1 The problem of phenomenal consciousness

Consciousness is perceived by many to provide the principal threat to materialist accounts of the mind. This threat has been developed, in somewhat different ways, by a lineage of writers from Nagel (1974) through Jackson (1982, 1986), Levine (1983, 1993) to McGinn (1989, 1991) and Chalmers (1996). While the precise nature of the threat posed by consciousness has tended to vary, the concept of consciousness perceived to underlie this threat has held relatively constant. It is *phenomenal* consciousness that is considered problematic. There are serious problems, if the authors of the above lineage are correct, involved in finding a place for phenomenal consciousness in the natural order. This book is concerned with these problems, with why they are problems, and with whether these problems admit of a solution.

1 What is phenomenal consciousness?

Any study of phenomenal consciousness faces an immediate problem. There is no perspicuous way of defining the associated concept. That is, there is no non-circular way of specifying the content of the concept of phenomenal consciousness that does not rely on concepts that are equally obscure. Attempts to explain its content, accordingly, tend to rely on a number of devices, linguistic and otherwise.

Examples

Attempts to explain what phenomenal consciousness is often proceed by way of examples: the way things look or sound, the way pain feels, and, more generally, the experiential properties of sensations, feelings and experiences. Sensations and feelings will include things such as pain, itches, tickles, orgasms, the feeling one gets just before one sneezes, the feeling one gets just after one has sneezed, the feeling of cold feet, and so on. When experiences are enlisted to provide an explanation of the concept of phenomenal consciousness, it is typically perceptual (and, to

a lesser extent, proprioceptive) experiences that are to the fore. These will include visual (colour, shape, size, brightness, darkness, depth, etc.), auditory (sounds of various degrees of complexity, decomposable into quantities such as pitch, timbre and the like), olfactory (newly mown grass, rotting fish, freshly baked bread, a paper mill, the sea, etc.), tactile (the feel of fur, velvet, cold steel, newly sanded wood, greasy hair, sand beneath one's toes) and gustatory (habanero sauce, ripe versus unripe apples, Hermitage La Chapelle 1988 versus my father's home-made wine, etc.) experiences.

The list could, obviously, be expanded indefinitely, both within each category and by the adding of new categories (emotions, imagery, conscious thought, etc.). But this is not necessary. One point is, perhaps, worth noting. There is often a tendency, particularly in the case of visual examples, to place undue emphasis on perceptually basic, or near basic, experiences: experiences of a patch of redness, and the like. But this, as Wittgenstein would put it, might provide a diet of philosophically one-sided examples. Often, the phenomenal character of an experience can depend on its *significance* for the experiencer, and this, at least ostensibly, cannot be reduced to the significance of a conglomeration of perceptually basic, or near basic, properties. I once saw Muhammad Ali at Nashville airport, and, believe me, this was an experience which very definitely had a phenomenal character, one which could not be reduced to the aggregation of significances of patches of colour, shape, contours, and the like. Nor is it clear that we *must* think of this as a combination of perceptual experience plus emotional response, with the richer phenomenal character lurking in the latter rather than the former. Or, if this strategy is available here, then it is not clear why it would not be available in the case of our experience of perceptually basic properties; and this would undermine the idea that visual experiences, as opposed to the emotional response they evoke, have a phenomenal character.

In any event, the idea that motivates these sorts of examples is simply that anyone who has had any of the above experiences will know that they *feel* or *seem* a certain way, that there is *something that it is like* to undergo them. This brings us to device no. 2.

Rough synonyms

The concept of phenomenal consciousness is sometimes explained, and I use the term loosely, by way of terms that are roughly synonymous with the original expression. Thus, phenomenally conscious states are ones which have, or are defined by, a *phenomenology*, which have a certain

qualitative feel or *qualitative character*. Such states are *experiential* ones, *subjective* ones. They are states that essentially possess *qualia*. Most importantly, perhaps, for any phenomenally conscious state, there is *something that it is like* to be, or to undergo, that state. 'Fundamentally', writes Thomas Nagel, 'an organism has conscious mental states if and only if there is something that it is like to be that organism – something it is like *for* the organism' (1974: 166).

Just do it

The third device embodies what we might call the *Nike*™ approach. Just do it. More precisely, one is invited to construct the circumstances that will produce in one states with a particular form of phenomenal consciousness. Sometimes, for example, one is invited to inflict mild bodily trauma on one's person to reacquaint oneself with the content of talk of phenomenal consciousness (Searle 1997: 97–9). The possibilities here are, of course, endless.

I think we would be advised to treat these devices with some suspicion, and some of the grounds for this will be examined more closely later on. Fundamentally, however, what seems to unite all three types of device is that they are, essentially, devices of *ostension*; they are means of *pointing*, or attempting to point, at phenomenal consciousness. And we are all familiar with the problematic status of attempts to point at private, inner, qualities, such as phenomenal consciousness purports, or is commonly taken, to be. So, the assumption that these devices are collectively sufficient to fix the meaning, or delineate the content, of the concept of phenomenal consciousness is far from certain. Indeed, this is precisely one of the assumptions that those who are sceptical of phenomenal consciousness will reject (see, for example, Dennett 1997: 117–18).

If the devices, even collectively, do not show that we know what we are talking about when we talk about phenomenal consciousness, they do show something much weaker, but something perhaps robust enough to provide a stepping-off point for further investigation. What the devices, or more importantly, the widespread presumed efficacy of the devices, do show is that a large number of people *think* they know what they are talking about when they talk about phenomenal consciousness. Indeed, I am one of those people. In fact, the people who explicitly deny that they know what they are talking about when they talk about phenomenal consciousness (and most of them do still talk about phenomenal consciousness, if only to deny the coherence of the concept) are, in all probability, limited to those antecedently *in the grip* of some quite specific theory of mind. A completely unscientific survey of some of my drinking

acquaintances, for example – who, I think they will not mind me saying, are very definitely *not* in the grip of some quite specific theory of the mind – indicates that they at least *seem* to have no difficulty in understanding what I am talking about when I talk about the what it is like of experience. Or perhaps they are just being polite. Or trying to shut me up.

In any event, that we, or most of us, *think* we know what we are talking about when we talk about phenomenal consciousness, even if we are mistaken in this thought, is the place where this book begins. This, then, is a book for all those who *think* they know what they are talking about when they talk about phenomenal consciousness. If the collection of devices outlined above is not sufficient to convince you that you at least think you know what you are talking about when you, or someone else, talks about phenomenal consciousness, then there is probably nothing in this book for you.

In fact, I labour our inability to define phenomenal consciousness, or to specify in any standard and perspicuous way the content of this concept, for a quite specific reason. *This is an essential datum that any account of consciousness should explain*. Our inability on this score is not something to be treated with embarrassment, swept under the carpet, lip-serviced, or mentioned at the outset and then forgotten. Rather, it is a feature of our understanding of the concept that any adequate account of consciousness should address and, hopefully, explain. Approaches that are, broadly speaking, eliminativist about phenomenal consciousness will explain this by saying that there is no coherent concept there to specify, or that what is there is a jumbled mish-mash of conceptually variegated strands that cannot be rendered into any coherent whole. While I am not convinced that such an explanation would work, even on its own terms, this book is, in any event, realist, not eliminativist, about phenomenal consciousness, and, as such, has no recourse to such strategy. The seeming ineffability of the concept of phenomenal consciousness imposes a fairly pressing requirement on realist accounts. If phenomenal consciousness is real, and if the corresponding concept is coherent, or reasonably so, then we should be able to *eff* it. And, if we cannot do this, then we have to come up with some explanation of why the concept of phenomenal consciousness cannot be *effed*.

2 The scope of 'There is . . . '

To say that an organism is conscious is, Nagel claims in his seminal (1974) paper 'What is it like to be a bat?', to say that 'there is something that it is like to be that organism – something it is like *for* the organism' (166).

And the claim that *there is something that it is like* to undergo a conscious experience is now one of the most common ways of explaining the idea that experiences, and the organisms that undergo them, are phenomenally conscious. The claim, however, is open to a variety of interpretation, some of which can, I think, be reduced to questions of the scope of the existential quantifier.

One obvious construal of Nagel's claim is that there is some object of conscious acquaintance and that all bats are acquainted with this object, while there is a distinct object of acquaintance such that all humans are acquainted with it. More generally, there is a certain form of consciousness that associates with being human, a distinct one that associates with being a bat, and so on. Indeed, it is possible to adopt an even broader conception of the what it is like of conscious experience. Flanagan (1992: 87), for example, claims that there is something that it is like to be conscious. And, again, one way of understanding this is as the claim that there is some object of conscious acquaintance and that all conscious creatures are acquainted with this object.

It is possible, however, to narrow considerably the scope of this claim. Thus, one might claim that what it is like is associated not with being conscious in general, nor with being a particular species of conscious organism, but, rather, with types of experience. One construal of this claim would entail that for every type of conscious experience there is some object of conscious acquaintance such that a creature which undergoes this type of experience is acquainted with that object. One might narrow the scope even further and claim that what it is like associates only with particular tokens of types of experience. On this view, for example, while there is no one thing that it is like to be in pain, there is something that it is like to suffer a particular token of pain. In an important, but strangely neglected, passage, Wittgenstein gestures towards the latter construal:

> Let us consider the experience of being guided, and ask ourselves: what does this experience consist in when for instance our *course* is guided? Imagine the following cases:
>
> You are in a playing field with your eyes bandaged, and someone leads you by the hand, sometimes left, sometimes right; you have to be constantly ready for the tug of his hand, and must also take care not to stumble when he gives an unexpected tug.
>
> Or again: someone leads you by the hand where you are unwilling to go, by force.
>
> Or: you are guided by a partner in a dance; you make yourself as receptive as possible, in order to guess his intention and obey the slightest pressure.
>
> Or: someone takes you for a walk; you are having a conversation; you go wherever he does.
>
> Or: you walk along a field track, simply following it . . .

'But being guided is surely a particular experience!' – The answer to this is: you are now *thinking* of a particular experience of being guided. (1953: #172–3)

There is no one thing that it is like to undergo the experience of being guided, but, rather, this what it is like fragments into the what it is like of particular (i.e. token) experiences of being guided.

There is, in fact, no straightforward inconsistency between the view that the what it is like attaches, in the first instance, to experiential tokens, and Nagel's claim that there is something that it is like to be a bat (or human). There are at least two ways of rendering these claims consistent, one in terms of the idea of set membership, the other which appeals to higher-order properties of what it is like. According to the first strategy, to say that what it is like to be a bat is different from what it is like to be a human is to say (i) that for each (actual or possible) bat experience-token there is an associated what it is like, and for each (actual or possible) human experience-token there is an associated what it is like, but either (ii) the set of bat what it is likes does not overlap with the set of human what it is likes or (iii) the overlap between the two sets falls below a certain threshold. According to the second strategy, the what it is likes of bat experience-tokens instantiate a certain essential higher-order property B, while the what it is likes of human experience-tokens instantiate a certain higher-order property H, and B is distinct from H. That is, what it is like instantiates various higher-order properties, properties which vary from human to bat. On this view, what it is like attaches primarily to mental tokens and derivatively (in virtue of its higher-order properties) to organisms.

The claim that the what it is like of conscious experience attaches primarily either to experience-tokens (or to experience-types), however, does give rise to the following, more radical, possibility. The claim that there is something that it is like to undergo a token of one experience-type, say pain, might mean something distinct from the claim that there is something that it is like to undergo a token of a different type of mental state, for example, to token-instantiate (occurrently) the belief that Ouagadougou is the capital of Burkina Faso. That is, it cannot be assumed at the outset that consciousness is a unitary property that attaches uniformly across all mental states.

The suspicion that it is not such a property can, in fact, be independently motivated by the following, well-known, considerations. Consider, first, the distinction between sensations and propositional attitudes. Propositional attitudes can certainly be associated with a phenomenology. There can be, in a given instance, something that it is like to have, say, a certain belief. However, propositional attitudes, it is commonly thought,

are not *defined* by a phenomenology, and their possession by a subject does not entail that this subject is presented with any phenomenology at all, let alone a particular phenomenology. However, this does not seem to be the case with at least some sensations. While, if Wittgenstein is correct, the phenomenology associated with an experience E may vary from one token of E to another, it seems that having some phenomenology or other, and indeed having a phenomenology constrained within certain reasonably definite limits, is essential to the tokening of at least some, and perhaps all, sensations. Even within the category of sensations there appear to be important differences. It is not only common, but also seemingly perfectly appropriate, to characterise the phenomenology of bodily sensations – pains, itches, orgasms, and so on – in terms of the notion of *feel*. With items such as perceptual experiences, however, the characterisation of their phenomenology in terms of the notion of feel sits a lot less comfortably. This is why the epithet 'feels' is, in the case of perceptual experiences, typically replaced by 'seems'. If we do want to say that it *feels* a certain way to see a green wall, or Muhammad Ali, then it is far from clear that *feel* means the same thing in this context as it does in the case of sensations. But, of course, *feel* is often used as an alternative appellation for the what it is like of conscious experience, sensational, perceptual or otherwise. To say that there is something that it is like to undergo a conscious experience is often taken as equivalent to saying that having that experience feels a certain way. And if this is correct, then we cannot assume, *a priori*, that the existential quantifier in the claim 'There is something that it is like to undergo X' ranges across the same quantity for all Xs.

Therefore, we should be alive to the possibility that what it means for a mental state to be phenomenally conscious can vary from one category of mental state to another, perhaps from one type of mental state to another, perhaps even from one token mental state to another. Perhaps the concept of phenomenal consciousness is a fundamentally hybrid concept.[1] And, if this is so, we would look in vain for a unified account of in what phenomenal consciousness consists. At the very least, this is not something to be ruled out *a priori*.

In later chapters, when the real argument starts, I propose to avoid these potential difficulties by focusing on, and working with, certain very general features that any instances of phenomenal consciousness must,

[1] Of course, many have claimed that the concept of consciousness is a hybrid one. What they typically have in mind, roughly, is the idea that consciousness comes in many forms: phenomenal, introspective, self, monitoring, reportability, etc., etc. The present point, however, concerns only the category of phenomenal consciousness, and the possibility being mooted is that this is itself a hybrid category.

I shall argue, possess. Whether or not phenomenal consciousness turns out to be a conceptually or theoretically unified item, I shall try to show that anything that could possibly count as an instance of a phenomenally conscious state must have certain features, and it is upon these features that the arguments will be built.

3 What is the problem of phenomenal consciousness?

The above problems, unclarities, and cautionary notes notwithstanding, we perhaps (hopefully) have enough in the way of a preliminary characterisation of the concept of phenomenal consciousness to proceed to a preliminary (again) characterisation of the problem or problems it raises. Phenomenal consciousness is widely, though far from universally, accepted to create at least the appearance of a problem for materialism. Agreement on precisely what this problem is, or appears to be, however, is far less widespread. The intuition that there is at least the semblance of a problem, here, is commonly supported by the way of various intuition pumps.

1 Abused scientists

Mary has been forced to live her entire life in a black and white room and has never seen any colours before, except for black, white, and shades of grey (Jackson 1982, 1986). Filling in the details would be a rather fatuous exercise, but presumably her skin has also been treated with some pigment that makes it appear a shade of grey, which pigment has also transformed her irises appropriately, her hair has been dyed black, etc., etc. Despite her dysfunctional upbringing, Mary has become the world's leading neuroscientist, specialising in the neurophysiology of colour vision. She knows everything there is to know about the neural processes involved in the processing of visual information, about the psychophysics of optical processes, about the physics of environmental objects, and so on. However, despite this extensive knowledge, when she is let out of her black and white room for the first time, it seems plausible to suppose, she learns something new; she learns what it is like to experience colour. And, if this is correct, then this knowledge is neither something she possessed before nor something that could be constructed from the knowledge she possessed before.

2 Zombies

A zombie, in the philosophical as opposed to the Hollywood sense, is an individual that is physically and functionally human, but which lacks

conscious experience (Chalmers 1996; Kirk 1974, 1994). Thus, my zombie twin is physically identical to me and, we can suppose, is embedded in an identical environment. Moreover, he is functionally identical to me in that he is processing information in the same way, reacting in the same way as me to the same inputs, and so on. Nevertheless, he lacks phenomenal experience; he has no phenomenal consciousness. My zombie twin is not, it is generally accepted, a natural possibility (that is, he is incompatible with the laws of nature) but he is, it has been argued, a logical possibility.

3 *Deviants*

It is logically possible for there to be a world where qualia are inverted relative to the actual world (Shoemaker 1982; Chalmers 1996). My inverted twin is physically identical to me but has inverted conscious experiences. Thus, for example, where I have a red experience (i.e. an experience *as* of red) my inverted twin has a green experience (i.e. an experience *as* of green). That is, when he looks at a fire engine, he has an experience of the same qualitative colour character as I do when I look at grass. Again, my inverted twin may not be a natural possibility, but he is, it has been argued, a logical possibility.

4 *Demons*

Laplace's Demon is able to read off all non-basic facts from basic ones (Chalmers 1996). That is, the Demon knows every detail about the physics of the universe, the configuration and evolution of all the basic fields and particles that make up the spatiotemporal manifold. And from this knowledge, the Demon can read off, or infer, every other fact about the universe. Or, rather, *almost* every other fact. For, it has been argued, the Demon would not be able to read off facts about conscious experience (Chalmers 1996). Indeed, the Demon could not even work out, from its knowledge of the basic facts alone, that there is any conscious experience at all, let alone what it is.

A motley crew. Surely, it is only in recent discussions of consciousness – and perhaps some fairly questionable B-movies – that one could possibly find such a collection of characters. But the question is: what does all this mean? And this is a good question, one that subsequent chapters will spend some time trying to work out, and one that as yet has nothing even close to an accepted answer.

However, it is possible to broadly identify two axes along which potential answers may be developed. On the one hand, one can understand

the examples as establishing, or suggesting, an ontological or *metaphysical* conclusion that is, essentially, dualistic in character. Phenomenal experiences are distinct from, and not reducible to, any physical event, state or process. This conclusion is (or has at one time been) endorsed, on the basis of one or more of the above scenarios, by Jackson (1982, 1986) and Chalmers (1996). On the other hand, one can understand the examples as establishing, or suggesting, an *epistemological* conclusion. Roughly speaking, our knowledge of physical facts does not, in some way, add up to knowledge of conscious experience, and, consequently (perhaps) physical explanations do not, in some way, add up to explanations of consciousness. There is, as it is often put, an *explanatory gap* between consciousness and the physical. This conclusion has been endorsed by Levine (1983, 1993) and McGinn (1989, 1991, 1993) among others. Of course, those who endorse the metaphysical conclusion are also going to endorse the epistemological claim, and this is the case with Jackson and Chalmers. However, it is possible to endorse the epistemological claim alone.

In fact, there are, in my view, good reasons for endorsing the epistemological claim alone. All the above examples turn, ultimately, on a difference between phenomenal and physical *concepts*, and it is difficult to turn this into any substantive difference between phenomenal and physical *properties*. But it is the latter difference that is required to underwrite the metaphysical conclusion.

To see this, consider the knowledge argument. There are, in fact, various strategies available to the materialist should she want to resist the metaphysical interpretation of the significance of the knowledge argument. The one I favour is due to Brian Loar (1990). According to Loar, the materialist can allow that Mary acquires new information when she leaves the room, but she does so only under an *opaque* reading. Transparent construals of the information acquired by Mary would, in effect, beg the question against materialism. Drawing (legitimate) metaphysical conclusions from opaque contexts is never easy. And, given the opaque construal of what Mary learns, we can construct *prima facie* analogous cases, where a metaphysical conclusion manifestly does not follow from the premises. Thus, to borrow from Loar, Kate learns that the bottle before her contains CH_3CH_2OH. But, on an opaque reading, she does not know that the bottle contains alcohol. That is, she does not know that the bottle contains stuff called alcohol, or that the bottle contains the intoxicating component of wine and beer, the component that makes people drunk. Indeed, we can suppose that innocent Kate even lacks the ordinary concept of alcohol. Then, when she inadvisedly consumes the bottle's contents, she acquires new information: that the bottle contains