Representationalism About Consciousness

William Seager (University of Toronto at Scarborough)
David Bourget (Australian National University)

Philosophers have traditionally drawn a sharp distinction between phenomenal and intentional states. Phenomenal states are states with phenomenal or subjective character – something it's like to be in them. The clearest examples of phenomenal states are perceptions, emotions, and sensations, which involve specific qualitative or sensory characters. Intentional states, such as beliefs, are mental states which represent something as being in a certain way. It has been commonly held that the intentional aspects of mental states lack phenomenal character while their phenomenal aspects lack intentionality. Modern representationalism about consciousness (MR) challenges this traditional distinction with the claim that phenomenal character is a species of and exhausted by representational content.

1. The Way of Ideas

MR is often conflated with classical representationalism (CR). We will discuss CR first in order to highlight the contrast between old and new representationalism and bring out some of the strengths of the latter.

CR is an answer to a simple question: what are we aware of in perception? Suppose that someone is consciously perceiving a bright red cardinal perched at a bird feeder on a clear winter day. Commonsense strongly suggests that this bird and how it looks and moves is what our perceiver is aware of. It is surely the business of anyone trying to understand perception to explain how it is that perceivers become aware of things such as red birds, feeders, the blue sky and white snow. However, the slightest acquaintance with the history of philosophy reveals that the route from "perceptual experience" to such normal objects of experience is far from clear or straightforward.

Philosophers have often assumed that, necessarily, when one is undergoing a perceptual experience that might be described either as involving or as an awareness of certain qualities (colors, shape, etc.), one must be in presence of an object with these qualities. C. D. Broad, for example, could not believe we could "see the property of bentness exhibited in a concrete instance, if in fact nothing was present to our minds that possessed that property" (Broad 1952, p. 241). Given the possibility of illusion and hallucination, it was commonly inferred that in perception we are not directly aware of external objects. CR endorses the conclusion that what we directly perceive are mental entities variously called "ideas," "impressions" or

1

"sense-data." This view has been called "representationalism" because such internal impressions act as pictures or signs of the external world.

While not without insight, CR has several unpalatable consequences. First, it literally cuts us off from the world and each other. No one has ever been directly aware of anything but their own minds, hence no two people have ever been perceptually aware of the same thing. Furthermore, insofar as CR makes it hard to justify beliefs about the external world, general skepticism threatens. CR is also radically revisionist about the phenomenology of perception. Intuitively, we directly perceive things in the external world rather than always and only our own mental states. The phenomenology of experience is *of* the features of the experienced objects themselves and no distinctively mental features intrude into perception (this is the *transparency* of experience, explored below). CR is potentially metaphysically revisionist as well, for the obvious solution to CR's epistemological catastrophe is to reform our conception of the material world itself. This was Berkeley's famous response, but the history of philosophy is rife with other suggestions along the same lines.

Perhaps the main problem with CR's posit of mental entities which possess all the features we are aware of in perception is that it seems to be irreconcilable with a physicalist world view. Consider this simple argument against materialism. Imagine as vividly as you can the Canadian flag, and note the shape and redness of the central maple leaf. Now consider that nothing in your brain is a bright red maple-leaf shaped blob. Nothing in the brain can be identified with the flag you have just mentally generated (and nothing outside the brain is a possible candidate either). So much for materialism!

2. Representation to the Rescue

MR rejects the assumption that leads to CR: not every perceptual experience requires that there be something perceived with qualities matching its phenomenal character. Instead, perceptual experience is understood in terms of representation, or intentionality. What an intentional state represents is its content, which can be thought of as *accuracy* or *satisfaction conditions*. For example, if someone believes that snow is white, the belief's content will be true just in case snow is really white. It happens that such a belief would be true, but that doesn't matter to the content: the content of the belief that most mermaids are beautiful is that most mermaids are beautiful, whether or not there are any mermaids and whether or not the majority of them are beautiful. So the idea is that just as one can have a story about mermaids without there being any such objects, one can have a perceptual experience as of a dagger without there being any dagger, or any immaterial stand in for the dagger. This conception of experience should not be conflated with that of CR: according to MR, perception is representational in the sense that it is intentional, not in the sense that it is mediated by "internal pictures." Early proponents of this reply to CR include Anscombe (1965), Armstrong (1968), and Hintikka (1969).

More recently, Harman (1990), Dretske (1995, 2003), and Tye (1995, 2000) have offered representational accounts of perceptual awareness.

Proponents of MR extend the denial of the assumption that leads to CR in two ways. The first is by characterizing the contents of perceptual experience, typically declaring them to be of external objects rather than internal mental entities. CR and its unpalatable implications are thus avoided. When someone perceives a cardinal they are perceiving a bird and not one of their own mental states. While dependent upon there being an active mental representation within the perceiver, proponents of MR deny that perception is indirect, proceeding via an apprehension of this representation. Rather, there is a "presentation" in consciousness of the content of the representation, a content which can be shared by many perceivers. There is no "maple leaf in the head." When one imagines a Canadian flag, one is aware of the content of a flag-representation which encodes shape and color information. This encoding does not have to be flag shaped and colored red. Whatever the vehicle of this representational content might be, perhaps a neural state, there is no need for there to be any awareness of it. While accepting MR does not refute skepticism, at least it allows for the possibility of "directly veridical" perceptual experiences in its endorsement of the claim that the contents of awareness are *of* a world independent of the perceiver's mind.

The second extension is a core claim we take to define MR (although weaker views have been dubbed "representational"). The claim is that a state's phenomenal character is exhausted by its content. The exact meaning of this *exhaustion thesis* is that for every phenomenal character P there is some content C such that a state with P is nothing more than a phenomenal state with C as content. It does not follow from the exhaustion thesis that the content which specifies the phenomenal character of a phenomenal state suffices to make any state that has it conscious. What follows is that, given that a state is in fact a phenomenal state, its phenomenal character is completely specified by its representational content.

We can now discern three key projects related to MR. The first is that of determining whether its defining claim – the exhaustion thesis – is true. The second is that of explicating the fundamental difference between phenomenal and non-phenomenal states. The third project is that of developing a theory of representation strong and stable enough to support MR. Unfortunately, there is no acknowledged theory of mental representation, so we must assume for the time being that an appropriate account can be developed.

3. The Exhaustion Thesis

While it perhaps seems rather obvious that experience carries information which is presented in consciousness, the exhaustion thesis is interesting and controversial. The thesis entails that for each phenomenal character, P, there is a content C, such that (1) all phenomenal states with P have C and (2) all phenomenal states that carry C have P.

Experiences of bodily sensations are often brought up against (1). What is the content of an experience of pain, for example (see Searle 1983)? An answer commonly given is that such experiences represent particular types of bodily damage, malfunction, suboptimality, stress, physiological fatigue or other sorts of misfortune (Armstrong 1968; Tye 1995; Bain 2003). The idea is simply that in pain the body, or a part of it, is represented as being a certain way, including a distinctive evaluative component discussed further below. But do headaches, for example, connote "bodily damage" (see Crane, forthcoming)? Lots of information is provided by the experience of a headache: location, intensity, and duration, in addition to the distinctive evaluative feature that one's head is "not right." This is what is intended under the rubric of bodily damage.

Moods and unfocused feelings are frequently suggested as counterexamples to MR. How could it be that an experience of elation, for example, has representational content? The key to an account of the content of moods and unfocused feelings is to distinguish global vs. local aspects of representation. If someone puts on some rose colored spectacles, this changes the way everything looks and – apart from one's knowledge of how colored glass works – the world itself appears to have undergone a general change in color. It is no accident that this is the metaphor we use for positive moods such as elation. In terms of representation, elation is a global transformation rather akin to turning up the brightness of a television: as a first approximation we can say that it involves the superimposition of goodness over everything one perceives.

Most opponents of MR grant that all experiences have representational contents but claim, contra to (2), that their phenomenal aspects "outstrip" their contents. Skepticism about (2) has been fueled by the perspectival nature of vision emphasized by Peacocke (1983, ch. 1, 1992). He argues that some perspectival differences in experience cannot be associated with informational differences. To illustrate, suppose you are looking at a tree from a given angle at a given distance and then step away from the tree without changing your viewing angle. It might seem that your experiences of the tree before and after share the same content: both represent a tree of given dimensions at a given position with many other unchanging features. Yet the two experiences are qualitatively different since, as Peacocke puts it, the tree initially occupies more of your visual field. However, informational differences that account for the change in phenomenal character are not hard to find. For example, the resolution with which you are representing the surfaces of the tree changes as you step away from it. Also, the two experiences represent the tree as being at different distances from you (Tye 1996 emphasizes such relational properties).

One might still suspect that there are experiential differences stemming not from position but rather from the representational system itself. Blurry vision is often raised as an aspect of phenomenal consciousness which outstrips representational content (Boghossian & Velleman 1989). But what we *believe* about the world must be separated from the way we experience the world. The satisfaction conditions of

the visual content of a shortsighted perceiver without glasses *is* a world in which things have fuzzy edges. To see this, imagine building a strange environment in which the edges of objects are deliberately "fuzzed out." If done right there would be a vantage point from which perceivers could not tell whether they were wearing their glasses in the fuzzy room or were without their glasses in the ordinary room (see also Tye 2002).

Another difficulty for (2) stems from inter-modal perception. Block (2003, 1995) offers a case in which vision and hearing seem to both represent an object "as above" in a way contradicting (2). Here the obvious counter is that vision and audition represent distinct properties beyond location. Among other things, it seems that vision must represent some color, or at least brightness and shading, while audition provides information about pitch and timbre. The fact that the total content of a visual experience and an auditory experience can agree on certain represented sub-features of the environment is harmless so long as additional represented features distinguish the experiences (see Lycan 1996, pp. 135–6).

The classic thought experiment of the "inverted spectrum" (Locke 1690/1990) puts pressure on (1) and (2) simultaneously. A color-invert experiences colors systematically transposed to their spectral opposites: red looks green, blue looks yellow, etc. Suppose there were inverts among us calling ripe tomatoes "red" even as they experience them as green, the sky "blue" even though they see it as yellow, and so on. The question would then arise as to whether their experience of red represents the property we call "green" and their experience of green represents the property we call "red." If this were the case, the character of experiences of red would not determine their content, because ours would be representing red while theirs represents green. Conversely, the content of our experiences of red would not determine any particular phenomenal character because inverts would have experiences of green with this content. This would contradict both (1) and (2) above.

However, it does not follow from the fact that inverts use the word "red" to describe ripe tomatoes that their experiences really represent tomatoes as red. MR can maintain that inverts *believe* that ripe tomatoes are red even though their experiences represent them as green. Such perceivers are merely victims of an inversion of semantic belief contents with respect to phenomenal contents, which is harmless to MR.

4. Wide vs. Narrow Representationalism

Though MR could answer the preceding inverted spectrum argument, things get more complicated when we take into account the commitment of many representational theories to *content externalism.*

Externalism asserts that the content of mental states is determined at least in part by environment or history. For example, one externalist view would be that a representational vehicle represents what causes it in normal conditions. By contrast,

internalism is the view that mental content is determined solely by one's intrinsic state. Internalism has traditionally been the "default" view of mental content, but recently many theorists have adopted externalism in the wake of Kripke's (1972), Putnam's (1975), and Burge's (1979) influential criticisms of internalism in philosophy of language.

Putnam (1975) famously argued that the term "water" has H_2O as its content because of the causal-historical relation between the introduction and use of "water" and the local prevalence of H_2O. On Putnam's imaginary Twin-Earth, where the lakes, seas, and organisms are full of an alternative, but superficially indistinguishable compound XYZ, the Twin-Earth term "water" means XYZ rather than H_2O. It has been argued that thought content similarly depends upon causal-historical relations, so that the natives of Earth and Twin-Earth are thinking *different* thoughts which they all express by saying "water is wet."

MR naturally bifurcates into externalist and internalist versions, depending upon the favored theory of mental representation. Many proponents of MR endorse the externalist view (e.g., Tye, Lycan, Dretske). The resulting view, *phenomenal externalism* (PE), faces a number of difficulties.

Perhaps the most bizarre consequence of PE is the possibility of "philosophical zombies": creatures physically identical to ourselves but which utterly lack consciousness. Since PE implies that consciousness depends upon content constituting relations, a creature lacking these will not be conscious. An example would be Davidson's (1987) *Swampman*, a spontaneously created physical duplicate of himself lacking causal links with the world or evolutionary history. PE faces the unpleasant choice of denying that Swampman is conscious (Dretske 1995) or rather unattractively modifying the theory of representation to include Swampman among the conscious (Tye 2000). Internalists, of course, face no such difficulties.

A revamped version of the inverted spectrum also threatens PE. Block (1990, 1996) envisages "Inverted Earth" – a place much like Earth except that the actual colors of things and color terms are – somehow – inverted. Now imagine some people are unknowingly transported to Inverted Earth and, during the trip, are given a secret operation which turns them into color-inverts via the implantation of an Inversion Device (ID). When they arrive, they will notice no difference in the colors of things. Block argues that eventually the travelers' color vision states will, via the mechanisms of externalist content fixation, come to veridically represent the colors of Inverted Earth even though there will be no phenomenal change. Such a representational change with no change in experience would refute MR.

A defender of MR can boldly reply that there has after all been an unnoticed phenomenal change with no internal change in the traveler (see for instance Lycan 1996). This simply amounts to biting the bullet by accepting the bizarre consequence of PE that two individuals who are in identical intrinsic states can have distinct phenomenal experiences.

This line of reply seems implausible. It asserts that the travelers' color vision gradually de-inverts while they live on Inverted Earth with the ID in place. At first,

red things looked green, although the inversion was disguised by the peculiarities of Inverted Earth. In time, however, red things come to look red again but this shift in vision is so "gradual" (or something) that it is not noticed by the travelers. Despite this, it remains clear that removal of the ID will still cause color vision inversion. It follows that for the *acclimatized* travelers, removal of the ID will make it the case that red things look green. So if one of them were suddenly switched to standard Earth immediately after her ID had been removed, she ought to exclaim that ripe tomatoes look green, much to her surprise. But it seems clear that with the ID removed and back on Earth, everything would look perfectly normal.

It seems preferable for the defender of MR to insist, with Block, that there is indeed a constant representational content (and phenomenality) to the travelers' experience just because there is no intrinsic change in the representational systems of either the traveler or her stay-at-home twin. This reply may seem obvious to the reader who has not been exposed to the wonders of PE. By adding epicycles, PE can also embrace this reply, but the point here is simply that an internalist MR has no difficulty with the inverted Earth thought experiment (see Dretske 1995, ch. 5; Tye 2000, ch. 6; Lycan 1996, ch. 6; Block 1996).

A further issue with wide representationalism stems from the inverted spectrum scenario. All externalist versions of MR are *relational*: they require that a mental state that represents X stands in a given relation to X in the actual world. This cashes out as a requirement that some experiences be veridical, which requires that colors and other such secondary qualities, if they are represented in experience, be objective.

The inverted spectrum thus leads into the problem of the objectivity of color (see Byrne & Hilbert 1997). If the phenomenal states of the inverts and the normals disagree about how the world is, which they must according to MR, then at most one group is correct about the colors of things. If the inverts are a minority of the population then it might seem easy to characterize their color vision as systematically in error. But what if the population is split 50–50, or what if, over time, the inverts come to form the majority? There does not seem to be any principled answer to these questions. This suggests that *neither* inverts nor normals are correctly perceiving the world, at least not if we take color experiences to be representing an objective continuous surface feature of objects.

5. Relational and Projectivist Approaches to the Exhaustion Thesis

An alternative to the relational, externalist approach is a *projectivist*, internalist approach. It may be that the experience of colors as intrinsic, continuous features of surfaces misrepresents the nature of color. We might borrow Hume's idea that color vision works by "gilding or staining all natural objects with the colours, borrowed from internal sentiment" (1777/1975, p. 294). Following Hume, we will want to extend the idea beyond the perception of colors, but we will not agree that the colors we see stem from internal sentiment, if that implies that color experience

involves directly knowable *mental* qualities which we "project" onto things. Rather, the view is that visual experience represents things as possessing color properties which in fact they do not possess as represented, although there is an objective ground for our experiences. The term "projectivism" has an unfortunately wide range of uses. The view that we project mental features onto external objects has been called "literal projectivism" by Shoemaker (1990) and defended by Boghossian and Velleman (1989). The view advanced here is more akin to Shoemaker's "figurative projectivism"(see Wright 2003 for a defense of projectivism).This is compatible with MR, exploiting the fact that representations can be more or less inaccurate, or simply false.

This projectivist approach, unlike the relational, leaves it open to what extent the mental representations which provide the contents of conscious experience are accurate, but it seems intuitively likely that experience harbors more or less serious errors about the natures of things. Even though we experience material objects as made of continuous substance, as possessing definite locations within a three dimensional space and a one dimensional time in which all events are well ordered, and as possessing surfaces upon which colors are continuously spread out, none of these features seems to be actually instantiated in the world. We should not be surprised if nature, cobbling together cognitive representational mechanisms to aid survival, failed to stumble upon the true nature of things.

It's worth emphasizing that the systematic inaccuracy of perceptual experience does not have disastrous epistemological implications. First, there is still room for perception to be informative and largely veridical. Even though some aspects of the world we experience have no echo in nature, many do. Furthermore, we have the ability to form true beliefs on the basis of perceptions that may be misleading. This ability culminates in the scientific picture of the world, which reveals and explains the erroneous aspects of perceptual experience.

Returning to the exhaustion thesis, there is reason to think that many experiences have non-objective features as contents. The perceptual experience of possible perceivers, including the whole range of conscious animal life on Earth and any number of alien creatures throughout the universe, presents a vast panoply of radically diverse modes of perception. MR must handle this by positing an equally vast range of ways of representing things via the cognitive mechanisms of all these more or less different minds. If all that experiences could represent were extant physical properties, this might seem to make it hard to find content that correlates with every possible experience. Furthermore, it is difficult to imagine what some of these experiences are like, that is, to imagine the kind of world that would satisfy their contents (as famously pointed out in Nagel 1974). Arguably, we would not have such difficulties if experiences only represented physical properties, since we can easily form beliefs about the physical properties various creatures might be representing. It thus seems that proponents of MR must posit non-physical contents, which is compatible with the combination of MR and projectivism, but not with the combination of MR, a relational theory of content, and physicalism.

We noted that pain has an evaluative aspect. The experienced world is suffused with value (positive, negative or sometimes neutral). This might be the most basic and primitive form of consciousness, in the service of which follow the wide range and fine nuances of perceptual experience and thought. Maybe the first twinges of inchoate sensation were "valuings" of stimuli as good or bad – to be pursued or to be avoided, perhaps in very primitive organisms (see Humphreys 1993). This would have been the ultimate origin of pleasure and pain which, roughly speaking, encode what is good and bad at the biological level. Just as in the case of color, we need not accept the naïve pronouncements of experience which presents value as an objective feature of things: value might be something that is projected on the external world.

Projectivism can also help explain the nature of emotional consciousness, including the case of moods briefly discussed in §3. Emotional response is very complex and the associated states of consciousness are similarly multifaceted, involving perception of the environment and the body as well as possessing rich cognitive dimensions. But arguably the core of emotion is the experience of value (see Edelman 1992; Damasio 1994; LeDoux 1996; Seager 2000, 2002).

6. Transparent Experience

MR entails an interesting prediction about experience. Since our consciousness is exhausted by the contents of underlying mental representations, there should be nothing apparent to the mind save the way things are represented. It follows that if one should try to attend to the nature of one's own experience, all that one will be able to find are these contents. This has been labeled the *transparency* of experience (an early discussion is in Harman 1990, although the idea can be traced back at least to G. E. Moore (1903); see Kind (2003) and Stoljar (2005) for useful discussions). Specifically, the transparency thesis is the claim that experience and introspection do not make us aware of anything beyond what mental states represent.

We confess that transparency seems so true to our own experience we have difficulty conceiving of consciousness in any other terms. MR does not *follow* from transparency, but MR is the best explanation for this feature of experience (a claim made by Tye 2000; and denied by Stoljar 2005, among others). Transparency can be explicated by considering cases where awareness is non-transparent. The most obvious example is awareness of meaning achieved through linguistic media. Consider how you come to be aware of the meaning of "most mermaids are beautiful." This awareness is indirect and mediated by an awareness of the vehicle of this content. You can't get to the meaning of "most mermaids are beautiful" except by *perceiving* those lexographical black marks which form the vehicle of this content. The transparency of experience entails that there are no mentalistic "marks" which we must be aware of in order to be aware of the "normal" objects of experience.

Although it is natural to explicate transparency in terms of examples of mediated perception, there is a distinction between mediated awareness and non-

transparency. As discussed, many think there are non-representational features of experience which do not contribute to the satisfaction conditions of our states of consciousness but enter into their phenomenal character. Let us call these Qualia (with a capital "Q"). Qualia, if there were such, might be introspectible through a kind of awareness that is not intentional. The traditional view is that Qualia mediate our awareness of the external world, but they need not play this role: we could be directly aware of the external world and simultaneously acquainted with Qualia. However, it's hard to see why there would be such non-mediating Qualia.

Could it be that Qualia are always part of our experience, but that it is "hard" to become aware of them? This flies in the face of characterizations of Qualia as the *most* immediately available and impossible to miss features of experience. Kind (2003) uses an analogy of seeing a landscape through a window where it is possible, if sometimes difficult, to also see the glass itself. Suppose we are looking through a very old and thin window. The landscape beyond looks blurry and wavy. But here the blurriness and waviness are not features of the glass but are the way the world looks *because* of the nature of the glass. But surely it is sometimes possible to see the pane of glass itself. If so, it is because there is some distinctive property of the glass which is visible. The problem in the case of consciousness is that there does not seem to be any such distinctive features of mental states themselves. These states do not intrinsically possess color, shape, sound, smell, or the other "common sensibles" which are experienced as elements of the perceived world. What is this mysterious qualitative feature of conscious experience which normally eludes us but can be appreciated with some "effort"?

An objection to the transparency thesis is that the mere possibility of introspection refutes it. It seems that any discussion of consciousness presupposes that we can take a reflective stance towards our own experiences and regard them as mental, as experiences. Perhaps Wittgenstein had this in mind when he wrote that I can "turn my attention in a particular way on to my own consciousness, and, astonished, say to myself: THIS is supposed to be produced by a process in the brain! – as it were clutching my forehead" (Wittgenstein 1953, I, 412). MR must make room for introspective awareness of our mental states. However, this does not undermine the transparency thesis so long as introspection is thought of as the higher-order representation of mental states (see §8 below). When we become aware of our experiences as such, we form mental representations in addition to, and about, our experiences. Introspective consciousness is of the content of these higher-order representations (but see Loar 2002).

7. The Demarcation Problem

The exhaustion thesis tells us very little about the relation between consciousness and representation. Most importantly, the exhaustion thesis does not reveal the nature of phenomenal consciousness: what is essential to, and characteristic of it. All it says is that phenomenal states can be specified in terms of their contents. The

demarcation problem is to determine what it is about certain representational states which makes them phenomenal (see Kriegel 2002).

The simplest possible solution is to say that what makes a representation phenomenal is just its content. This *radical representationalism* seems preposterous so long as "representation" is construed sufficiently broadly to include anything from markings on paper to mental states, for whatever can be represented in experience can be written about, and markings on paper are not conscious. As far as we know, nobody has ever advocated radical representationalism.

It is therefore important to distinguish radical representationalism from another view we will call *pure representationalism* (our use of the terms "pure" and "impure" maps roughly onto Chalmers' 2004, but not Lycan's 2005). Pure representationalism is the view that phenomenal states are mental representations with contents that explain why they, but not other mental representations, have the phenomenal character they have. What these contents are is left open by pure representationalism, but the most natural account is that mental states which possess the kind of phenomenal character had by perceptual states have qualitative properties as part of their contents. Qualitative properties would be properties of possible perceptible objects as we experience them (such as redness, painfulness, etc.). The idea is that the one and only reason why thoughts about numbers, human rights, and economic systems differ in phenomenal character from perceptual states is that their objects are not qualitative. Thau (2002) defends a kind of pure representationalism. Byrne (2001,2002) also points toward this view, though his stance is not entirely clear.

A tempting objection to pure representationalism is that it is difficult to spell out what "qualitative" means without explicating it as what is common to the properties sensory states represent, which threatens to make the account circular. This is a complex issue we cannot delve into here. Instead, we focus on three main problems pure representationalism faces apart from this one.

The first is what we might call the "blind man problem." It seems possible to think about the properties represented in experience without experiencing any phenomenal character. For example, those born blind presumably can have beliefs about the properties represented in color vision even though (let us suppose) they cannot experience color (see Neander 1998). Similarly, it seems that sighted individuals must be able to think about these properties without experiencing them: otherwise we would spend our time visualizing colors when writing articles like this one. Thau (2002) denies this seemingly obvious fact: he holds that all perceptual experiences represent properties we cannot represent in (non-phenomenal) thoughts. His defense of this conclusion rests on substantial premises in philosophy of language we cannot discuss here, but he appears to be biting a very large bullet.

An alternative approach deploys Frege's sense/reference distinction. Perhaps the properties represented in experience can indeed be represented in thoughts with no, or distinct, phenomenal characters, but only via round-about descriptions of the

form "the content of John's experience" or "the way X tastes." One might then argue that perceptual experiences and states that differ from them in phenomenal character always differ in content at the level of *sense* even though they can have the same content at the level of *reference*.

This Fregean approach gathers support from the intuition that one has to experience red to properly *grasp* the nature of the experience (recall Jackson 1982). Given the exhaustion thesis, this would mean that its content can only be grasped when undergoing it. Intuitively, the difference between sense-level content and reference-level content is that the former is what we can grasp and reason on, while the latter is determined in context by sense but need not be grasped and determined by sense independently of context. From the facts that one can only grasp what is represented in the experience of red by undergoing this experience and sense-level content is content that is grasped, we may conclude that only *experiences* of red have the content they have at the level of sense. (Thau & Byrne 2002 each makes part of this argument; for a discussion of the sense/reference distinction within the context of MR, see Thompson 2003; Chalmers 2004).

A second problem for pure representationalism is that sub-personal processes may provide examples of unconscious mental representations that lack phenomenal character. A simple example stems from binocular vision. If you close one eye and note what you see, then open the other and close the first, you will note a difference caused simply by the locations of your eyes. It seems that, under conditions of normal vision, the brain somehow combines or links the contents of these two ocular viewpoints into the 3D view that informs consciousness. This would mean that we lack consciousness of the individual components even though the individual representations remain active within the system (see Seager 1999).

A response to this argument questions the assumption that binocular vision combines preexisting mental contents. It is clear that it combines two sources of information, but we cannot assume that all states or events that carry information are mental representations.

The preceding problem leads naturally to a third issue. Pure representationalism seems to rely on a rather restrictive account of mental representation. Without asking for a full theory, it seems reasonable to ask the pure representationalist to sketch an account of mental representation that excludes sub-personal informational states. A promising approach here is to invoke the distinction between *derived* and *original* intentionality. States with derived intentionality are states that can have content only if other states have content. Arguably, natural language expressions have derived intentionality: they are meaningless apart from speakers' intentions. One could argue that mere informational states also have only derived intentionality because they can be construed as signs or indicators only given a certain interpretation of their functions. Searle (1990) and Georgalis (2005) posit the inapplicability of the sense-reference distinction (or something like it) to informational states and argue along similar lines. These suggestions are very controversial.

Such difficulties have led most theorists to reject pure representationalism in favor of "impure" accounts which put less burden on representational content while by and large respecting the exhaustion thesis. Tye (1995, 2000), Dretske (1995), and Jackson (2004) endorse variants of impure representationalism which conform to the exhaustion thesis. Crane (2002), Chalmers (2004), and Lycan (1996) hold impure representationalist views which infringe on it to some extent.

What is characteristic of impure representationalism is an appeal to properties of mental representations above and beyond their contents to account for the difference between conscious and unconscious states. Chalmers, Crane and Jackson describe the relevant features as manners of representation, which are ways of relating to contents comparable to attitudes such as believing and desiring. Dretske, Lycan, and Tye give largely functionalist accounts of the distinction.

All forms of impure representationalism can be classified either as reductive or non-reductive: some hold that the extra ingredient which accounts for consciousness is completely physical or functional, others hold that it is not. Non-reductive impure representationalism takes consciousness as at least relatively fundamental, which dashes the hopes of naturalizing consciousness. It does not forego all the other advantages of representationalism however. Here we can only offer a cursory discussion of a representative reductive account championed by Tye (1995, 2000).

Tye claims that phenomenal states are *PANIC* states; they have Poised, Abstract, Nonconceptual Intentional Content. Poised content stands at the periphery of, and ready to affect "higher" or "central" cognitive systems, especially those which underlie beliefs. Abstract content does not require the presence of any particular object for its satisfaction. Finally, nonconceptual content, on Tye's definition, is such that the subject need have no matching concept (e.g. we can experience millions of colors yet lack correspondingly specific concepts). This account is partly functional and partly representational: the property of being abstract is an intrinsic property of contents, but the properties of being poised and nonconceptual are causal properties relating contents to cognitive centers and concepts, respectively.

Tye's three conditions are supposed to explain why sub-personal informational processes do not have phenomenal character. On the kind of account of content that comes with the PANIC theory, the states involved in such processes could share nonconceptual, abstract content with experiences, so their lack of phenomenal character must be explained by their not being poised.

But there are plenty of unconscious, sub-personal processes which leak information into and influence cognitive centers. There are many experiments which show how stimuli which are presented for too short a time for conscious awareness nonetheless modify cognition (see Murphy & Zajonc 1993). More examples come from dichotic listening, in which two distinct sound streams are played to a subject, one to each ear. In these experiments, only one of the two channels is consciously apprehended, but the other channel can produce cognitive effects (Lackner & Garrett 1972). The phenomenon of blindsight might also be appealed to here (see Siewert 1998 for an extensive philosophical discussion). The existence of such

leaks does not immediately refute Tye's theory because we have not shown that their source is located at the fringe or boundary of the higher cognitive system, as required by Tye.

However, consider that perceptual experience appears to depend on multiple stages of processing in the brain. From the receptors of the retina to the structures of the dorsal system that are involved in object-recognition, there are several layers of increasingly abstract representation along pathways spanning a good part of the brain, each of which plays a role in determining what we experience (see Tye's (1995) own discussion). If this is correct, it seems that some events relevant to conscious experience are upstream of others. Mother Nature is too parsimonious in her allocation of resources to design a brain where low-level representations are always carried along with the information abstracted from them – this would defeat the very purpose of abstraction. Now conscious representations which are upstream of others cannot sit at the boundary of the cognitive centers, wherever it is. If this is correct, all we are left with of the "poised" condition is that poised states are apt to impinge on cognitive centers. But we saw that unconscious states with nonconceptual, abstract content can do this (for other criticisms of Tye see Block 1995; Seager 1999, 2003; Kriegel 2002; and Byrne 2003).

This objection proceeds on more or less empirical grounds, so it might be hoped that all such objections could be avoided by devising a PANIC-type theory on the basis of more empirical data. Ultimately, however, the real problem is that of locating consciousness in our metaphysical picture of the world. The important lesson to draw from the foregoing discussion is that proponents of reductive impure representationalism have to resort to traditional functionalist and physicalist solutions to this problem. Similarly, proponents of non-reductive impure representationalism have to deal with the difficulties traditionally associated with dualism. Put differently, the main problem faced by impure representationalism is simply the hard problem of consciousness in its traditional form (for more on the general problem, see Nagel 1974; Jackson 1982; Levine 1983; Chalmers 1995).

8. Introspective Minds

MR is compatible with a range of theories of introspection which we cannot survey here. But there is a view of introspection which seems a natural extension of MR, tying together several strands of the theory into a unified account of the conscious mind. The seeds of this account can be found in Sellars (1956), but it is developed explicitly within MR by Tye (2000) and Dretske (1995). The latter labels it the displaced perception theory of introspection (see also Seager 1999).

To begin outlining this account, think about what introspection provides: *knowledge* of our own mental states. Via introspection we come to know what mental states we are in: what we are thinking, feeling, seeing, hearing, wondering, hoping, etc. Thus a necessary condition of being able to introspect is the possession of the concepts of those mental states we can discover we are in or experiencing via in-

14

trospection.

The family of mental state concepts, with their complex inter-relations, forms the "theoretical" core of Folk Psychology, and it would seem that very few animals on Earth (perhaps *only* human beings) have any acquaintance with it. It follows that very few animals can engage in introspection. And yet intuitively it seems that there are many conscious beings on the Earth. Introspection is thus not essential to consciousness.

Introspection requires a special and sophisticated way of thinking about conscious experience. What is needed is the Wittgensteinian attitude discussed above. Consciousness presents, in the first instance, information about the world, the body, and sometimes the mind, but even in the latter case it does not provide this information as being "about the mind." It takes a special reflective stance wherein we apply the concepts of mental states to our ongoing experience to transform consciousness into introspective knowledge of our own mental states.

What exactly is involved in introspection if not some kind of reflexivity intrinsic to consciousness? Dretske (1995) frequently writes as if the transformation needed to generate introspective knowledge is an *inference* from experience. But we need not suppose that inference is essential (a view which faces difficulties – see Bach 1997; Aydede 2003). A better model is that of concept application itself. It seems to be a pervasive feature of experience that the world is presented to us in terms of the concepts we bring with us: we see tables, chairs, cats, and dogs. We do not infer from some primordial visual ur-material to a world of furniture and pets. Similarly, we come to apply mentalistic concepts to our experience with the same kind of effortless spontaneity.

Whether this account of introspection is correct depends in large part upon the acceptability of MR in general. But there is a nice fit between MR's depiction of consciousness as the representation of an external world and the claim that the mind is not something which is apparent in consciousness unless and until one takes up a reflective stance which permits one to apprehend experience *as* mental.

9. Conclusion

MR provides a powerful account of the mind which incorporates conscious experience in a way that seems intuitively satisfying, avoids the difficulties associated with such views as CR, opens the door for a variety of naturalistic theories of mind, and integrates introspection without making the implausible requirement that all conscious beings have the conceptual equipment necessary to think about mental states as such.

There is much room for argument and progress within the representationalist framework. Beyond the exhaustion thesis, it seems to us that the two most pressing questions are, first, whether the pure or impure approach is to be favored and, second, whether the internalist (and projectivist) or externalist (and relational) approach is best. These two issues should be investigated jointly in light of more

general considerations concerning mental content, including the important project of investigating the specific contents of phenomenal states.

Further Readings

Kim, Jaegwon (2005) Philosophy of Mind, 2nd edn., chs. 8 and 9. Boulder, CO: Westview Press.
Lycan, William (2005) Representational theories of consciousness in the (online) Stanford Encyclopedia of Philosophy.
Seager, Will (1999) *Theories of Consciousness*, chs. 6 and 7. London: Routledge.

References

Anscombe, G. E. (1965) The intentionality of sensation: a grammatical feature. In R. Butler (ed.), *Analytic Philosophy*, second series.
Armstrong, D. M. (1968) *A Materialist Theory of the Mind*. London: Routledge and Kegan Paul.
Aydede, M. (2003) Is introspection inferential? In B. Gertler (ed.), *Privileged Access*. Aldershot: Ashgate.
Bach, K. (1997) Engineering the mind. *Philosophical and Phenomenological Research* 57: 2, 459–68.
Bain, D. (2003) Intentionalism and pain. *Philosophical Quarterly* 53, 502–23.
Block, N. (1990) Inverted earth. *Philosophical Perspectives* 4, 53–79.
Block, N. (1995) On a confusion about the function of consciousness. *Behavioral and Brain Sciences* 18, 227–47.
Block, N. (1996) Mental paint and mental latex. In E. Villanueva (ed.), *Perception*. Ridgeview.
Block, N. (2003) Mental paint. In M. Hahn and B. Ramberg (eds.), *Reflections and Replies*. Cambridge, MA: MIT Press.
Boghossian, P. and Velleman, J. D. (1989) Color as a secondary quality. *Mind* 98, 81–103.
Broad, C. D. (1952) *Scientific Thought*. London: Routledge and Kegan Paul.
Burge, T. (1979) Individualism and the mental. *Midwest Studies in Philosophy* 4, 73–122.
Byrne, A. (2001) Intentionalism defended. *Philosophical Review* 110, 199–240.
Byrne, A. (2002) Something about Mary. *Grazer Philosophische Studien* 63, 27–52.
Byrne, A. (2003) Consciousness and nonconceptual content. *Philosophical Studies* 113, 261–74.
Byrne, A. and Hilbert, D. R. (1997) *Readings on Color, vol. 1: The Philosophy of Color*. Cambridge, MA: MIT Press.
Carruthers, P. (2000) *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge: Cambridge University Press.

Chalmers, D. J. (1995) Facing up to the problem of consciousness. *Journal of Consciousness Studies* 2, 200–19.

Chalmers, D. J. (2004) The representational character of experience. In B. Leiter (ed.), *The Future for Philosophy*. Oxford: Oxford University Press.

Crane, T. (2002) The intentional structure of consciousness. In Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Essays*. Oxford: Oxford University Press.

Crane, T. (2002) (forthcoming) Intentionalism. In Ansgar Beckermann and Brian McLaughlin (eds.), *Oxford Handbook to the Philosophy of Mind*. Oxford: Oxford University Press.

Damasio, A. R. (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Putnam.

Davidson, D. (1987) "Knowing One's Own Mind," *1987*.

Dretske, F. (1995) *Naturalizing the Mind*. Cambridge, MA: MIT Press.

Dretske, F. (2003) Experience as representation. *Philosophical Issues* 13, 67–82.

Edelman, G. (1992) *Bright Air, Brilliant Fire*. New York: Basic Books.

Georgalis, N. (2005) *The Primacy of the Subjective*. Cambridge, MA: MIT Press.

Harman, G. (1990) The intrinsic quality of experience. *Philosophical Perspectives*.

Hintikka, K. J. J. (1969) On the logic of perception. in K. J. J. Hintikka (ed.), *Models for Modalities: Selected Essays*, 161–83. Dordrecht: Reidel.

Hume, D. (1777/1975) An enquiry concerning the principles of morals. In L. Selby-Bigge (ed.), *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 3rd edn., revised by P. H. Nidditch. Oxford: Clarendon Press.

Humphreys, N. (1993) *A History of the Mind*. New York: HarperPerennial.

Jackson, F. (1982) Epiphenomenal qualia. *Philosophical Quarterly* 32, 127–36.

Jackson, F. (2004) Representation and experience. In H. Clapin (ed.), *Representation in Mind*, Elsevier.

Kind, A. (2003) What's so transparent about transparency? *Philosophical Studies* 115, 2225–44.

Kriegel, U. (2002) Panic theory and the prospects for a representational theory of phenomenal consciousness. *Philosophical Psychology* 15, 55–64.

Kripke, S. A. (1972) *Naming and Necessity*. Cambridge, MA: Harvard University Press.

Lackner, J. and Garrett, M. (1972) resolving ambiguity: effects of biasing context in the unattended ear. *Cognition* 1, 359–72.

Ledoux, J. (1996) *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.

Levine, J. (1983) Materialism and qualia: the explanatory gap. *Pacific Philosophical Quarterly* 64, 354–61.

Lewis, D. (1946) *Knowledge and Valuation*. Open Court.

Loar, B. (2002) Transparent experience and the availability of qualia. In Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Essays*. Oxford: Oxford

University Press.

Locke, J. (1690/1990) An Essay Concerning Human Understanding. In P. Nidditch (ed.), Oxford University Press.

Lycan, W. G. (1996) *Consciousness and Experience*. Cambridge, MA: MIT Press.

Lycan, W. G. (2005) Representational theories of consciousness. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2005 edition).

Moore, G. E. (1903) The refutation of idealism. In G. E. Moore (ed.), *Philosophical Studies*, 1–30. 1965, Littlefield, Adams.

Murphy, S. and Zajonc, R. (1993) Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology* 64, 723–39.

Nagel, T. (1974) What is it like to be a bat? *Philosophical Review* 83, 435–50.

Neander, K. (1998) The division of phenomenal labor: a problem for representationalist theories of consciousness. *Philosophical Perspectives* 12, 411–34.

Peacocke, C. (1983) *Sense and Content: Experience, Thought, and Their Relations*. Oxford: Oxford University Press.

Peacocke, C. (1992) Scenarios, concepts, and perception. In T. Crane (ed.), *The Contents of Experience*. Cambridge: Cambridge University Press.

Putnam, H. (1975) The meaning of "meaning," *Minnesota Studies in the Philosophy of Science* 7, 131–93.

Seager, W. E. (1999) *Theories of Consciousness: An Introduction and Assessment*. London: Routledge.

Seager, W. E. (2000) Introspection and the elementary acts of mind. *Dialogue* 39, 53–76.

Seager, W. E. (2002) Emotional introspection. *Consciousness and Cognition* 11: 4, 666–87.

Seager, W. E. (2003) Tye on consciousness: time to panic? *Philosophical Studies* 113, 237–47.

Searle, J. R. (1983) *Intentionality*. Cambridge: Cambridge University Press.

Searle, J. R. (1990) Consciousness, explanatory inversion and cognitive science. *Behavioral and Brain Sciences* 13, 585–642.

Sellars, W. (1956) Empiricism and the philosophy of mind. *Minnesota Studies in the Philosophy of Science* 1, 253–329.

Siewert, C. (1998) *The Significance of Consciousness*. Princeton, NJ: Princeton University Press.

Stoljar, D. (2005) The argument from diaphanousness. In Maite Ezcurdia, Robert Stainton, and Chris Viger (eds.), *Language, Mind and World: Special Issue of the Canadian Journal of Philosophy*.

Thau, M. (2002) *Consciousness and Cognition*. Oxford University Press.

Thompson, B. (2003) The nature of phenomenal content. PhD thesis, University of Arizona.

Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.

Tye, M. (1996) Perceptual experience is a many-layered thing. In E. Villaneuva (ed.), *Perception*, Ridgeview.

Tye, M. (2000) *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.

Tye, M. (2002) Blurry images, double vision, and other oddities: new problems for representationalism? in Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Essays*. Oxford: Oxford University Press.

Wittgenstein, L. (1953) *Philosophical Investigations*. G. E. M. Anscombe, trans. Oxford: Blackwell.

Wright, W. (2003) Projectivist representationalism and color. *Philosophical Psychology* 16, 515–29.