. It’s impossible to have an experience with the content *blue is intrinsically overall more like green than purple*.
- IV. It’s impossible to have an experience whose only content is a wildly disjunctive content—for example, *that thing is round and green and directly in front of me or it is square and purple and 45 degrees to the left*.
- V. It’s impossible to have an experience whose *only* content is *that thing’s facing surface is round* (i.e., without any information about the chromatic or achromatic colors of the thing or its background).
- VI. It’s impossible to have an experience that only has the content *that thing is cubicle*, but that doesn’t have any *de se* content about the thing’s apparent shape from a particular point of view (as it were, a “God’s eye” visual experience of a cubicle object, from no point of view).
- VII. It’s impossible to have a single experience whose content is *a is red all over and b is wholly behind a* (Johnston unpublished).
- VIII. It’s impossible to have a visual experience in which one *phenomenally* represents a “high-level” content like *that is pine tree or that is a Republican*, but in which one *phenomenally* represents *no* “low-level” content at all. It’s also impossible to have an experience in which one *phenomenally* represents such a high-level content, but in which one *phenomenally* represents a *totally* incongruous low-level content, like *that very same thing is a giant round sphere right in front of me*.⁹

You might take the skeptical view that these strong assertions of *metaphysical* impossibility are false and that we only think they’re true owing to limits on our imagination. But I myself think that at least *some* such modal claims are true. Call them *laws of appearance*, since they are metaphysically necessary prohibitions on how things might appear.

Now we have general Occamist reasons to keep “brute modal facts” to a minimum. So, when faced with the laws of appearance, we have reason to derive them from some smaller set of more basic facts. Maybe they can be derived from more general metaphysically necessary truths (e.g., the most general, overarching laws of appearance). Or maybe they can be derived from a pre-modal *real definition* or *account* of *what it is* for an experience to have a certain content, or represent a certain state of affairs.

Here is an analogy. I think, with many others, that having certain totally bizarre sets of *beliefs* and *desires* is metaphysically impossible. For instance, it’s impossible that someone with my same past and current experiences and my same behavioral dispositions should be correctly interpreted as *believing* and *desiring* the *negations* of all the propositions I believe and desire. So there are also constraints on beliefs and desires (though they differ from the constraints

on experiences I–VIII). Now, in my view, such metaphysical impossibilities involving beliefs and desires *can* be explained. They can be explained by a *rationality-based account* of belief and desire in the tradition of Davidson and Lewis. So they are *not* “brute” modal facts but rather follow from the “real definition” of belief and desire.

Here, finally, is the problem for representationalists that I have been leading up to. *Representationalists apparently can provide no good explanation of why the “laws of appearance” obtain, so they must hold that they are all brute facts.* The only potential explanations I can think of face immediate problems.

One potential explanation is this. Impossible states of affairs and highly disjunctive states of affairs are acausal and therefore cannot be tracked. So, given a general “causal” or “tracking” account of what it is for an experience to have a certain content (Tye 2014, p. 51), they cannot be phenomenally represented. This might go some way toward explaining at least *some* laws of appearance, specifically I–IV.

But this explanation fails for two reasons. First, even if a tracking account of phenomenal representation were correct, it couldn’t explain *all* the laws of appearance. In fact, such a theory entails that V–VIII are *false*, for some possible visual systems *could* directly track the relevant states of affairs. Second, given internal-dependence, we must in any case reject “tracking representationalism”. For instance, according to what I take to be the correct representationalist account of the cases discussed earlier (§1), you and your twin *track* the same objective physical properties (reflectance-types, chemical types) but *perceptual predicate* distinct sensible properties (sensible colors, smell qualities).

Another potential explanation has it that the laws of appearance follow from the fact that *phenomenal representation is a “quasi-pictorial” form of representation* (a suggestion put to me independently by Chris Hill and Michael Tye).

But what does this mean? Maybe it just means that *phenomenal representation is necessarily a form of representation that obeys laws of appearance like I–VIII*. But then the claim doesn’t provide a substantive, non-circular *explanation* of these laws.

Alternatively, maybe the claim here is entirely about the way phenomenal representation is realized in the brain—that is, about the *format* of phenomenal representation (Kosslyn 1994). In other words, maybe it is about the nature of the *neural “content-vehicles”*, not in the first instance about the kinds of “contents” we phenomenally represent. The idea is that this claim nevertheless explains why we cannot phenomenally represent the kinds of contents described in I–VIII.

But, first, I do not see how any such a claim might logically entail *all* of the laws of appearance. In fact, there is a fundamental problem with the idea that it could explain *any* of them. The laws of appearance I–VIII claim that it’s *metaphysically necessary* that no one has certain bizarre experiences. By contrast, if

it is a truth at all, it is bound to be a *contingent truth* that phenomenally representation is realized in the brain in a quasi-pictorial format. And, as a matter of logic, a contingent truth cannot explain a necessary one. For instance, surely there are *possible* creatures for whom the realizers of phenomenal representation are sentence-like, rather than quasi-pictorial. (Representationalist theories are typically broadly functionalist theories allowing for “multiple realizability”.) Presumably, in such a creature’s “sensation box”, there might appear “sensation-sentences” that represent the bizarre contents mentioned in I–VIII. So even if in actual humans the format of phenomenal representation is “quasi-pictorial”, that is not enough to explain why the bizarre experiences described in I–VIII cannot occur in *other* possible creatures. Therefore, it is not enough to explain I–VIII, understood as claims of *metaphysical* impossibility.¹⁰

Let’s suppose, for the sake of argument, that representationalists *cannot* explain the “laws of appearance” and must take them to be brute axioms. Would this be a strong reason for rejecting representationalism? I don’t think so, because I think that all theories of experience face a parallel explanatory challenge. For instance, consider naïve realism, which I criticized previously on empirical grounds. Even on this view, we can ask exactly parallel explanatory questions. Why couldn’t you have certain very bizarre veridical or non-veridical experiences? For instance, why couldn’t you have a hallucination of a round square? Why couldn’t you be acquainted with the objective cubicle shape of something “neat”, without being ostensibly acquainted with any other fact about it (for instance, its viewer relative shape from here, or its apparent color)? I think it can be shown that contemporary naïve realists aren’t any better placed to adequately answer such questions than representationalists. So the problem of how to explain “the laws of appearance” is *everyone’s problem*. Yet it is a problem that philosophers of perception have hardly addressed.¹¹

Notes

1. In an “author meets critics” meeting at the 2014 meeting of the Central Division of the American Philosophical Association.
2. If the naïve realist nonetheless asserts that my two stipulated physical conditions are incompatible (Fish 2013, p. 59), then he or she must back this up with a plausible account of the causal relations that ground acquaintance from which this assertion follows. Elsewhere (Pautz 2011) I argue that this cannot be done.
3. Naïve realists might handle my cases by further revising or complicating their view. Logue (2012, pp. 223–225) says that, to handle my hypothetical cases, naïve realists might say that “features of the subject” (presumably, neural features) partly determine qualitative character. Another idea would be to handle the case by building on Campbell’s idea that experience is not simply a “two-place” relation between a subject and the perceived scene but a “three-place relation” modified by the “ways in which the scene is given” (Campbell and Cassam 2014, pp. 27–28). However, naïve realists must answer a crucial question. To you, the round surface of the ball looks *bluish*; to your twin, it looks *greenish*. Given the empirical facts, that is the most plausible verdict. The crucial question is this: what *are* these *distinct* sensible colors that appear to you and your twin to fill the round region of the ball? I think that even

modified naïve realist views cannot adequately answer this question. The naïve realist cannot answer that these distinct sensible colors are *objective, response-independent* features of the ball, since by stipulation you and your twin causally detect *the same* response-independent chromatic feature of the ball. Maybe Logue would suggest that naïve realists should instead answer that that they are *in fact* “features of the subjects” (e.g., features of you and your twin’s *different inner postreceptoral neural processing*), which, however, you and your twin experience as *properties of the round ball!* But this option would require a bizarre kind of projective error. Yet another suggestion (put to me by Craig French) would be that naïve realists adopt a Shoemaker-style view according to which the distinct sensible colors experienced by you and your twin are distinct “appearance properties” or “color-looks” of the ball, where these are *neural-relative, response-dependent* properties of the ball of the form *normally causing internal V4 processing X in the relevant population*. Typically, it is representationalists who propose such a view. Elsewhere I develop several objections showing that representationalists cannot accept this type of view (Pautz 2013). My objections there could be used, *mutatis mutandis*, to argue that naïve realists cannot adopt it either.

4. In the same meeting mentioned in note 1. Keith Allen suggested a similar response in discussion.
5. Elsewhere (e.g., Pautz 2010, pp. 266–272) I develop this argument in more detail. Papineau (2016, sect. 15) is an inner state theorist who agrees with me that inner state theorists must reject external directedness. However, he offers a surrogate claim using *similar-looking* spatial terminology: necessarily, if an individual (e.g., BIV) has the ball experience (that is, on his inner state view, the intrinsic neural-computational property *N*), then there is an internal “phenomenal object” that is “in” the individual’s experience and *that is round**. Now you might think, “This is very close to external directedness—so isn’t this good enough?” But don’t be deceived. By a “phenomenal object”, Papineau means something like a population of neurons, and by “visual roundness*” he means a neural-computational property of this population of neurons—a *property that is nothing like the property of being round* (that is, the property *having edges roughly equidistant from a common point*). So, despite his similar-looking spatial terminology, Papineau’s claim doesn’t *come close* to accommodating external directedness: the obvious fact that, necessarily, if you have the ball experience, then you are in a state that matches the world only if something before you *is round*, where “is round” expresses a *genuine spatial condition*.
6. Michael Tye (2014, pp. 51–52) is a representationalist who claims that we can be said to be *aware of* the uninstantiated properties we phenomenally represent. But it is important to realize that representationalists are not committed to this claim. I myself reject it because I find it very odd (Pautz 2010, p. 266). The oddness of the claim is especially clear in the case of uninstantiated *relations*, a case that Tye doesn’t consider. For instance, we *phenomenally represent* the relation of *spatial containment* when we represent one thing as being spatially within another thing. By representing this relation in experience, we somehow become directly aware of certain facts about it—for instance, that it is necessarily transitive. But it is odd in the extreme to say that in non-veridical cases one is *aware of* the abstract relation of spatial containment, even though it is not instantiated by anything before one.
7. While I favor identity representationalism, it should be noted that it faces couple of important objections. (1) One might think that there are possible worlds (“Eden worlds”) where naïve realism is true and representationalism is false, which would undermine the strong thesis of identity representationalism. (This may be Logue’s objection in her contribution to this volume.) For different possible responses, see Pautz (2010, fn. 33) and Chalmers (2013, p. 350). (2) Another objection to identity representationalism is simply that, intuitively, having a blue ball experience just can’t be identical with standing in a representation relation to a complex property, if this is understood as an *abstract object* of some kind. Intuitively, having an experience must be *distinct from* standing in a relation to such a peculiar abstract object. I developed and responded to this objection in previous papers (e.g., 2010, pp. 292–297). More recently, Papineau (2015, sect. 13) has taken it up; in fact, it is his chief objection

to representationalism. I think the objection fails for a couple of reasons. First, the objection only tells against “identity representationalism” and doesn’t tell against a “grounding” form of representationalism (along the lines of [ii] in the text) that *concedes* that having an experience is *distinct from* standing in a relation to an abstract object. Second, we have reason to be dubious of the “intuition” here. For if it is an “intuition of distinctness”, then it cuts equally against pretty much all views, including Papineau’s own *inner state view* (for, intuitively, having the ball experience is *also* distinct from being in an inner neural-computation state involving soggy grey brain matter). On the other hand, the intuition might be that it is odd or surprising that the account or real definition of what it is to have a certain experience should mention *abstract objects* like complex properties. But there is a track record of equally surprising accounts being correct in other domains. For instance, there are convincing arguments for supposing that the right accounts of what it is for two people to believe alike, or for two things to be exactly alike, also essentially appeal to “abstract” items—namely, propositions and properties.

8. You might think that the waterfall illusion shows that we can perceptually predicate two incompatible properties of the same thing. But my view is that the content of the waterfall illusion actually involves no real inconsistency at any particular time (Pautz 2010, p. 303).
9. Nanay (2012) argues that cases of unilateral neglect show that it is possible to phenomenally represent a certain *kind of* “high-level” contents (viz., one attributing what he calls *action-oriented properties*), while phenomenally representing “no level content”. This is an interesting interpretation of such cases, but it is also open to doubt—see Masrour (2011) and Raftopoulos (2015). Further, the cases Nanay discusses don’t cast doubt on my distinct claim that it is impossible to have an experience in which one phenomenally represents a high-level content like *that is a pine tree*, but in which one phenomenally represents a *totally incongruous* low-level content, like *that very same thing is a giant blue and round sphere right in front of me*.
10. Of course, in response, the representationalist might say: “Ok, I think that the experiences described in I–VIII are only *nomically* impossible. Contrary to what you say, they are not *metaphysically* impossible. The reason is that they could occur in other possible creatures for the reason you suggest”. I realize this is a possible response. But, as said at the outset, I think that it is just obvious that at least *some* of I–VIII are true, understood as strong claims of *metaphysical* impossibility. So my preference is to accept them, even if I cannot now provide any good explanation of them. (Thanks to Bence Nanay for discussion of this point.)
11. My thanks to Craig French and Bence Nanay for some very helpful comments on an earlier version of this paper.

References

- Allen, K. 2015. “Colour Physicalism, Naïve Realism, and the Argument from Structure.” *Minds and Machines* 25: 193–212.
- Brouwer, G. and D. Heeger. 2013. “Categorical Clustering of the Neural Representation of Color.” *Journal of Neuroscience* 33: 15454–65.
- Campbell, J. 2010. “Demonstrative Reference, the Relational View of Experience and the Proximality Principle.” In R. Jeshion (ed.) *New Essays on Singular Thought*, 193–212. Oxford: Oxford University Press.
- Campbell, J. and Q. Cassam. 2014. *Berkeley’s Puzzle*. Oxford: Oxford University Press.
- Chalmers, D. 2012. *Constructing the World*. Oxford: Oxford University Press.
- Chalmers, David J. 2013. “The Contents of Consciousness: Reply to Hellie, Peacocke and Siegel.” *Analysis* 73: 345–68.
- Chang, E. F., J. W. Rieger, K. Johnson, M. S. Berger, N. M. Barbaro, and R. T. Knight. 2010. “Categorical Speech Representation in Human Superior Temporal Gyrus.” *Nature Neuroscience* 13: 1428–32.
- Fish, W. 2013. “Perception, Hallucination, and Illusion: Reply to My Critics.” *Philosophical Studies* 163: 57–66.

- Goldstein, B. 2009. *Sensation and Perception*. Pacific Grove, CA: Wadsworth.
- Howard, J. D., J. Plailly, M. Grueschow, J. D. Haynes, and J. A. Gottfried. 2009. "Odor Quality Coding and Categorization in Human Posterior Piriform Cortex." *Nature Neuroscience* 12: 932–9.
- Johnston, M. unpublished ms. *The Manifest*.
- Kosslyn, S. 1994. *Image and Brain*. Cambridge: MIT Press.
- Logue, H. 2012. "Why Naive Realism?" *Proceedings of the Aristotelian Society* 112: 111–37.
- MacAdam, D. L. 1985. "The Physical Basis of Color Specification." In *Color Measurement: Theme and Variations*, 1–25. New York: Springer-Verlag.
- Margot, C. 2009. "A Noseful of Objects." *Nature Neuroscience* 12: 813–4.
- Masrour, F. 2011. "Does Perception Represent Affordances?" <https://consciousnessonline.files.wordpress.com/2011/02/comments-on-nanay-farid.pdf>
- Moore, B. 2003. *An Introduction to the Psychology of Hearing*. San Diego, CA: Academic Press.
- Nanay, B. 2012. "Perceptual Phenomenology." *Philosophical Perspectives* 26: 235–46.
- Papineau, D. 2016. "Against Representationalism (about Experience)." *International Journal of Philosophical Studies* 24(3): 324–47.
- Pautz, A. 2010. "Why Explain Visual Experience in Terms of Content?" In B. Nanay (ed.) *Perceiving the World*, 254–310. Oxford: Oxford University Press.
- Pautz, A. 2011. "Can Disjunctivists Explain our Access to the Sensible World?" *Philosophical Issues* 21: 384–433.
- Pautz, A. 2013. "The Real Trouble for Phenomenal Externalists: New Evidence for a Brain-Based Theory of Consciousness." In R. Brown (ed.) *Consciousness: Inside and Out*, 237–298. Berlin: Springer-Verlag.
- Pautz, A. 2016. "How Does Color Experience Represent the World?" In D. Brown and F. MacPherson (eds.) *The Routledge Handbook to the Philosophy of Color*. London: Routledge.
- Raftopoulos, A. 2015. "What Unilateral Visual Neglect Teaches us About Perceptual Phenomenology." *Erkenntnis* 80: 339–58.
- Russell, B. 1912. *The Problems of Philosophy*. London: Williams and Norgate.
- Tye, M. 2014. "Transparency, Qualia Realism and Representationalism." *Philosophical Studies* 170: 39–57.

CHAPTER 3

Are Perceptual Experiences
Just Representations?

HEATHER LOGUE

One task of the philosopher of perception is to give a metaphysics of perceptual experience: the sort of mental state associated with vision, hearing, touch, smell, and taste. Giving a metaphysics of perceptual experiences is, roughly, to say what they are "made of". Plausibly, perceptual experiences involve goings-on in the brain. But we may ask: are they *just* brain states? If so, which ones? Perhaps ones that represent the subject's environment in a distinctive manner? If they *aren't* just brain states, what more is there to them? Perhaps they include bits of the subject's environment, too?

Before attempting to answer such questions, it is important to be clear about what would make it the case that a given answer is correct. I take it that at least one constraint on a metaphysics of perceptual experience concerns accounting for its *phenomenal character*—"what it is like" to have a given perceptual experience (Nagel 1974), or its distinctive "feel". That is, at least part of what we're trying to do in giving such a theory is to explain where the phenomenal character of perceptual experience comes from; in other words, we're trying to give an account of the facts in virtue of which a given perceptual experience has the phenomenal character it does.¹

On what is arguably the orthodox view, perceptual experiences are states of a subject that involve *representing* her environment as being a certain way. For example, my current visual experience consists in my *visually* representing that there is something yellow and crescent-shaped before me (namely, the banana on my desk), much as I might *doxastically* represent the same proposition.² In previous work (Logue 2014), I have argued that perceptual experiences involve