

CHAPTER 19

REPRESENTATIONALISM ABOUT CONSCIOUSNESS

ADAM PAUTZ

To a first approximation, representationalism about sensory consciousness holds that sensory consciousness (for instance, being conscious of a red and round thing, or being conscious of the aroma of coffee) is a bit like *belief* or *judgment*. Having a belief is a matter of representing the world to be a certain way. Likewise, according to representationalists, sensory consciousness is a matter of representing the world to be a certain way.

Representationalism is important because it changes the shape of the mind–body problem, the problem of explaining how the brain engenders conscious experience. Once we accept representationalism, the hard problem of consciousness becomes a special case of the hard problem of representation, the problem of how the brain enables us to represent the world. Some (Armstrong, Tye, Dretske) have suggested that representationalism fits well with the idea that consciousness can be reduced to something physical. Others think that representationalism makes the mind–body problem harder because our usual models for reducing representation do not apply in the special case of conscious representation.

In Section 19.1, I formulate the basic theory. In Section 19.2, I explain an argument for it. In Section 19.3, I consider objections. In Sections 19.4 and 19.5, I describe two ways of developing the basic theory: *reductive* and *nonreductive* representationalism.

19.1 WHAT IS THE REPRESENTATIONALISM ABOUT CONSCIOUSNESS?

Here is one way to get an initial sense of representationalism. Suppose you have a hallucination as of a tiger in the room (as one does). Suppose it is more like a dream or imagery than ordinary sensory consciousness. On a traditional view, you are presented

with a tiger-like ‘mental image’ or ‘visual field region’ arrayed with ‘qualia’ on a peculiar sort of internal cinema (Russell 1912; Peacocke 2008). But this view is implausible for such a degraded hallucination. If such an ‘image’ really exists, how many stripes does it have? A better view is that you are just in a ‘belief-like’ state with an impoverished content according to which a tiger-like thing is present, even though no such (mental or physical) thing is present. And once we accept this ‘belief-like’ view for such degraded hallucinations, it then becomes natural to generalize the view to all sensory consciousness: even an ordinary veridical experience of tiger is just a belief-like state of the same kind, only with a much richer content. There is nothing more to the reality of your experience. While in a hallucination case, the belief-like state occurs ‘off-line’, in a veridical case it is controlled by the impact of the environment on the visual system. Your neural patterns somehow enable us to have belief-like states in which it seems to us that the external world is thus-and-so. But you do not have any access to these neural patterns; you only have access to what they are about.

So far, I have tried to give you an initial sense of representationalism. But we need a more precise formulation. I have said that, according to representationalism, sensory consciousness is *analogous to belief*. So, we will start with some remarks about belief; then we can formulate representationalism more exactly.

To believe is to represent things as being a certain *way*. For instance, if you believe that the thing on the table in the next room is a red ball, then you represent it as being a certain way: as *red* and *round*. In other words, in believing, you generally *predicate properties of things*. As Scott Soames says, ‘we predicate redness of *o*, when we form the belief that *o* is red’ (2010, 81). Let us say that in having beliefs we *cognitively predicate* properties of things. Here is an important fact: you can cognitively predicate a property quite independently of whether anything has that property; your belief may be false. Indeed, you can predicate redness even if there does not really exist anything you predicate it of (Soames 2010: 117).¹

Now we can turn to representationalism. Let us start with hallucination. For instance, suppose you hallucinate a blue ball. It seems to you that there is a thing before you with the properties *being bluish*, *being round*, and *being before me*. Yet these properties do not characterize any physical thing in your environment or brain. How is this possible? Above we noted that you can cognitively predicate a property even if nothing has that property. Representationalists conjecture that sensory consciousness has a similar nature. In particular, they suggest that sensory consciousness involves *phenomenally predicating properties*. (I will also sometimes call it the relation of *phenomenal representation*.) When you hallucinate a blue ball, nothing before you really has the properties *being bluish*, *being round*, and *being before me*. Still, these *properties* exist.

¹ In formulating representationalism, I will assume the existence of properties understood as ‘abstract objects’. I assume ‘realism’ about properties as against ‘nominalism’. For an argument for this assumption, see Yi 2017. However, I think that representationalism could also be formulated in a higher-order language acceptable to nominalists. I will touch on this issue later in the chapter: see the discussion of ‘relation-to-abstract-items representationalism’ and ‘concrete representationalism’ in Section 19.3.

Representationalists propose that your hallucination consists in your phenomenally predicating such properties as a result of aberrant activity in your visual system, in the absence of anything that has the properties. Your brain is ‘telling you’ that those properties are out there, even though they are not out there. Hallucination is the appearing of properties, in the absence of anything that has the properties. In particular, your hallucination consists in your phenomenally predicating the complex property $\lambda x(x$ is round and x is bluish and x is before me).² This is a first approximation: the content of the hallucination is much more rich and detailed than this.

In one respect, then, phenomenal predication is similar to cognitive predication: you can phenomenally predicate properties that nothing has. But, in another respect, phenomenal predication is *radically different* from cognitive predication: unlike cognitively predicating properties, *phenomenally* predicating properties necessarily involves undergoing a state with *presentational phenomenology* (that is why I call it *phenomenal* predication). For instance, when you phenomenally predicate $\lambda x(x$ is round and x is bluish and x is before me), it seems to you that something with those properties is ‘present’. This need not be so when one *cognitively* predicates $\lambda x(x$ is round and x is bluish and x is before me). For instance, you could guess that such a thing is before you with your eyes closed.

Representationalists hold that, in normal cases as in the hallucination case we have discussed, having the ball-experience consists in nothing but phenomenally predicating the complex property $\lambda x(x$ is round and x is bluish and x is before me). The only difference is that in the normal case, unlike in the hallucination case, something really *does* have these properties. The way you phenomenally represent things to be is the way that they are.

Representationalists generalize their view beyond vision. For instance, those who lose a limb often continue to feel a pain where the limb used to be. It is part of the phenomenology of their experience that a pain quality seems bound together with a bodily location. A representational view of pain accounts for this nicely. On this view, having phantom pain consists in phenomenally predicating the complex property $\lambda x(x$ is throbbing pain and x is at location l). Or again, representationalism can be applied to the experience of *smell* (Batty 2010). Having a smell experience consists in phenomenally predicating an olfactory quality (*minty*, *citrus-like*, etc.) together with a certain diffuse location. In a smell hallucination (phantosmia) nothing in the relevant location possesses the relevant quality.

According to representationalists, qualitative similarities and differences in sensory consciousness are similarities and differences in the properties we phenomenally predicate. For instance, if I view a ball, and you view an identical-looking but distinct ball, we might have the same experience because we phenomenally predicate the same properties, even if we see distinct balls. The character of sensory consciousness is indifferent to the identities of the things we experience (if any); what matters is the way we experience the things to be.

² I use λ -abstraction to refer to complex properties. See Sider 2010.

So far, we have developed an initial formulation of representationalism:

Representationalism (first pass): undergoing an episode of sensory consciousness with a certain character is identical with *phenomenally predicating* a complex array of perceptible properties. All differences in the character of sensory consciousness consist in nothing but differences in what arrays of perceptible properties we phenomenally predicate.³

However, there is a problem with this formulation. It contains a made-up technical term, namely ‘phenomenally predicating’. We need to say what this term means. Until we do, representationalism is not a clear thesis; we have no idea how we can determine whether it is true or false. As a matter of fact, representationalists have not really adequately explained what they mean by terms like ‘phenomenally predicate’ or ‘phenomenally represent’. So the debate over representationalism has proceeded in the absence of any adequate explanation of what it is saying. This has made the debate over ‘representationalism’ unclear and difficult to resolve (see note 13 for an example).

Here is a solution. To formulate representationalism without using any unexplained technical terms, we can employ the general *Ramsey–Carnap–Lewis method* for defining theoretical terms. Representationalists think that the postulated relation of phenomenally predicating plays *the character* role: having an experience with a certain character *consists in* phenomenally predicating a certain array of perceptible properties. It also plays *the cognitive-access* role: if a thinker phenomenally predicates a perceptible property at time, then at that time they thereby have the capacity to have beliefs according to which a thing has that property and the capacity to know what that property is like. Another key feature of the phenomenal predication relation is *neutrality*: you can phenomenally predicate a property, so that it *seems* to you that an item has the property, even if nothing in your vicinity does in fact have that property. Given all this, here is all that representationalism amounts to:

Representationalism (final formulation): There is a unique relation R such that: (I) R plays the cognitive-access role, (II) R plays the character role, and (III) R is neutral.

Representationalism is now clear. This formulation does not appeal to the undefined notion that sensory consciousness is representational. In fact, it does not even contain the word ‘representation’. But, if we like, we can now introduce the technical term ‘phenomenally predicating’ (and its synonym ‘phenomenally representing’): we can use it to denote the unique relation R (*if there is one*) that satisfies (I)–(III). On this approach,

³ For defenses of representationalism, see Armstrong 1968; Dretske 1995; Tye 1995; Byrne 2001; Chalmers 2010; Pautz 2010; Horgan 2014; Speaks 2015; and Mendelovici 2018. I have formulated representationalism in terms of the predication of *properties*. More often, representationalism is formulated in *propositional* terms. However, there is in fact no big difference between these formulations (Perkins and Bayne 2013: 74).

talk of the representational content of experience stands or falls with the truth of a substantive theory of the nature of experience.⁴

Bare-bones representationalism leaves open some big questions. How does the activity of soggy grey matter in the brain enable us to ‘reach out’ and phenomenally predicate sensible colors and shapes and other sensible properties that need not be instantiated *in* the brain? And are all these properties really out there? For instance, maybe *colors* have never been out there, but our brains evolved to ‘tell us’ they are out there, in order to help us isolate objects in scenes, remember these objects, and to communicate about them with other observers.

We will address these questions later (Sections 19.4–5). First, we must look at the argument for the bare-bones theory (Sections 19.2–3).

19.2 THE CASE FOR REPRESENTATIONALISM: AN INFERENCE TO THE BEST EXPLANATION

The argument we will consider puts forward representationalism as the best *empirical hypothesis* about the nature of sensory consciousness, akin to the hypothesis that water is H₂O or that energy is mass. Representationalism provides the simplest explanation of a host of facts: the fact that many forms of sensory consciousness are essentially spatial and ‘externally-directed’, the fact that states of visual consciousness can be indeterminate, the fact that they can depict ‘impossible’ scenarios, and much else. Here I will present the argument informally in a series of steps.

Step 1: The essentially spatial character of visual consciousness. We start with a pretheoretical datum. Consider the *ball-experience* again. This experience-type is essentially spatial. Having the ball-experience *essentially* involves the *seeming* presence of a *round* item. So spatial terms like *round* must be used in any complete definition of what it is to have the experience. Moreover, necessarily, if you have enough experiences like this, this directly enables you to acquire a concept of the shape *round*. It directly enables you to think about the property *being round*, and predicate it of things in thought. In general, so-called ‘color qualia’ (like *being bluish*) necessarily *appear to* fill spatial regions. And when you experience many color qualia at the same time, they appear to stand in certain *spatial relations* to each other.⁵

⁴ Travis (2004: esp 85 and 92) and Brewer (2017: sect 2.3) object to representationalism on the grounds that it requires—what they think cannot be supplied—a general algorithm (e.g. in terms of ‘looks’-reports or in terms of underlying facts about causal-covariation) for determining the representational content of any given experience. Against this, the availability of the present Ramsey–Carnap–Lewis formulation of representationalism shows that talk of the representational content of experience can be perfectly intelligible even in the absence of such an algorithm.

⁵ Here and in what follows, I use ‘color qualia’ and similar expressions (like ‘sensible colors’) to refer to properties involved in experience that we can get a grip on from examples, while remaining entirely neutral on their nature. For instance, as we shall see, sense datum theorists hold that they are always

This pretheoretical datum is generally accepted, which testifies to its truth. For instance, even Christopher Peacocke, who opposes representationalism, says, ‘visual experience is intrinsically [essentially] spatial... if we do not use spatial properties in characterizing the visual [experience], we omit a subjective feature of the experience’ (2008: 10).

The point applies equally to non-veridical visual experiences. Imagine that all your actual experiences are hallucinatory: you are a life-long ‘brain in the void’ (BIV). Even if your experiences are hallucinatory, they involve the seeming presence of items with shapes. And they are still enough to give you concepts of shapes. They still enable you to think about shape *properties*. For instance, you might still *have a favorite shape*. In fact, even if you are a BIV, your hallucinatory visual experiences still give you plenty of *knowledge* about shapes. They do not give you knowledge about particular physical *things* with shapes—if you are really a hallucinating BIV in empty space, you are not touch with any physical things at all. But they are still enough to enable you to know general things about shape *properties*, for instance, the timeless, necessary truth that being round is more like being oval than being square (Russell 1912: ch X; Yi 2017). And your experiences still help you to know certain other general, necessary geometrical truths.

In fact, there are real-world examples. A person with ‘Charles Bonnet Syndrome’ might hallucinate an extremely unusual and detailed shape. Before having the hallucination, she could not think of that *very fine-grained, detailed* shape property (unaided by experience, her powers of description and imagination are not so acute). It is her *hallucinatory experience* that explains her now having that fine-grained capacity for as long as the hallucination lasts. She can now wonder: ‘Is there something shaped *that way* there?’ And this shape property is not a property of any physical thing in her brain or the environment. Nor is it a property of her experience (her experience does not have a shape).

This may seem strange. How could visual experiences ‘put you in touch with’ spatial properties, even in hallucination cases where they do not ‘put you in touch’ with physical things having the properties? Maybe it is strange, but it is a fact. And it is a fact that any adequate theory of visual consciousness must somehow accommodate.

Next we will see that the spatial character of experience is not plausibly accommodated by non-representational theories. It is not accommodated at all by the ‘internal physical state theory’ of visual consciousness (step 2). It is accommodated by the ‘sense datum view’, but this view faces other problems (step 3). Representationalism accommodates the spatial character of experience while avoiding these problems (step 4).

Step 2: Against the internal physical state theory. The *internal physical state theory* holds that experience-types are (necessarily) identical with intrinsic physical-functional types. For instance, to have the ball-experience is just to undergo a distinctive, distributed system of neural patterns (e.g. McLaughlin 2012; Papineau 2016).

instantiated by sense data ‘in the mind’; by contrast, representationalists hold that they are properties *represented by* experience, and that they are either properties of ordinary physical things (Section 19.4) or nothing at all (Section 19.5). They are talking about the same salient properties, even if they disagree about where to locate them in the world.

The spatial character of visual consciousness rules this view out. The argument is an application of Leibniz's Law:

1. Visual experience-types *essentially* involve the seeming presence of items with *shapes* (*round, oval, etc.*), and *necessarily* directly enable individuals to have concepts of shapes. (Essential spatial character.)
2. These things are *not* true of mere neural pattern-types.
3. Therefore, visual experience-types are not *identical with* neural-pattern types; for instance, it is not the case that to have the ball-experience is just to undergo a neural state.⁶

The case for Premise 2 is this. Unlike visual experience-types, the essential nature of neural patterns can be fully described without mentioning shapes like *round, oval, and square*. They are not *themselves* essentially round, oval, and so on. They only involve neural properties, and visible shapes are evidently distinct from neural properties. (This is so even if visible shapes are 'Edenic'. See note 30.) It also could not be said that neural patterns essentially *represent* such shape properties (*round, oval, square*) so that they *seem* present. This view is inconsistent with the fact that the essential nature of a neural pattern can be fully described in terms of *types of neurons* and the *times, directions, and intensities* at which they fire, without mentioning such *shapes* at all. (Analogy: it cannot be said that the *linguistic expressions* 'round', 'oval', and so on *essentially* represent the relevant shapes.) Moreover, neural patterns could not, all by themselves, necessarily endow individuals with concepts of shapes. Imagine the neural patterns occurring at random in an isolated BIV. How could undergoing a series of neural patterns *alone* necessarily be enough to give BIV concepts of *shapes* and put BIV 'in touch with' *shape properties*? The isolated system could not think about shapes and know what they are like directly by virtue of undergoing these neural patterns alone. In this respect, too, visual experience-types differ from neural patterns, since visual experience-types *are* enough to directly endow individuals with concepts of shapes. Therefore, visual experiences must be *something more* than neural patterns.

True, neural states are part of the *enabling-conditions* for visual experiences. It may even be that neural states are *sufficient* for experiences. But the spatial argument shows that visual experiences are *something more* than neural states, something that essentially involves apparent *spatial properties* that need not be instantiated *in* the brain.⁷

⁶ For discussion of the spatial argument against the internal physical state theory, see Pautz 2010: 266–72; and Block 2019.

⁷ Papineau (2014, 2016) accepts the internal physical state view. He rejects Premise 1 (essential spatial character) of the spatial argument against this view (personal discussion). On his view, having visual experiences only essentially involves being in contact with *non-spatial neural properties*, so that having visual experiences alone *only* necessarily endows thinkers with concepts of such neural properties (2016, sects 10 and 15). (Papineau calls these neural properties *square**, *round**, *oval** and so on. But this is misleading, since they are *nothing like* shapes. For instance, *being round* has the definition *having edges equidistant from a common point*, but this is not true of the neural property *being round**.) Thus, according to Papineau, a life-long, accidentally-created brain in a vat that *has all the same experiences as you* can only

Step 3: Against the sense datum view. According to the *sense datum theory* (Russell 1912; Jackson 1977; Peacocke 2008), to explain the spatial character of visual experience, we must indeed say that visual experiences are ‘something more’ than neural patterns, contrary to the internal physical state theory.

For instance, suppose that you have the ball-experience during a hallucination. According to the sense datum view, this experience-type is not a mere neural-type, although it is dependent on one. In particular, on this view, it essentially involves being presented with a ‘visual field’ containing a region (a ‘visual sense datum’) that has the properties *being bluish* and *being literally round*. And this visual sense datum is not a physical item in your brain (there is no bluish and round item in the brain in such a hallucination case), even if it might be ‘created’ by your brain. It is something else.⁸

In general, sense datum theorists hold that *all* visual consciousness essentially involves standing in a relation *R* to shapes and colors. Russell (1912) called this relation *sensory acquaintance*. You always stand in this relation to shapes and colors by being presented with sense data *instantiating* those shapes and colors. We might call this the *instantiation principle*. Even in normal visual consciousness, you are presented with a mental visual field that is distinct from the physical layout of objects before you. It is like a peculiar sort of internal cinema screen. This is what determines the character of visual consciousness.

This view is seductive. If you had a vivid hallucination of a blue ball, then it would seem obvious that there is *something* present to your consciousness that is *round*—if not a physical thing, then an ‘image’ or ‘visual field region’. It would seem to be a denial of reality to suggest there is *nothing* round there (Price 1932: 3; Campbell and Cassam 2014: 10).

However, the arguments against the sense datum view are overwhelming.

First, regions of the visual field (‘visual sense data’) must be peculiar non-physical objects. As mentioned, if you have the ball-experience in a hallucination, there is a no literally round and bluish item in your brain or your environment. So if there really exists such an item, it must be a *non-physical* region in a *non-physical* ‘visual field’.⁹

think about such neural properties, and entirely lacks the ability to think about any *shapes* distinct from neural properties. But Papineau’s view seems to go against a phenomenologically manifest fact about ordinary visual experience. As Peacocke (2008: 10) says, ‘if we do not use spatial properties in characterizing the visual [experience], we omit a subjective feature of the experience.’ So even a life-long BIV could think and know about *shapes*, where shapes are evidently not neural properties.

⁸ A terminological note. As I have said, on the sense datum view, the familiar, salient bluish property you are presented with as you view the ball is in fact a property of a round ‘sense datum’. Traditional sense datum theorists (Russell 1912; Jackson 1977) called it ‘the color blue’. However, for reasons I will not go into, Peacocke (2008: 10) says it should *not* be called a ‘color’, even though it has hue and saturation. He calls it *blue**. (Peacocke, though, is happy to say that the sense datum is literally ‘round’.) I take this to be a trivial verbal issue. These different sense datum theorists have the same familiar property in mind, and just refer to it differently. In what follows, I will follow the traditional sense datum theorists, and simply call it ‘the color blue’.

⁹ Peacocke (2008: 14) tries to avoid this problem. He writes that, for such a hallucinating individual, ‘it is *as if* there is something [a “visual field”] parts of which enjoy the relevant sensational properties [e.g. being round, being bluish], even though there is no such thing.’ Thus, Peacocke says that, in this case, there does not *really* exist a round and bluish ‘visual field region’; it only *seems* that such a thing exists.

Second, if you look a waterfall, and then a stationary rock, you have an experience as of the rock staying in place, but at the same time you have an experience as of movement upwards. How might sense datum theorists like Russell and Peacocke explain this ‘impossible’ experience? On their view, your visual field contains a blackish and rock-shaped region that is distinct from the rock. Is this region stationary? If so, what accounts for the experience as of movement upwards?

Third, the sense datum view does not apply to ‘indeterminate’ visual experiences, for instance, experiences of things in the periphery, dream-like hallucinations, and imagery. In such cases, it can seem to you that something is (say) roughly triangular, even if there is no *specific* triangular shape that it seems to you to have. Your experience just does not go into that much detail. Would proponents of this view like Russell and Peacocke say that a ‘region of the visual field’ can be roughly triangular, without having any particular triangular shape? That seems incoherent. But if we do not apply the theory to such degraded cases, then considerations of uniformity suggest that we should not apply it to more ‘vivid’ experiences either.

For these reasons, while it may *seem* that there exist ‘sense data’ or ‘visual field regions’, we should deny the real existence of such things.

Step 4: Representationalism to the rescue. The conclusion of the spatial argument was that visual experiences are *something* more than neural states, something that essentially involves *shape properties* that need not be instantiated in the brain. We have seen that the ‘something more’ cannot be the presentation of ‘sense data’ or ‘visual field regions’ *possessing* spatial properties. That idea faces big problems. But then what is the alternative?

Enter representationalism about visual consciousness. On this view, the ‘something more’ is a matter of *phenomenally representing* spatial properties, sensible colors, and so on. Such representational states are *enabled by* our neural states but are *distinct from* those neural states. On representationalism, in hallucination cases, it *seems* that there exist ‘sense data’ or ‘visual field regions’ replete with qualia, but there do not *really* exist such things. In such cases, there is misrepresentation. So representationalism avoids the real existence of ‘sense data’ and ‘visual field regions’.¹⁰

Since he says it does not exist, Peacocke avoids commitment to a ‘non-physical object’ in this case. But this is just a straight representationalist account of the case. And so the question arises: why does Peacocke not apply this account to *all* cases, and so accept a version of the kind of representationalism we are leading up to? In that way, he could avoid positing ‘visual field regions’ *in normal as well as abnormal cases*. (For a similar point in a different connection, see the reply to the first objection in Section 19.3, and also note 13.)

¹⁰ Experiences are distinct from neural states even on ‘internalist representationalism’. For instance, Ned Block (2019) is an internalist about experience; in particular, he thinks that our experiences are totally fixed by our neural states (so that even a life-long BIV could have all the same experiences as you). To explain the essentially spatial character of visual experience, Block also accepts a form of representationalism: he suggests that visual experiences essentially involve standing in an irreducible phenomenal predication (representation) relation to arrays of spatial properties (for discussion see Section 19.5). (This is why, for a BIV, it *seems* that there exists a visual field whose parts have various spatial properties and qualities, even though in reality there exists no such field.) The result is a kind of internalist representationalism. Even though this view holds that visual experiences are wholly determined by neural states, it implies that they are something more than neural states. This is because it implies that having experience

Taken together, the points we have covered lead straight to this view. The spatial character of visual consciousness means that, in both good cases and hallucination cases, the ball-experience involves the seeming presence of a round thing, and necessarily grounds the ability to think about the property of being round. The simplest explanation is that, in every case, having this experience involves standing in a relation R to the property *being round*; when you bear this relation to a property, then you can easily thereby predicate this property of things in thought. That is, R plays the *cognitive-access role*. As we saw, sense datum theorists also posited such a relation, and they thought that it obeys a kind of *instantiation principle*. Russell (1912) called it *acquaintance*. But we also saw that the instantiation principle leads to disaster—it requires peculiar non-physical objects ('sense data', 'visual field regions'). So now we add to our proto-theory a clause that goes against the sense datum theory: you can stand in R to a property even if you are not related to an item (a physical thing, an image, or visual field region) that instantiates the property. That is, R satisfies *neutrality*. When you have the ball-experience, you stand in the *same* R relation to the property *being bluish* (or *being bluish**, or whatever you want to call it). After all, we might as well explain color experience and spatial experience in the same way. So we have arrived at the view that having the ball-experience *necessarily* involves standing in a relation R to the complex property $\lambda x(x$ is round and x is bluish and x is before me), both in hallucination cases and in normal cases. Now, the simplest explanation of this necessary connection is that having the ball-experience *just is* standing in R to this complex property. There is *nothing more to having the ball-experience*. In particular, standing in R to $\lambda x(x$ is round and x is bluish and x is before me) is not grounded in any additional, more basic mental condition ('intrinsic properties of the experience', 'undergoing a sensation', 'being presented with a visual field region with qualia', etc.). That is, R plays the character-role.

We have given an argument for the theory that the ball-experience involves a relation R that has certain features: it is 'neutral', grounds cognitive access, and constitutes phenomenal character. As we have formulated 'representationalism', this hypothesized relation thereby counts as 'phenomenal representation', and this thesis counts as 'representationalism'. The claim that experience is representational is *not* a pretheoretical claim that lies at the surface (*pace* Speaks 2017: 493). It cannot be established just by pointing out that differences in experience are accompanied by differences in how things seem (*pace* Byrne 2001). (After all, even sense datum theorists could agree with this.) It is a totally non-obvious empirical hypothesis about the hidden nature of experience, akin to the hypothesis that matter is energy.

Representationalism accommodates the spatial character of visual experience. Contrary to the internal physical state view, the definition of what it is to have the ball-experience involves more than *types of neurons* and the *times, directions, and intensities*

essentially involves standing in a non-neural, phenomenal predication relation to spatial properties that are not instantiated in the brain. (Compare how the sense datum view implies that experiences are something more than neural states even if they wholly determined by neural states.) We will discuss internalist representationalism in Section 19.5.

at which they fire. Rather, we must use spatial terms to fully characterize what it is to have the ball-experience. For to have the ball-experience is to phenomenally predicate $\lambda x(x$ is round and x is bluish and x is before me). At the same time, since phenomenal predication is a ‘neutral’ relation to properties, representationalism avoids ‘sense data’ or ‘visual field regions’. There need not exist anything (a sense datum or a visual field region) that has those spatial and other features. So representationalism avoids the problems confronted by the sense datum view. For instance, representationalism allows a very natural explanation of ‘indeterminate’ experiences (peripheral vision, dream-like hallucinations, dreams, imagery): they involve information loss. Just like beliefs, experiences can be more or less rich in informational content. No ‘indeterminate sense data’ or ‘indeterminate visual field regions’ are required. As for ‘impossible’ experiences like the waterfall illusion, the explanation of the peculiar phenomenology is now straightforward: you phenomenally represent that the rock is standing still and, at the same time, you phenomenally represent that *something in the same region (not necessarily the rock) is moving upward* (without phenomenally representing anything else about this moving item). The movement is not ‘bound’ to the rock.¹¹

Step 5: Generalizing to all forms of sensory consciousness. Representationalism is a very strong thesis. It says that *every* type of sensory-perceptual experience is identical with standing in a relation R to a distinct complex of perceptible properties. To support this, we note that the previous steps apply equally to all other sensory-perceptual experiences. Different types of experiences provide cognitive access to different arrays of sensible properties (pain qualities, olfactory qualities, audible qualities) bound with various spatial properties. The most economical and most uniform hypothesis is that different types of visual experiences consist in nothing but *bearing R —the same relation involved in the ball-experience—to different complexes of properties.*

Step 6: Against naïve realism. So, a strong argument can be made for representationalism. Nevertheless, many think we should instead accept *naïve realism* (e.g. Brewer 2011; Campbell and Cassam 2014; Allen 2016). Naïve realists agree with representationalists that what it is to have the ball-experience, unlike what it is to have a neural state, can only be defined by using spatial terms like *round* and *three-feet away*. So they agree with representationalists that having the ball-experience is something more than undergoing a neural state. But they disagree with representationalists about what this ‘something more’ is. On one version of naïve realism, to have the ball-experience is (roughly) to *either* see the blueness and roundness as an external item *or* be in a state that is indistinguishable from seeing the blueness and roundness of an external item. This is a kind of ‘disjunctivism’. The first clause provides a naïve realist account of normal perception and the second clause provides a ‘negative epistemic’ account of illusion and hallucination.

How are we to decide between representationalism and naïve realism of this sort? This is a big issue but let me mention two points.

¹¹ Although I call this an ‘impossible’ experience, it is not really the case that its content is impossible. For instance, it would be true in a case where the rock is not moving, but an invisible thing in the vicinity is moving upward.

First, naïve realists have problems with illusion and hallucination. Return to the example (mentioned in step 1 above) in which someone with Charles Bonnet Syndrome has a vivid and long-lasting hallucination of a very idiosyncratic and complex shape *S* and that this explains her new capacity to (for example) wonder whether anything really has *that specific shape*. According to the ‘negative epistemic’ account of hallucination that many naïve realists favor, all that is going on is that she is in a state that cannot be discriminated from seeing. But how could this negative epistemic condition *explain* her new capacity to think uniquely of this idiosyncratic and complex shape *S*? (She could not think of it prior to having such an experience: her experience-independent powers of thought are not so discriminating.) If her hallucination explains this new cognitive capacity, does it not have to have a more ‘positive’ nature, as we said in step 4 of the foregoing argument for representationalism? Do we not have to say that her having the hallucination involves her standing in *some interesting perceptual relation R* to this specific idiosyncratic shape *S*, a relation that is fit to play the ‘cognitive access role’ (for example, *perceiving a sense datum with the shape*, or *phenomenally predicating the shape*)? Otherwise, how does the hallucination *explain* her capacity to think of *S* and not some other shape? (For discussion of these issues, see Brewer 2011: 112ff and Alford-Duguid and Arsenaul 2017.)

In addition, naïve realists face empirical problems. They cannot accommodate in a plausible way the role of internal neural processing in shaping sensory consciousness. (For discussion, see Logue 2017; Beck 2018; and Campbell 2018.) By contrast, representationalists can easily accommodate the role of the brain, by accepting an ‘internalist’ form of representationalism.

We will be discussing the forms of representationalism, and how they might accommodate the role of the brain, in Sections 19.4–5. But first let us consider some objections to the basic theory of representationalism.

19.3 OBJECTIONS TO REPRESENTATIONALISM ABOUT SENSORY CONSCIOUSNESS

I will put each objection in the mouth of a hypothetical opponent, and then offer a possible representationalist rejoinder.

Objection. How can you be so confident in representationalism’s strong thesis that *all* phenomenal differences among experiences are representational differences?

For instance, what about blur? You look at a blue ball with your glasses on. Then you take your glasses off. There is a phenomenal difference in the character of your experience. Yet to a mature perceiver it does not look as if any mind-independent object has really changed. So (the objection continues) there is no difference in the properties you phenomenally predicate (Boghossian and Velleman 1989; Burge 2003; Smith 2008).

Reply. There are two plausible representationalist explanations of blur: the under-representation view (Tye 2000) and the over-representation view (Allen 2003). However, here I will not go into them. Instead, I will make a more general point. For the sake of argument, suppose that the representationalist explanations of blur due to Tye and Allen fail. Then opponents will demand an alternative account of blur from representationalists. But we can turn the tables and ask them *what is their own account?* *Whatever* account they offer, it might be co-opted by the representationalist and turned into a representationalist account.

For instance, anti-representationalists might suggest a sense datum theory of blur. They might say that, when you take off your glasses, what is going on is that you are presented with a ‘visual field region’ that has the following trio of properties: *round*, *bluish* (or *bluish**), and *blurry*. Here the property of being blurry is a property of regions of the visual field and not physical things. This anti-representationalist account, then, appeals to the property of being blurry without saying much about its nature.¹²

Now if anti-representationalists can appeal to such a special property (without saying much about its nature), there is no obvious bar to representationalists appealing to such a property too. While anti-representationalists can say it is a property of a non-physical ‘visual field region’, representationalists can say it is a property of *nothing* at all. That is, when one has a blurry experience of a blue ball, all that is going on is that one is phenomenally representing the co-instantiation of the properties *being bluish* (or *being bluish**), *being round*, and the special property *being blurred*. But nothing—no physical thing and no region of the visual field—has this trio of properties.

This representationalist account describes the content of the blurry experience *in exactly the same way* as the anti-representationalist account. It just adds that the content is not real. It merely *seems* that there is a blurry visual field region, according to this view. So how could it be *inferior* to the anti-representationalist account? If anything, it is vastly superior, since it avoids the real existence of a peculiar, non-physical ‘visual field region’.

The point here is general. If you have reason to believe that there is a difference in character between two of your experiences, then you must be aware of a *difference*, or an *apparent difference*, of some kind. (If not, you have no reason to think that there was a difference in the character of your experiences in the first place!) But a difference is just a difference in *properties*. So it will always be possible to hold that the relevant difference in character consists in a difference in what properties you bear the representational relation *R* to. That is, it will always be possible to say that the different properties are *represented* properties. For the Ramsey–Carnap–Lewis version of representationalism

¹² Boghossian and Velleman (1989: 96) suggest a sense datum view of this kind (but in his 2008 Boghossian takes it back and suggests a representational view). Burge (2003) criticizes representationalism on the grounds that it cannot handle blur but he does not offer his own positive, illuminating account of blur. In fact, he nowhere proposes any theory of the phenomenal character of visual experience and its built-in spatial structure (although at p. 444 he does suggest, without argument, that it supervenes on the internal chemical properties of the brain). However, he does repeatedly speak of a ‘visual field’ (2003: 408, 440), which suggests something like the sense datum theory.

is a very minimal thesis which does not by itself place any restrictions on what properties can be phenomenally represented.¹³ (True, it may sometimes be difficult to *specify* or *characterize* the different properties; but this does not affect my present point.) So it is hard to see how there *could* be a counterexample to the bare-bones representationalist theory. If there is a good objection to representationalism, it must take a different form.¹⁴

Objection. You representationalists give a lousy *explanation* of the ball-experience. Your explanation is that you stand in this mysterious relation, ‘phenomenal predication’, to $\lambda x(x$ is round and x bluish and x before me). When asked to define this term, you say it is a theoretical term, and define it as ‘the relation such that having an experience consists in standing this relation to some property-complex’. But this makes the explanation look trivial or circular. How is this progress? (John Campbell pressed this objection in discussion. See also Langsam 2018.)

Reply. This objection misunderstands representationalism. Representationalism, as we have formulated it, is not an *explanation* of experience. To explain X is to cite *something else*, Y, that stands in an explanatory relation (causation, grounding) to X. But representationalism does not do that. Representationalism is an existentially-quantified *identity claim*. It only says this: there is a relation R with certain features (it is ‘neutral’ and grounds ‘cognitive access’) and having the ball-experience (for instance) is identical with standing in R to $\lambda x(x$ is round and x bluish and x before me). This cannot be criticized for giving a trivial or circular *explanation* of experience—it is not trivial, it is not circular, and it does not even pretend to be an explanation of experience. It is just a remark about the structure of experience.

Objection. Ok, but we can still criticize the *identity claim* proposed by representationalism. For starters, it fails to account for the sense in which the properties that characterize what it is like to have the ball-experience are *present* when one has this experience. This is a way in which having the ball-experience is different from merely *believing* that a bluish and round thing is there. (See Papineau 2016 and Campbell 2018.)

Reply. What do you mean by ‘present’? *Instantiated*? In that case, the claim here is not true of all experiences. Suppose you have the ball-experience in hallucination. As we

¹³ For example, against representationalism, Peacocke (2008: 9–10) suggests that *after-image experiences* do not involve phenomenally representing *anything at all*, on the grounds that in such cases ‘it does not look [to us] as if there are [mind-independent] objects or events in your spatio-temporal environment’. This relies on the assumption that you phenomenally predicate a property in visual experience just in case it looks to you as if some *mind-independent* object or event has the property. But representationalism is not committed to this assumption. In fact, it does not require that the content of experience is ‘looks-indexed’ (Travis 2004) in any such simple way (see note 4 of this chapter). So representationalists are free to characterize the content however they like. In fact, they could *agree* with sense datum theorists like Peacocke that the content is such that it looks to mature perceivers as if after-images are *mind-dependent visibilia*, but they will just add that the content is unreal (does not really obtain), thereby avoiding Peacocke’s reification of ‘visual field regions’. That is, they can also talk of visual field regions, but place all such talk within an intensional context, thereby avoiding ontological commitment (see note 9).

¹⁴ This representationalist move (‘just stick the different properties in the representational content’) may always be available, but it may not always be plausible. For instance, Speaks (2015: 191) notes the availability of this move in the case of *covert attention shifts*, but suggests that in this case it may be more plausible to revise or supplement representationalism.

saw above (‘step 1’), the spatial feature *round* must be mentioned in any complete characterization of what it is like. But here nothing is actually round—at least if we reject sense data (‘visual field regions’). So there is a strong case against the ‘presence’ claim if it requires *actual instantiation*.¹⁵

At this point, the objector might say, ‘well, what I meant is that, necessarily, in every case of having the blue-experience, it vividly (“phenomenally”) *seems* that an instance of the relevant properties is “present”. That is different from merely believing.’

But, of course, this is something representationalists can and do accept; in fact, truth be told, bare-bones representationalism *does not go far beyond this!* It just adds that this can happen even if nothing with the relevant properties *is* really present. In that minimal sense, the ‘seeming’ state is a ‘representational’ state. And (as a ‘common factor theory’) representationalism holds having the experience, in both veridical and hallucination, is nothing more than being in this representational state.

Objection. I am still not appeased. When you formulated representationalism, you said more: you brought in ‘abstract items’. You said that having the ball-experience is nothing more than standing in a relation to a *non-spatial, abstract item*, the complex property $\lambda x(x \text{ is round and } x \text{ is bluish and } x \text{ is before me})$. How weird is that? I just think that this is obviously false—like the claim that Julius Caesar is identical with the number three. I have got an ‘intuition of distinctness’. To be more specific, I object to the idea that experiences are identical with relations to *abstract items*. (See Pautz 2010: 292–3; Papineau 2014.)

Reply. This objection does count against representationalism, but several points soften the blow.

- (i) Consider an analogy. The right account of *exact resemblance* mentions abstract items: for two things to *resemble exactly* is for them to instantiate the same *properties*. This is somewhat surprising. Maybe, then, the right account of *experience* mentions abstract items, even if this is surprising.
- (ii) Even though I formulated representationalism as the view that experiences are relations to abstract items, there is another way of thinking about the view. On this alternative formulation, to have the ball-experience (for instance) is to *phenomenally represent that a bluish and round item is present*, where this representational state is *not* understood to be a relation to any kind of abstract item.

¹⁵ The thesis that the spatial feature *round* must be mentioned in any complete characterization of what it is to have the ball-experience is generally accepted. It is accepted by sense datum theorists, representationalists and naïve realists. This testifies to its truth. True, internal physicalist state theorists, such as David Papineau, will reject this claim; instead, they will say that what it is to have the ball-experience can be specified entirely in terms of *types of neurons* and the *times, directions* and *intensities* at which they fire, without mentioning spatial properties at all. But, as we saw in step two of the argument for representationalism, if we have to choose between the essential spatiality of the ball-experience and the internal physical state view, we should keep essential spatiality and reject the internal state view. So our counterexample to the ‘instantiation’ claim stands: if someone were to have the ball-experience in a hallucination, *round* would essentially enter into a characterization of what the hallucination is like, yet nothing in the vicinity would actually be round.

(Notice that the specification in italics does not contain a singular term referring to an abstract property. True, it contains a ‘that’-clause, but the view I have in mind denies that this is a singular term referring to a proposition understood as an abstract item.) This kind of ‘concrete representationalism’ avoids the specific objection that ‘relations to entities in a Platonic realm outside space and time cannot matter to consciousness’, as Papineau (2014: 8) puts it. In fact, concrete representationalists absolutely agree with Papineau’s remark here. Indeed, they might be nominalists who altogether reject the existence of abstract entities in Plato’s realm.¹⁶

- (iii) At this point, the objector might rephrase the objection. It is not just that it cannot be that to have the ball-experience is to stand in a relation to an abstract item—an objection that is avoided by concrete representationalism. It is more general than that: it is a general ‘intuition of distinctness’ to the effect that having the ball-experience must be distinct from representing that a blue and round thing is present (where that can happen even if nothing blue and round is present). This intuition, it might be said, counts equally against ‘concrete’ representationalism and ‘relation-to-abstract-items’ representationalism. In reply: it would be odd if opponents of representationalism relied on this kind of basic ‘intuition of distinctness.’ For equally forceful intuitions of distinctness tell against their own views. For instance, such an intuition counts against the internal physical state view: how could this technicolor phenomenology *just be* a pattern of neural activity in soggy grey matter? And such an intuition counts against naïve realists’ ‘negative epistemic theory’ of phenomenal character in hallucination cases (which we briefly discussed in step 6 of the argument for representationalism). How could a mere negative epistemic property constitute technicolor phenomenology?¹⁷

¹⁶ I said in the text that concrete representationalists might be nominalists who altogether reject the existence of abstract items. It is worth remarking that they could also accept the existence of abstract items. Analogy: it is natural to think that the fact that an apple is red is a wholly concrete fact, involving only the apple. It is not identical with the fact that it instantiates an abstract item, the property of being red (Lewis 1986: 190, fn. 13). But you could accept this and then go on to say that the fact the apple is red *grounds* the fact that it stands in the *instantiation relation* to an abstract item in Plato’s realm, the property of being red. Likewise, concrete representationalists hold that the fact that you have the blue experience is identical with the wholly concrete fact that you phenomenally represent that a blue thing is present. But they might then go on to say that this *grounds* your standing in a derivative phenomenal representation relation to the property of being blue, an abstract item. They would still evade the specific objection ‘it cannot be that the fact that you have the ball-experience is *identical with* the fact that you stand in a relation to an abstract item.’

¹⁷ Only an across-the-board act-object view—such as the sense datum view or ‘Austinian disjunctivism’ (Moran 2018, forthcoming)—does not face the ‘intuition of distinctness’ objection. For, on this view, having the ball-experience is identical with experiencing the blueness and roundness of some presented object. And, if anything, intuition *favors* this view (Price 1932: 4). But this theory faces the problems mentioned in step 3 of the inference-to-the-best-explanation argument for representationalism.

Objection. There are metaphysically necessary restrictions on how things can appear—‘laws of appearance’. For instance, it is metaphysically impossible that a surface should seem both *round* and *square*, or *pure red* and *pure yellow* at the same place and the same time. (Of course, something can look *reddish* and *yellowish*—that is, *orange*—but that is not a counterexample.) And it is metaphysically impossible that something should seem red but not extended. It is metaphysically impossible that something should appear either red or green but nothing more specific. How might representationalists explain all this? It is not clear. Why could your (or some possible individual’s) visual system not get really screwed up and phenomenally predicate incompatible properties of something? Some philosophers—‘Meinongians’—*believe* that there are round squares, and pure red and green things; so if experience is representational like belief, why can you not have *experiences* with these weird contents? By contrast, sense datum theorists have an explanation ready to hand, since they endorse the ‘instantiation principle’. Their explanation is that something’s perceptually seeming to have incompatible properties would require that there actually be something—a ‘sense datum’ or ‘visual field region’—having incompatible properties. But nothing—not even a sense datum or ‘visual field region’—could have incompatible properties. So it looks like we have stumbled upon a new argument for the sense datum view. Maybe we should accept the sense datum view after all, in spite of its problems, because it explains the laws of appearance.¹⁸

Reply. This is indeed a puzzle for representationalists. One option for them is to reject the metaphysical necessity of the laws of appearance, if they can find no explanation of their metaphysical necessity.¹⁹ Alternatively, they might just accept the laws of appearance as basic, inexplicable metaphysical necessities.

One point worth noting here is that this is a problem for other views too—including the sense datum view. The laws of appearance are many and various, and *no* view can explain them all. Even sense datum theorists cannot explain certain ‘laws of appearance’.

¹⁸ For the problem of explaining the laws of appearance, see Pautz (2017), Speaks (2017), Morgan (ms). It is worth mentioning that the hypothesis that experience is ‘non-propositional’ and ‘feature-placing’ (Block ms) is not enough to explain the laws of appearance. For, whatever it may mean, it faces the very same explanatory problem. Why should experience not ‘place’ incompatible features (*pure red* and *pure green*) in one place, or super-disjunctive properties (*pure red or pure green*), or merely negative properties (*not pure red*)? Block (ms: ch 3) rejects some ostensible laws of appearance, and then suggests that the remaining laws of appearance can be explained by the thesis that phenomenal representation is realized in an *iconic format*. For difficulties with this suggestion see Pautz 2017.

¹⁹ For example, Speaks (2017: 495) and Block (ms: ch 3) favor the option of *rejecting* the metaphysical necessity of some of the laws of appearance: for instance, it is metaphysically *possible* that a surface look pure red and pure yellow all over, or round and square. (As mentioned in the previous note, Block accepts some few *other* laws of appearance but suggests that they are explained by the thesis that phenomenal representation is realized in an *iconic format*.) But in that case why do we *think* it is so obviously metaphysically impossible that a surface look pure red and pure yellow, or round and square? Block (ms) and also E. J. Green and Jacob Beck (in discussion) suggested to me that the explanation might be that, given the way the visual system works, this never happens in humans (though it could happen in other possible creatures), and so we have difficulty imagining the relevant experiences. But this may be insufficient, for there are many experiences that no one has ever had and that we have difficulty imagining (e.g. experiences of unusual shapes) but that we take to be metaphysically *possible*.

Why must everything that looks colored also look extended in space? Why could you not be perceptually acquainted with the color of a sense datum ‘neat’, without being acquainted with its extent in space or its location? That does not seem to be a possible visual experience, but why not? Furthermore, a sense datum has loads of properties: for instance, the relational property *being created by so-and-so brain process*, various logically complex properties (e.g. disjunctive properties), and so on and so forth. Why can you not be perceptually acquainted simply with *these* properties of the sense datum? Why are some properties (color and shape) ‘acquaintables’, while these other properties are not? Since all views face a version of the puzzle of the laws of appearance, the puzzle does not provide a strong reason to accept the sense datum view or any other view over representationalism. It remains the case that representationalism achieves an overall better ‘balance sheet’ than the alternatives.

19.4 REDUCTIVE-EXTERNALIST REPRESENTATIONALISM

We have looked at a defense of representationalism. However, representationalism is incomplete until it is combined with an answer to the question of how the brain enables us to phenomenally predicate sensible colors and shapes and other sensible properties that need not be instantiated *in* the brain.

In our final two sections, we will look at two different answers to this question. We begin in the present section with *reductive-externalist representationalism*, which has been defended by Armstrong (1968), Dretske (1995), Tye (1995), and Byrne and Hilbert (2003).

Consider the world before the evolution of conscious creatures. Reductive-externalist representationalism maintains that even at this time external objects and events were rich with sensible properties (‘qualia’). For instance, tomatoes were red, the sky was blue, a falling tree made a sound (even though no one was around to hear it), methane had a bad objective smell, and so on. The sensible beauties (and the sensible nasties) are not creations of the brain; they were out there even before brains came on the scene. The so-called ‘qualia’ are not features of ‘sense data’ or ‘sensory field regions’ in the head; they are just features of physical things in the world. They are just as objective as shape and size. In fact, this view holds that sensible properties are *reducible to* physical properties that things had even before conscious creatures came on the scene:

The Reduction of Sensible Properties. Colors are reflectance properties (ways of reflecting light), odor qualities are molecular properties (see Figure 19.1A), audible qualities are complex physical properties, and so on.

Compare: water is H₂O, which was on the scene long before sentient creatures showed up.

So, before the evolution of consciousness, all these sensible properties were out there, but no one was around to appreciate them. Then we and other creatures evolved the capacity to be *conscious* of them. How did that happen? Reductive-externalist representationalists hold that we came to be conscious of them by *phenomenally representing* them. So the notorious ‘hard problem of consciousness’ becomes the hard problem of representation: how do we manage to represent these properties so that they seem *present* to us?

Here reductive-externalist representationalists appeal to a standard *externalist model* for reducing representation of the kind that many philosophers developed in the 1990s. Analogy: thermometers came to represent temperatures by undergoing states that have the function to indicate those temperatures. Likewise, we came to phenomenally represent, and thereby be conscious of, the objective sensible properties of things by undergoing states that have the (biological) function of indicating them. Roughly, you stand in the indication relation to property just in case you are in a brain state that *normally* occurs only when that property is present before you (Dretske 1995: 48).

This origin story for sensory consciousness is still incomplete. The phenomenal representation relation cannot be a mere biological indication relation. Take an early single-celled organism. Suppose it had states that indicate *light* and *dark*. We do not think it was *conscious* of these features. It may ‘represent’ them in some sense but it does not *phenomenally* represent them. So reductive-externalist representationalists need to address the question of what turned mere *indication* into the magic of conscious presentation.

In response, reductive-externalist representationalists hold that phenomenal representation is indication *plus cognitive-rational accessibility*:

The Reduction of Phenomenal Representation Relation. The dyadic mental relation *x phenomenally (consciously) represents property y* is identical with the following complex indication relation: *x is in a internal physical state that has the biological function of indicating property y and that internal state is poised to be fed into a ‘cognitive system’ for the rational control of thought and action.* (Tye 1995; Dretske 2006)

On this view, consciousness was not present in single-celled organisms; it came into existence only when creatures became capable of *thought* and *reason* (Tye 2000: ch 8). Since this account explains consciousness in terms of thought and reason, it needs to be supplemented by an account of thought and reason that is independent of consciousness, in order to avoid circularity. But we will leave this difficult issue to one side.

In a nutshell, then, that is reductive-externalist representationalism. It implies a radical externalism about the qualities of experience. As Michael Tye puts it:

Peer as long as you like at the detailed functioning of the brain... that is not where phenomenal character is to be found. Neuroscientists are looking in the wrong place (Tye 199:, 162–3)... phenomenal character is in the world (Tye 2009: 119).

For instance, suppose that you look at the ball. On reductive-externalist representationalism, the bluish *quale* that you are conscious of *just is its reflectance-type* (its distinctive way of reflecting light), and you are conscious of it by undergoing a neural state that indicates it. Any possible creature that indicates the same reflectance-type must experience the same bluish quale, *no matter what its internal neural processing is like*.²⁰

Reductive-externalist representationalists provide a parallel account of other experiences. For instance, suppose you smell a cloud of R-limonene molecules (Figure 19.1A). It will smell *citrus-like* to you. On reductive-externalist representationalism, the citrus-like *quale* that you are conscious of *just is this molecular-type*, and you are conscious of it ('phenomenally represent' it) by undergoing a neural state that normally indicates it. This implies that *any possible creature that undergoes a state indicating the same molecular-type must experience the same citrus-like quale, no matter what its internal neural processing is like*.

Likewise, reductive-externalist representationalists hold that, if you have a pain in your leg, your pain-system detects, and thereby phenomenally represents, some disturbance-type down there. The felt quality is identical with the disturbance-type, and its sensory intensity is constituted by the extent and size of the disturbance. Pain is not in the brain.

A consequence is that a life-long *brain in a void* could not be conscious of any olfactory qualia, or color qualia, or even pain qualia, because qualia are in the world or the body, and such a brain would not be connected to them in the right way (its internal states do not have a history of indicating them). The brain needs *help from the world* to generate the consciousness of qualia; it cannot do it all on its own.

It may be wondered how reductive-externalist representationalism explains *perceptual variation*. Reductive-externalist representationalism provides a *selectionist* explanation. The pre-conscious world was rich with sensible properties. Different creatures have different experiences because their visual systems indicate, and thereby enable them to phenomenally represent, different external properties. For example, dogs' olfactory systems indicate different ranges of molecular-types (and hence, different smell qualia) than our own. Or again, pigeon visual systems have the function of indicating reflectances involving UV light, which (on this view) constitute alien color *qualia* that we cannot imagine. We are all conscious of the real world—only we are conscious of different aspects of that world (Cohen 2009: 78ff).

Reductive-externalist representationalism also explains illusion and hallucination. For instance, suppose you have the ball-experience in a hallucination. On this view, the vivid impression that there is *right there* before you a bluish and round thing essentially consists in the fact that you are undergoing an internal neural pattern that *normally* (but not in the present circumstances) indicates an object with a blue-reflectance and a round

²⁰ Because of 'metamerism', reductive-externalist representationalists must identify sensible colors with reflectance-types that are *highly disjunctive* (see Byrne and Hilbert 2003). A similar point applies to *smell qualities* (to be discussed below). However, I will ignore this complication, since it is irrelevant to the points to be discussed.

shape. The ball-experience is a wide physical state that can only be fully characterized by mentioning features of external things like *round* and *bluish*. This explains the essentially spatial character of the hallucination without positing a bluish and round ‘visual field region’ or ‘sense datum’.

Reductive-externalist representationalism is attractive. But there is reason to think that it just does not agree with the facts. Its basic idea—as expressed by Tye in the above quotation—is that the explanation of phenomenal character is to be found in the external world, not the brain. But decades of studies in psychophysics and neuroscience suggest exactly the opposite.

This point is especially clear in the case of *smell*. To illustrate, suppose that you consecutively smell *citral*, *R-limonene*, and *R-carvone*, as shown in Figure 19.1A. You will then experience the smell qualities *citrus*₁, *citrus*₂, and *minty* (where *citrus*₁ and *citrus*₂ are two similar but distinct citrus smell qualities). Howard et al. (2009) used fMRI to look at the neural patterns (distributed spatial-temporal patterns of neural firing) caused by these odorants in a normal human such as yourself. Recall that reductive-internalist representationalism holds that the smell qualities (*citrus*₁, *citrus*₂, and *minty*) that you experience are *identical with* the corresponding molecular-types in the air. It is the *molecular-types indicated*—and *not* your distributed neural patterns—that constitutes what olfactory qualities you experience. But what Howard et al. (2009) found seems to show that the *opposite* is true. There is a *big mismatch* between the resemblances among the smell qualities you experience and the resemblances among the corresponding molecular-types in the external world, as Figure 19.1A illustrates. In general, olfactory scientists have found that there is *no* measure of molecular similarity that predicts qualitative similarity (Coward and Rawson 2001: 568).²¹ On the other hand, Figure 19.1A also illustrates the finding of Howard et al. (2009) that there is, by contrast, a *perfect agreement* between the resemblances among the smell qualities you experience and the resemblances among the distributed neural patterns in your olfactory system. Many other studies have supported the same basic finding: while *molecular* similarity-space totally fails to match qualitative similarity space, *neural* similarity-space nicely matches qualitative similarity space (e. g. Youngentob et al. 2006). This seems to directly contradict the externalist position of reductive-externalist representationalists.

We can make vivid the conflict between these empirical findings and reductive-externalist representationalism by considering a hypothetical scenario (illustrated in Figure 19.1B). In particular, consider a counterfactual scenario where everything is the same but for one thing: humans naturally evolved so that, in this scenario, your neural

²¹ Some have proposed that a molecule’s *vibrational frequency* in the infrared range predicts its quality. Against this, enantiomers (mirror-image molecules) can smell quite different but have the same vibrational frequency. For other problems, see Pautz 2014b: 276, fn. 11. In general, ‘numerous chemical and molecular features (e.g., molecular weight, molecular mass and shape, polarity, resonance structure, types of bonds and sidegroups) can all influence the odorous characteristics of a chemical [but] no systematic description of how these characteristics relate to particular odor qualities has been developed’ (Coward and Rawson 2001: 568). It is only in the brain that we find good predictors of smell quality. (Here I am indebted to Alex Byrne.)

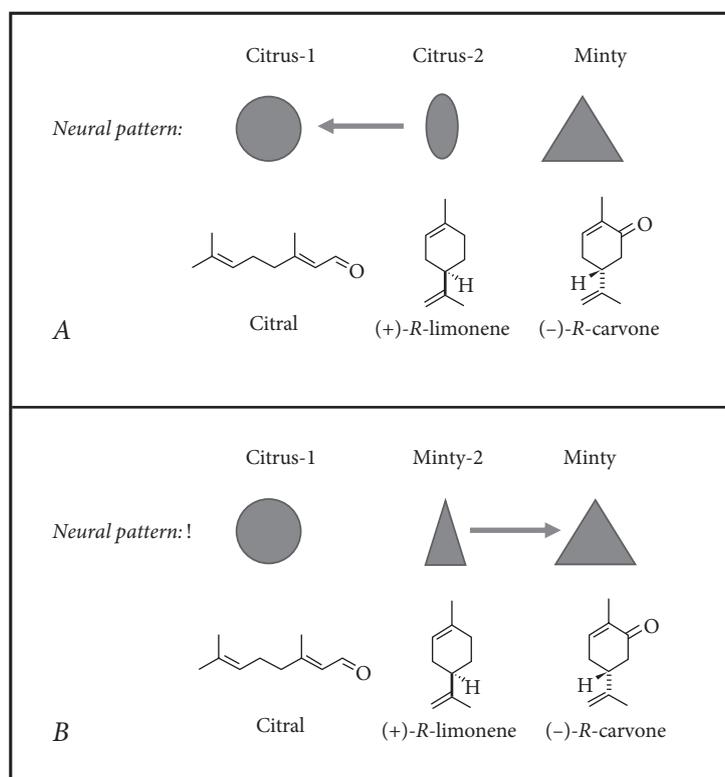


FIGURE 19.1 A The situation in the actual world: smell experiences are better correlated with internal neural patterns than with external molecular-types. B A counterfactual situation in which you have a different neural representation of R-limonene, occupying a different position in neural similarity-space for smell

Source: Adapted from Margot (2009).

representation of R-limonene is neurally more similar to (but distinct from) your neural representation of the *minty-smelling, third molecular-type* than your neural representation of citrus-smelling, first molecular-type—the exact opposite of how things stand in the actual world (Figure 19.1A). As a result, in this scenario, you categorize R-limonene, not with the citrus-smelling first molecular-type, but with the minty-smelling third molecular-type. Let us also stipulate that your (different) neural representation of R-limonene in this scenario has the function of indicating the very *same* molecular-type as humans' actual neural representation of R-limonene. Analogy: different words, in different languages, can indicate the same thing. We can call this a *coincidental variation case*: in the actual situation and the counterfactual situation, there is a complete coincidence in what your olfactory system has the function of indicating, but there is variation in the realizing neural states. Clearly, given the empirical findings, in this scenario, R-limonene would be similar to (but distinct from) the *minty-smelling, third molecular-type*. That is, it would smell *minty* to you, and not citrus-like as in the actual world.

But this verdict contradicts reductive-externalist representationalism. For, on this view, in both the actual situation and the counterfactual situation, you should *phenomenally represent* the same smell quale, despite radical difference in your neural pattern and behavioral dispositions, because those different neural patterns have the *function of indicating* the same molecular-type. The case for the possibility of this counterexample is not based on its conceivability; it is supported empirically.

The moral: when it comes to the qualities of olfactory consciousness, our internally neural processing is running the show, not the molecular-types that this internal-neural processing has the function of indicating in the world. If the empirical findings mentioned above do not convince of this, what empirical findings *would*? Call this the *argument from neuroscience* against reductive-externalist representationalism.

The argument from neuroscience is general. For instance, pain researchers generally hold that pain intensity is fixed by neural firing rates in the pain-matrix (this is linearly related to sensory pain intensity), not by ‘the size and extent’ of the bodily disturbance indicated by the firing rate (this is an extremely poor predictor of pain intensity and is not even well-defined). From an empirical point of view, reductive-externalist representationalism about pain is a non-starter (Price and Barrell 2012: 203). In fact, the empirical findings are much the same in the domain of color experience (Byrne and Hilbert 2003: Figure 5; Brouwer and Heeger 2009; Bohon et al. 2016). If you agree that the empirical findings about smell and pain do support an ‘internalist’ approach in these cases, then consistency demands you take these similar empirical findings to support a parallel internalist approach in the case of color experience. For instance, contrary to reductive-externalist representationalism, a creature’s visual system could indicate the very same reflectance property of the ball, and yet that creature will be conscious of a totally different color quale (say, *green* rather than blue), because of differences in its color-processing. (This would be another *coincidental variation case* with the structure indicated in Figure 19.1B.)²² True, our *naïve, pretheoretical* view of color experience may be resolutely externalist. But, given the empirical facts, externalism about color experience is no more sustainable than externalism about the experience of pain and smell.²³

²² For recent discussion of the argument from neuroscience against externalism about the qualities of experience, see Chalmers 2005; Cohen 2009: 81ff; Pautz 2014b; Allen 2016: 71–2; Logue 2017; Beck 2018; Berger 2018; Campbell 2018.

²³ There are a few arguments against reductive-externalist representationalism besides the argument from neuroscience described in the text. (i) *The argument from the internalist intuition* (Horgan and Tienson 2002: n. 23; Hawthorne 2004: 352). One problem with this argument is that ‘internalism’ about experience is *not* pretheoretically intuitive—it can only be supported empirically (Pautz 2014a). (ii) *The conceivability of spectrum inversion* (Levine 1997: 109; Chalmers 2010: 400 n. 17 and 415–16; Shoemaker 2019). In response, Tye (2000: 110) questions the conceivability-possibility link on which this argument relies. (iii) *The ancient problem of perceptual variation*, for instance due to shifted spectra (Shoemaker 2019) and attention shifts (Block 2010). Byrne and Hilbert (1997: 272–3) first defended a *selectionist* response to this problem and then later (2003: 17) converted to a *misrepresentation* response.

19.5 NONREDUCTIVE-INTERNALIST REPRESENTATIONALISM

Finally, we turn to a radically different way of developing representationalism: *nonreductive-internalist representationalism*.²⁴

Reductive-externalist representationalism is radically externalist. But the argument from neuroscience reviewed in Section 19.4 supports *internalism* about the experience of traditional ‘secondary qualities’, such as pain, smell, and color. In particular, it supports the following picture: external physical properties—types of bodily disturbance, molecular-types, reflectance-types—cause our internal neural states; then those internal neural states totally fix the character of our experiences of secondary qualities.

If we now combine internalism with the representationalist view that experiences are representational states, then we reach *internalist representationalism*: what pain-qualities, smell-qualities, and color-qualities we phenomenally represent ‘out there’ are somehow entirely fixed by our neural states ‘in here’. For instance, what smell qualities you phenomenally represent in the region before your nose is entirely fixed by your neural processing (see Figure 19.1B), and what pain qualities you phenomenally represent in regions of your body is entirely fixed by your neural processing. This is a very natural view. Since this view is internalist, it accords with the empirical results. Since it is representationalist, it explains the possibility of illusion (e.g. phantom-limb illusion). It differs from the *internal physical state view* criticized in Section 19.3: on internalist representationalism, experiences are not *identical with* internal neural states; rather, they are representational states that are *dependent on* but not identical with internal neural states (see Figure 19.2).

The empirical evidence for internalism and against externalism is strongest for the experience of ‘secondary qualities’. It is weaker when it comes to the experience of ‘primary qualities’ (shape, size, distance, number, motion).²⁵ But, once we accept internalism for the experience of traditional secondary qualities, considerations of uniformity suggest generalizing it to the experience of primary qualities. For instance, since the phenomenal representation of sensible colors is inseparable from the phenomenal representation of extension and location (a ‘laws of appearance’ of the kind discussed at the end of Section 19.3), internalism about the phenomenal representation of sensible

²⁴ Proponents of nonreductive internalist representationalism include Chalmers 2010; Kriegel 2011; Horgan 2014; Pautz 2014b; and Mendelovici 2018.

²⁵ For instance, if the apparent length of a line doubles, then the length that your neural state has *the function of indicating* doubles. (This is true even if the experience of the line is a total hallucination.) So here, unlike in the case of smell, there is a good correlation between the character of experience and the character of the physical properties in the world that our neural states have the function of indicating.

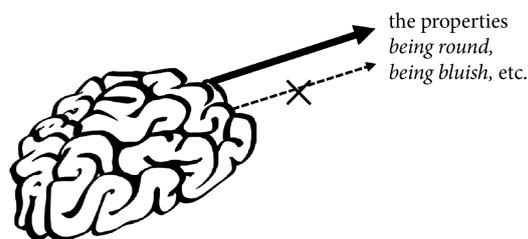


FIGURE 19.2 In having a hallucination of a blue ball, BIV stands in the phenomenal representation relation (the solid arrow) to spatial properties and color properties. But BIV bears no interesting physical relation (such as the indication relation) to these properties (the dashed arrow).

colors leads to internalism about the phenomenal representation of extension and location. Reductive-externalist representationalism fails even here.²⁶

Recall that reductive-externalist representationalists hold that the phenomenal representation relation (the ‘conscious-of relation’) is reducible to the *indication relation*. Internalist representationalists cannot accept this view. According to internalist representationalists, phenomenal representation is unique: it is internally-determined. Therefore, no such standard externalist model for reducing representation applies in the special case of phenomenal representation. In fact, internalist representationalism requires a *non-reductive* picture of sensory consciousness.

One way to see this is to return to the kind of *coincidental variation case* mentioned at the end of Section 19.4 (Figure 19.1). You and your twin in the hypothetical scenario bear the indication relation to the same molecular-type. But, on nonreductive-internalist representationalism, you bear the phenomenal representation relation to different smell qualities. Therefore, the phenomenal representation relation must be distinct from the indication relation.

Or consider the ‘brain in the void’ (BIV). Recall that on reductive-externalist representationalism BIV cannot have *any* experiences. By contrast, on *internalist* representationalism, BIV has all the same experiences as you. For instance, BIV has experiences of pain. And, just like you, BIV has the ball-experience. In having this experience, BIV has a vivid impression as of a round thing being present. On representationalism, this is because BIV phenomenally represents *being round and in front of me* (and not because, say, BIV is presented with a round ‘visual field region’). The brain, then, does not need ‘help from the world’ in order to enable us to phenomenally represent a range of basic properties; rather, it has an *intrinsic capacity* to represent those properties. See Figure 19.2.

²⁶ On the need for a uniform theory, see Cutter (2016: 7–8) and Pautz (2014b: 286–91). For some independent lines of argument against externalism and for internalism about the experience of ‘primary qualities’, see Chalmers (2012: 296–7, 333), Masrour (2015), and McLaughlin (2016: 292).

If the BIV-example is possible, then the phenomenal representation relation is an irreducible, non-physical relation. The argument for this is simple:

1. BIV bears the phenomenal representation relation to the property *being round*.²⁷
2. BIV does not bear the indication relation, or indeed *any* interesting physical relation, to the property being round.
3. Therefore, the phenomenal representation relation is distinct from any physical relation.

As Ned Block (2019) says in a discussion of this argument, ‘we internalists must acknowledge an irreducible representation relation.’ (See also Speaks 2015: 271–2.)

Nonreductive-internalist representationalism may seem to require a bizarre dualism. This is not true. It is also compatible with *grounding physicalism*. Even if the phenomenal representation relation is irreducible, it may yet be that, whenever an individual (e.g. you or your BIV twin) bears this relation to some cluster of properties, then this is fully grounded in that individual’s internal brain state. In fact, there might be systematic ‘grounding laws’ of the form: if one undergoes brain state *B*, then this grounds one’s phenomenally representing property *f(B)*, where *f* is a systematic function from brain states onto perceptible properties. In that sense, there is an explanation of consciousness; it just does not take the form of a reductive definition. Compare: some hold that normative facts about what *ought* to be the case are irreducible to, but still *fully grounded in*, non-normative facts about what *is* the case. Nonreductive-internalist representationalism is like dualism in one respect: just as dualists need ‘psychophysical laws’, so this view needs special ‘grounding laws’. This adds to the complexity of the view. But in another way, it is like physicalism and unlike dualism: by contrast to dualism, it implies that the facts about consciousness cannot ‘float free’ from the underlying physical facts.

I have said that internalist representationalism requires a *non-reductive* view of phenomenal representation (the ‘arrow’ in Figure 19.2). It is also naturally combined with an *illusionist* view of the sensible properties, like sensible colors, smell qualities, tickle-qualities, and so on (the properties that the arrow is ‘pointed at’). On this view, such sensible properties are *neither* properties of external physical items (objects’ surfaces, odorants, bodily regions, and so on) nor are they properties of internal items (e.g. brain states or ‘sense data’). They only live in the content of experience.

This is a big issue, but let me briefly explain why internalist representationalism is most naturally combined with illusionism. The basic point is that, if you take an inventory

²⁷ Some internalist representationalists with nominalists sympathies (Kriegel 2011: ch 3; Mendelovici 2018) might reject premise 1; they might agree that BIV has an experience according to which a round thing is present but deny that this means that BIV stands in a representation relation to the abstract property, *being round*. (This is an example of ‘concrete representationalism’ which we discussed in Section 19.3.) As against this kind of nominalism, by having visual experiences, BIV is in a position to know timeless, necessary truths about the *resemblances* of shapes. And these truths have no nominalistically acceptable paraphrases; they are truths about shape *properties* (Yi 2017). So BIV *is* mentally related to shape properties like *being round*.

of the properties instantiated in the external world or in the brain, you find that they are all poor candidates to constitute the sensible properties. Let us look at a few options.

- (I) To begin with, the kind of empirical results mentioned in Section 19.4 clearly suggest that sensible properties cannot be identified with objective, physical properties, contrary to reductive-externalists. For instance, look back at Figure 19.1A. Given the mismatch, the smell qualities cannot be identified with the corresponding molecular-types.²⁸ Further, *internalist representationalists* certainly cannot identify the sensible properties with such physical properties. They hold that BIV phenomenally represents sensible properties (sensible colors, smell qualities, etc.), but it is impossible to see how BIV could phenomenally represent any such external physical properties (reflectance properties, molecular-types, etc.), as it has never interacted with them.
- (II) Some internalist representationalists hold that sensible properties are identical with *dispositions of external objects to produce neural states in us* (e.g. Kriegel 2009: 90). But this approach faces several problems.²⁹
- (III) Finally, some have suggested the strange view that sensible properties are *neural properties of our own brain states*. For instance, smell qualities are just identical with neural patterns (see Figure 19.1A). Or again, when one has the blue-ball experience, the bluish (or bluish*) quality one experiences is a *neural property of one's own brain state* (Block 2010: 24, 56, fn. 2). (Of course, this differs from the view of Jackson 1977 and Peacocke 2008, criticized in Section 19.2, that it is a property of a *literally round visual field region or sense datum*.) One problem with this concerns the spatial character of experience. For instance, if the bluish quality is in fact a neural property of one's own (non-round) brain state, then (absent some kind of projective-binding error) it becomes difficult to explain the phenomenologically obvious fact that it *seems* to one to fill a round region. Another point is that internalist representationalists *already* hold that the brain has the innate capacity to phenomenally represent certain properties that are not properties of the brain and that may not be properties of anything at all (Block 2019). That is their view in the case of the property *being round* in the BIV case (see Figure 19.2)? So why should it be any different for the property of being bluish (or the quality of being minty, and so on)? Should they not take a *parallel* representationalist view here: that the brain has an innate capacity to phenomenally represent *being bluish* (or being bluish*), even

²⁸ In reply, realists might concede that smell qualia are *distinct from* molecular-types (because of the mismatch), but then try to save realism by holding that they are *grounded in* molecular-types (Allen 2016: 128–9). But this view would require a giant mismatch between the resemblances among the small qualia and the resemblances among the distinct molecular-types they are grounded in. While this view is logically coherent, it is very odd; it would require an endless raft of highly irregular, unsystematizable grounding connections between distinct properties (one for each *smell quale*). Alternative views avoid this.

²⁹ For a 'psychosemantic' problem with this view, see Byrne and Hilbert 2017: 182–3; McGrath ms; Pautz 2014b: 292–3.

though it is *not* a property of the brain? Why insist that *being bluish* needs to be a property of the brain (but somehow ‘projected outward’), if one thinks that other properties involved in the phenomenal character of experience (namely, spatial properties like *being round*) need *not* be properties of the brain.

In sum, sensible properties (‘qualia’) cannot be plausibly located in the external world or in mind-brain. The best view may be that they are not located anywhere, even if they *appear to be* located in regions and surfaces in space.

Nonreductive-internalist representationalism tells a totally different origin story for sensory consciousness than externalist-reductive representationalism. Recall that, on the externalist picture, the sensible properties were in the world before brains evolved; the brain enabled us to be conscious of (‘phenomenally represent’) sensible properties only because it had a history of indicating their occurrence in the world. The brain, then, did not generate the experience of sensible properties all on its own; it needed help from the world. By contrast, the kind of internalist representationalism I have described holds that the sensible properties were *not* in the world before sentient creatures evolved. (In fact, on an illusionist version, they are not in the world *now*.) The physical world is intrinsically devoid of qualities like *red-as-we-see-it*, *citrus-like*, and so on; it is only filled with quantitative properties like mass, charge, and spin. So, in generating the phenomenal representation of sensible properties, the brain did not have any help from the world. It evolved to enable us to phenomenally represent properties of a wholly novel sort that never have been instantiated in the world. Most would agree that the brain enables us to experience *pain qualities* that were not antecedently out there. Internalist representationalists just extend the same view to the experience of olfactory qualities, audible qualities, color qualities. The brain is inventive. It provides a useful but false model of the physical world. Much of our experience of the world is a case of adaptive illusion. The overriding function of the sensory systems is to enhance adaptive fitness, not to represent the way the world really is. However, internalist representationalists need not say that experience is *wholly* illusory. We also evolved brain states that enable us to phenomenally represent *shapes* and other spatial properties. Here, they might say, there is a happy agreement between appearance and reality. For here enhancing adaptive fitness amounted to enhancing veridicality. The result is an updated version of the traditional Lockean thesis that our ideas of primary qualities resemble external objects while our ideas of sensible qualities do not.³⁰

That, then, is nonreductive-internalist representationalism. Why believe it? One of the most puzzling features of sensory consciousness is its Janus-faced nature. On the one hand, we have seen that sensory consciousness is essentially externally-directed

³⁰ For a priori arguments for illusionist forms of internalist representationalism, see Chalmers 2010 and Horgan 2014. For empirical arguments, see Pautz 2014b. Chalmers (2012: 296–7, 333) generalizes illusionism to the *spatial properties* that we phenomenally represent. He calls them ‘Edenic’. Russell (1912, 29) advocated the same view (except that Russell was a sense datum theorist while Chalmers is a representationalist).

(Section 19.2): experiences essentially present qualities arrayed in *space*. On the other hand, we have seen that empirical investigation shows that sensory consciousness is also internally-dependent (Section 19.4). Experiences spring from the inside but they also point outwards. Nonreductive internalist representationalism explains this while at the same time avoiding problematic items such as ‘sensory field regions’ and ‘sense data’ in a private space.

Nonreductive-internalist representationalism has another advantage over reductive-externalist representationalism. Recall that reductive-externalist representationalists hold that phenomenal representation (and hence phenomenal consciousness) is reducible to *indication plus cognitive-rational accessibility*. So they think that thought and reasons are *explanatorily prior* to consciousness. In particular, Tye (2000: 62) advocates a *cognitive analysis* of consciousness and Dretske (2006: 174) advocates a *reasons-based analysis* of consciousness. On pain of circularity, they must also think that there is an account of thought and reasons that does not advert in any way to consciousness: for instance, an account in terms of reliable indication relations to the external environment. (Compare: *being a man* is explanatorily prior to, and can be explained independently of, *being a bachelor*.) But it is just intuitively implausible that thought and reasons could be explained independently of consciousness. Moreover, such accounts are open to counterexamples and other problems. Nonreductive-internalist representationalists can reverse the direction of explanation. Since they do not reductively explain phenomenal representation (sensory consciousness) in terms of thought and reason (indeed they think it is an irreducible relation), they are free to hold that phenomenal representation explains, and is prior to, thought and reason. They can take a *consciousness-first* approach, on which our conscious experience as of basic properties (itself grounded in our internal brain states) grounds our most cognitive-capacities and reasons. This is by far the more natural view.³¹

True, nonreductive-internalist representationalism may be more complex than reductive-externalist representationalism. Reductive-externalist representationalism *identifies* the phenomenal representation relation with the indication relation. And it *identifies* the sensible properties with external physical properties. Identities are unique in that they do not add to the complexity of our theory; in fact, they conduce to simplicity. By contrast, nonreductive-internalist representationalism says that the phenomenal representation relation is irreducible. It requires basic ‘grounding laws’ (as yet unknown) linking quite different states: brain states and states of bearing this irreducible relation to certain basic perceptible properties (see Figure 19.2). And these grounding laws add to the complexity of our theory of the world, in much the same way as the ‘psychophysical laws’ of traditional dualism. But maybe that is just the way the world is.

³¹ For the idea that consciousness is the source of our most basic reasons and cognitive capacities, see Russell 1912; Pryor 2000; Chalmers 2012: 467; Kriegel 2011. See also Bourget and Mendelovici (Chapter 26, this volume) and the references therein.

19.6 CONCLUSION

Representationalism can provide an excellent account of a host of otherwise puzzling perceptual phenomena (Sections 19.2–3). If true, it changes the mind–body problem. It posits a dyadic mental relation, the relation of *phenomenally representing* (‘predicating’) a property. Some think that we can reductively explain this relation using one of our standard externalist models for reducing representation (Section 19.4). Others think that it makes a new problem for reductive physicalism (Section 19.5). In particular, phenomenal representation may be unique in that it is internally-determined. If so, our usual models for reductively explaining representation do not apply in the special case of phenomenal representation.³²

REFERENCES

- Alford-Duguid, D. and Arsenault, M. (2007), ‘On the Explanatory Power of Hallucination’, *Synthese*, 194: 1765–85.
- Allen, K. (2003), ‘Blur’, *Philosophical Studies*, 162: 257–73.
- Allen, K. (2016), *A Naïve Realist Theory of Colour*. Oxford: Oxford University Press.
- Armstrong, D. (1968), *A Materialist Theory of Mind*. London: Routledge.
- Batty, C. (2010), ‘A Representational Account of Olfactory Experience’, *Canadian Journal of Philosophy*, 40: 511–38.
- Beck, O. (2018), ‘Rethinking Naïve Realism’, *Philosophical Studies*, <https://doi.org/10.1007/s11098-018-1030-x>
- Berger, J. (2018), ‘A Defense of Holistic Representationalism’, *Mind and Language*, DOI: 10.1111/mila.12163.
- Block, N. (2010), ‘Attention and Mental Paint’, *Philosophical Issues*, 20: 23–63.
- Block, N. (2019), ‘Arguments Pro and Con on Adam Pautz’s External Directedness Principle’ in A. Pautz and D. Stoljar (eds), *Blockheads! Essays on Ned Block’s Philosophy of Mind and Consciousness*. Cambridge, MA: MIT Press.
- Block, N. ms. ‘The Border Between Seeing and Thinking’.
- Boghossian, P. (2008), ‘Introduction’ in his *Content and Justification*. Oxford: Oxford University Press.
- Boghossian, P. and Velleman, D. (1989), ‘Colour as a Secondary Quality’, *Mind*, 98: 81–103.
- Bohon, K, Hermann, K., Hansen, T., and Conway, B. (2016), ‘Representation of perceptual color space in macaque posterior inferior temporal cortex (the V4 Complex)’, *eNeuro* preprint.
- Brewer, B. (2011), *Perception and Its Objects*. Oxford: Oxford University Press.
- Brewer, B. (2017), ‘The Object View of Perception’, *Topoi* 36: 215–27.
- Brouwer, G. and Heeger, D. (2009), ‘Decoding and reconstructing color from responses in human visual cortex’, *Journal of Neuroscience*, 29: 13992–14003.

³² I am very grateful to two referees for their incisive and helpful comments on an earlier version of this essay.

- Byrne, A. (2001), 'Intentionalism Defended', *Philosophical Review*, 110: 199–240.
- Byrne, A. and Hilbert, D. (1997), 'Colors and Reflectances' in A. Byrne and D. Hilbert, *Readings on Color: Volume 1*. Cambridge, MA: MIT Press.
- Byrne, A. and Hilbert, D. (2003), 'Color Realism and Color Science', *Behavioral and Brain Sciences*, 26: 3–21.
- Byrne, A. and Hilbert, D. (2017), 'Color Relationalism and Relativism', *Topics in Cognitive Science*, 9: 172–92.
- Burge, T. (2003), 'Qualia and Intentional Content: Reply to Block' in M. Hahn and B. Ramberg (eds), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, MA: MIT Press, 405–15.
- Campbell, K. (2018), 'Does That Which Makes a Sensation of Blue a Mental Fact Escape Us?' in Derek Brown and Fiona MacPherson (eds), *The Routledge Handbook of Philosophy of Colour*. Abingdon: Routledge.
- Campbell, J. and Cassam, Q. (2014), *Berkeley's Puzzle*. Oxford: Oxford University Press.
- Chalmers, D. (2005), Representationalism Showdown. <http://fragments.consc.net/djc/2005/09/representationa.html>
- Chalmers, D. (2010), *The Character of Consciousness*. Oxford: Oxford University Press.
- Chalmers, D. (2012), *Constructing the World*. Oxford: Oxford University Press.
- Cohen, J. (2009), *The Red and the Real*. Oxford: Oxford University Press.
- Cowart, B. J. and Rawson, N. E. (2001), 'Olfaction' in E. Goldstein (ed.), *The Blackwell Handbook of Perception*. Oxford: Blackwell Publishers.
- Cutter, B. (2016), 'Color and Shape: A Plea for Equal Treatment', *Philosophers' Imprint*, 16: 1–11.
- Dretske, F. (1995), *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Dretske, F. (2006), 'Perception without Awareness', in J. Hawthorne and T. Szabo Gendler (eds), *Perceptual Experience*. Oxford: Oxford University Press, 147–80.
- Hawthorne, J. (2004), 'Why Humeans Are Out of their Minds', *Noûs*, 38: 351–8.
- Horgan, T. (2014), 'Phenomenal Intentionality and Secondary Qualities' in B. Brogaard (ed.), *Does Perception Have Content?* Oxford: Oxford University Press,
- Horgan, T. and Tienson, J. (2002), 'The Intentionality of Phenomenology and the Phenomenology of Intentionality' in D. Chalmers (ed.), *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press.
- Howard, J. D., Plailly, J., Grueschow, M., Haynes, J. D., and Gottfried, J. A. (2009), 'Odor Quality Coding and Categorization in Human Posterior Piriform Cortex', *Nature Neuroscience* 12: 932–9.
- Jackson, F. (1977), *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- Kriegel, U. (2009), *Subjective Consciousness: A Self-Representational Theory*. Oxford: Oxford University Press.
- Kriegel, U. (2011), *The Sources of Intentionality*. Oxford: Oxford University Press.
- Langsam, H. (2018), 'Why Intentionalism Cannot Explain Phenomenal Character', *Erkenntnis*. DOI 10.1007/s10670-018-0031-7
- Levine, J. (1997), 'Are Qualia Just Representations?' *Mind and Language*, 12: 101–13.
- Lewis, D. (1986), *On the Plurality of Worlds*. Oxford: Blackwell.
- Logue, H. (2017), 'Are Experiences Just Representations?' in B. Nanay (ed.), *Current Controversies in the Philosophy of Perception*. New York: Routledge, 43–56.
- McGrath, M. ms. The Metaphysics of Looks.

- McLaughlin, B. (2012), 'On Justifying Neurobiologicalism for Consciousness' in S. Gozzano and C. Hill (eds) *New Perspectives on Type Identity*. Cambridge: Cambridge University Press, 206–28.
- McLaughlin, B. (2016), 'The Skewed View from Here: Normal Geometrical Misperception' *Philosophical Topics*, 44: 231–99.
- Margot, C. (2009), 'A Noseful of Objects', *Nature Neuroscience*, 12: 813–14.
- Masrour, F. (2015), 'The Geometry of Visual Space and the Nature of Visual Experience', *Philosophical Studies*, 172: 1813–32.
- Mendelovici, A. (2018), *The Phenomenal Basis of Intentionality*. Oxford: Oxford University Press.
- Moran, A. (2018), *Austinian Disjunctivism Defended: A Presentational Theory of Visual Experience*. Ph.D. Thesis, University of Cambridge.
- Moran, A. (Forthcoming), 'Naive Realism, Hallucination and Causation: A New Response to the Screening Off Argument', *Australasian Journal of Philosophy*, DOI 10.1080/00048402.2018.1458142.
- Morgan, J. ms. The Representational Limits of Perception.
- Papineau, D. (2014), 'Sensory Experience and Representational Properties', *Proceedings of the Aristotelian Society*, 114: 1–33.
- Papineau, D. (2016), 'Against Representationalism (about Experience)', *International Journal of Philosophical Studies*, 24: 324–47.
- Pautz, A. (2010), 'Why Explain Experience in terms of Content?' in B. Nanay (ed.), *Perceiving the World*. Oxford: Oxford University Press, 254–309.
- Pautz, A. (2014a), 'The Real Trouble with Armchair Arguments Against Phenomenal Externalism' in M. Spervak and J. Kallestrup (eds), *New Waves in Philosophy of Mind*. Houndmills: Palgrave Macmillan, 153–84.
- Pautz, A. (2014b), 'The Real Trouble for Phenomenal Externalists: New Empirical Evidence for a Brain-Based Theory of Sensory Consciousness' in R. Brown (ed.), *Consciousness Inside and Out*. New York: Springer, 237–98.
- Pautz, A. (2017), 'Experiences are Representations: An Empirical Argument' in B. Nanay (ed.), *Current Debates in Philosophy of Perception*. New York: Routledge, 23–42.
- Peacocke, C. (2008), 'Sensational Properties: Theses to Accept and Theses to Reject', *Revue Internationale de Philosophie*, 62: 7–24.
- Perkins, R. and Bayne, T. (2013), 'Representationalism and the Problem of Vagueness', *Philosophical Studies*, 162: 71–86.
- Price, D. and Barrell, J. (2012), *Inner Experience and Neuroscience*. Cambridge, MA: MIT Press.
- Price, H. H. (1932), *Perception*. London: Methuen.
- Pryor, J. (2000), 'The Skeptic and the Dogmatist', *Noûs*, 34: 517–49.
- Russell, B. (1912), *The Problems of Philosophy*. London: Williams and Norgate.
- Shoemaker, S. (2019), 'Phenomenal Character and Physicalism' in A. Pautz and D. Stoljar (eds), *Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness*. Cambridge, MA: MIT Press.
- Sider, T. (2010), *Logic for Philosophers*. Oxford: Oxford University Press.
- Smith, A. (2008), 'Translucent Experience', *Philosophical Studies*, 140: 197–212.
- Soames, S. (2010), *What is Meaning?* Princeton, NJ: Princeton University Press.
- Speaks, J. (2015), *The Phenomenal and the Representational*. Oxford: Oxford University Press.
- Speaks, J. (2017), 'Reply to Critics', *Philosophy and Phenomenological Research*, 95: 492–506.

- Travis, C. (2004), 'The Silence of the Senses', *Mind*, 113: 57–94.
- Tye, M. (1995), *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.
- Tye, M. (2000), *Consciousness, Color and Content*. Cambridge, MA: MIT Press.
- Tye, M. (2009), *Consciousness Revisited*. Cambridge, MA: MIT Press.
- Yi, B. (2017), Nominalism and Comparative Similarity. *Erkenntnis* , <https://doi.org/10.1007/s10670-017-9914-2>.
- Youngentob, S. L., Johnson, B. A., Leon, M., Sheehe, P. R., and Kent, P. F. (2006), 'Predicting Odorant Quality Perceptions from Multidimensional Scaling of Olfactory Bulb Glomerular Activity Patterns', *Behavioral Neuroscience*, 120: 1337–45.