The Nature of Consciousness
This page intentionally left blank
In *The Nature of Consciousness*, Mark Rowlands develops an innovative and radical account of the nature of phenomenal consciousness, one that has significant consequences for attempts to find a place for it in the natural order. The most significant feature of consciousness is its dual nature: consciousness can be both the directing of awareness and that upon which awareness is directed. Rowlands offers a clear and philosophically insightful discussion of the main positions in this fast-moving debate, and argues that the phenomenal aspects of conscious experience are aspects that exist only in the directing of experience towards non-phenomenal objects, a theory that undermines reductive attempts to explain consciousness in terms of what is not conscious. His book will be of interest to a wide range of readers in the philosophy of mind and language, psychology, and cognitive science.

The Nature of Consciousness

Mark Rowlands

University College, Cork
## Contents

**Preface**

<table>
<thead>
<tr>
<th>1 The problem of phenomenal consciousness</th>
<th>page ix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 What is phenomenal consciousness?</td>
<td>1</td>
</tr>
<tr>
<td>2 The scope of ‘There is . . .’</td>
<td>4</td>
</tr>
<tr>
<td>3 What is the problem of phenomenal consciousness?</td>
<td>8</td>
</tr>
<tr>
<td>4 Explaining consciousness</td>
<td>13</td>
</tr>
<tr>
<td>5 Vertical strategies I: the mind–body problem</td>
<td>14</td>
</tr>
<tr>
<td>6 Vertical Strategies II: the mind–mind problem</td>
<td>16</td>
</tr>
<tr>
<td>7 Horizontal strategies</td>
<td>22</td>
</tr>
<tr>
<td>8 The shape of things to come</td>
<td>23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2 Consciousness and supervenience</th>
<th>26</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Logical supervenience: ontological and epistemological interpretations</td>
<td>27</td>
</tr>
<tr>
<td>2 (Merely) natural supervenience</td>
<td>28</td>
</tr>
<tr>
<td>3 The incoherence of (Chalmers’ versions of) supervenience</td>
<td>29</td>
</tr>
<tr>
<td>4 Natural supervenience and weak supervenience</td>
<td>35</td>
</tr>
<tr>
<td>5 Natural supervenience as an epistemological concept</td>
<td>37</td>
</tr>
<tr>
<td>6 More on ‘reading off’</td>
<td>41</td>
</tr>
<tr>
<td>7 Logical supervenience and reductive explanation</td>
<td>48</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3 The explanatory gap</th>
<th>51</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Intuitions and arguments</td>
<td>52</td>
</tr>
<tr>
<td>2 Analysing the intuition</td>
<td>57</td>
</tr>
<tr>
<td>3 Truth and adequacy</td>
<td>58</td>
</tr>
<tr>
<td>4 Explanatory adequacy and epistemic satisfaction</td>
<td>59</td>
</tr>
<tr>
<td>5 Proto-epistemic satisfaction</td>
<td>64</td>
</tr>
<tr>
<td>6 Mechanistic explanations and correlations</td>
<td>68</td>
</tr>
<tr>
<td>7 Explaining consciousness</td>
<td>70</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4 Consciousness and higher-order experience</th>
<th>75</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 HOR models of consciousness</td>
<td>75</td>
</tr>
<tr>
<td>2 The structure of HOP theories</td>
<td>79</td>
</tr>
<tr>
<td>3 Presuppositions of the HOP model</td>
<td>81</td>
</tr>
<tr>
<td>4 The independence condition</td>
<td>83</td>
</tr>
<tr>
<td>5 The explanatory primacy of vehicles</td>
<td>94</td>
</tr>
</tbody>
</table>
vi Contents

6 The primacy of transitive consciousness 97
7 What has gone wrong? 98

5 Consciousness and higher-order thoughts 101
1 HOT models 103
2 The problem of circularity 105
3 The problem of regress 113

6 The structure of consciousness 122
1 Introduction 122
2 Consciousness as object of consciousness: empirical apperception 123
3 Transcendental apperception 126
4 Consciousness as experiential act 133
5 What it is like 136
6 The ubiquity of objectualism 141
7 Summary 147

7 What it is like 148
1 Against objectualism 149
2 What it is like as a phenomenal particular 150
3 What it is like as a phenomenal property 152
4 What it is like as a representational property 157
5 For actualism 166
6 Phenomenology by proxy 168
7 Objections and replies 173
8 Summary 176

8 Against objectualism II: mistakes about the way things seem 178
1 Introduction 178
2 Three mistakes about experience 179
3 The objectualist gloss: qualia 181
4 Perceptual completion and neural filling in 183
5 Dennett’s criticism of filling in 186
6 Change blindness and the richness of experience 187
7 Category (2) mistakes: how an experience seems and how it really is 189
8 Mistakes of category (3) 191
9 Why the way an experience seems cannot be explained as awareness of qualia 194

9 Consciousness and representation 197
1 Brentano’s thesis 198
2 Consciousness as revealing and as revealed 201
3 Phenomenal revealing 204
4 Consciousness of and consciousness that 208
5 Representationism 209
6 Object representationism 211
7 Mode representationism 213
8 Actualism and representationism 214
10 Consciousness and the natural order 216
  1 What it is like and reductive explanation 216
  2 Consciousness and materialism 219
  3 Consciousness and causality 221
  4 The epiphenomenalist suspicion 222
  5 The standard problem of epiphenomenalism 226
  6 The epiphenomenalist suspicion allayed 232

Bibliography 236
Index 242
Preface

Colin McGinn first got me thinking about consciousness. I was finishing up a D.Phil. at Oxford, where Colin was my supervisor. He had just thought up the basic line of argument behind ‘Can we solve the mind–body problem?’, and I may have been one of the first people he explained it to. I thought he was mad! A decade or so later, when I returned to look at his work, I was struck by how sane the old man had become in the intervening years. Also, much to my chagrin, I was struck by how much my own developing position owed to his. Somewhat in this spirit of chagrin, then, I did my best to distinguish my view from his, and this resulted in chapter 3.

My thinking on the nature of supervenience, and, in particular, on the distinction between ontological and epistemological interpretations, has been profoundly influenced by the work of John Post, as anyone who has read his Faces of Existence – a work of the highest quality – will know. The influence of Sydney Shoemaker will also be evident in many of the pages that follow.

An earlier version of chapter 5 appeared in Mind and Language, as ‘Consciousness and higher-order thoughts’. I am grateful to Sam Guttenplan, Editor of the journal, and to Blackwell publishers for permission to use this work.

Thanks to Hilary Gaskin at Cambridge University Press. Colin Allen, in his capacity as reader for Cambridge University Press, made several helpful suggestions. My thanks to him. And thanks to Joanne Hill for some outstanding copy-editing.

This work was supported by a grant from the Faculty of Arts Research Grants Committee at University College, Cork. My thanks. Thanks also to my colleagues in the Department of Philosophy at University College, Cork who have helped foster a very pleasant working environment, and to Des Clarke, whose creative approach towards my current leave of absence greatly facilitated the completion of this work.
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
The Nature of Consciousness

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 phenomenol-
yogy 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 cate-
gory of mental state to another, perhaps from one type of mental state
to another, perhaps even from one token mental state to another. Per-
haps 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$_3$CH$_2$OH. 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
The problem of phenomenal consciousness

alcohol. If the knowledge argument, on the metaphysical construal, had a generally valid form, we could then infer from Kate’s epistemic situation that alcohol is not identical with CH₃CH₂OH. And this, evidently, does not follow.

What seems to be going on here is that we have two distinct concepts associated with the same substance; one a theoretical-physical concept, the other what Loar calls a recognitional concept. The substance alcohol can be picked out both by way of theoretical description, and in terms of the properties by which one typically recognises it. However, the two types of concept are conceptually independent of each other, and this explains both why the above opaque reading of what Kate learns is possible and why this opaque reading does not yield a substantive metaphysical conclusion.

A recognitional concept has the form ‘x is one of that kind’; i.e. they are type-demonstratives grounded in dispositions to classify, by way of perceptual discriminations, certain objects, events, and situations. Recognitional concepts, crucially, are typically conceptually independent of, and irreducible to, theoretical-physical concepts, even where both concepts, as in the above case, pick out the same property.

Loar argues that phenomenal concepts are essentially recognitional in character. Thus, materialism at the metaphysical level is underwritten by the claim that phenomenal and physical-functional concepts can pick out the same property, while the conceptual independence of these concepts is explained by the fact that recognitional and theoretical concepts are, in general, conceptually independent, and that the former cannot be reduced to the latter. Thus the epistemological reading of the knowledge argument is safeguarded and explained, and the metaphysical reading shown to be invalid.

This, of course, takes us only part of the way. It is not difficult to find a difference between the case of Kate and the case of Mary. Kate lacks knowledge of the contents of the bottle under a contingent description of it: stuff that gets you drunk. However, Mary’s acquired information of what it is like to experience colour does not conceive it under a contingent mode of presentation. It is not as if she is conceiving of a property that presents itself contingently thus: it is like such and such to experience P. Being experienced in this way is essential to the property Mary conceives. Thus, when Mary later acquires new information (construed opaquely) the novelty of this information cannot be explained – as in the case of Kate – as her acquiring a new contingent mode of presentation of something she has known all along. This is why, according to its proponents, the knowledge argument can be valid on an opaque reading. There is no contingency in Mary’s conception of the new phenomenal information
that explains it as a novel take on old facts. Therefore, we must suppose that she learns new facts simpliciter, and not new conceptions of old facts.

As Loar points out, however, there is an implicit assumption in this argument: a statement of property identity that links conceptually independent properties is true only if at least one concept picks out its associated property by way of a contingent mode of presentation of that property. Conversely, the underlying idea is that if two concepts both pick out the same property by way of its essential properties, neither mediated by contingent modes of presentation, then one ought to be able to see a priori – at least after optimal reflection – that they pick out the same property. If the two concepts pick out the same property by way of essential modes of presentation, then those concepts themselves must be logically connected.

However, Loar argues, convincingly in my view, that this assumption should be rejected. It rests on the idea that (i) if a concept picks out a property by way of an essential mode of presentation, then that concept must capture the essence of the property picked out, and (ii) if two concepts capture the essence of the same property, then there must exist constitutive conceptual connections between those concepts, such that one concept is derivable from the other a priori. However, when expressed in this way, it is fairly clear that these are equivocating uses of ‘capture the essence of’. On one use, it expresses a referential notion that comes to no more than ‘directly rigidly designate’. On the other, it means something like ‘be conceptually interderivable with some theoretical predicate that reveals the internal structure of’ the designated property. But the former does not imply the latter. Claims about rigid designation do not, in general, imply the conceptual interderivability of the designating concepts.

Once we allow that phenomenal and physical concepts can both (i) pick out a property by way of an essential mode of presentation, but (ii) still be conceptually independent of each other, then essentially the same deflationary strategy can be adopted with respect to the rest of the assorted cast listed above. The logical possibility of zombies, that is, need only be taken as indicative of the conceptual independence of phenomenal and physical-functional concepts, and not of any deeper metaphysical division. A similar account will be applicable to the case of the qualia-inverted deviants; their logical, as opposed to natural, possibility, need be indicative only of the logical independence of phenomenal from physico-functional concepts. And the failure of Laplace’s Demon to read off phenomenal facts from non-phenomenal ones, again, need only indicate the conceptual independence of phenomenal concepts from physical or functional ones.
Loar’s account, of course, will not satisfy everyone. Indeed, despite my general sympathy to this line of reasoning, I think that Loar’s claim that phenomenal concepts are recognitional ones needs to be severely qualified (see chapter 7). Nevertheless, I suspect that a story substantially similar to the one Loar tells can be made to work. And, for this reason, I am going to treat the problem of phenomenal consciousness as a primarily epistemological one. This may be incorrect. Perhaps consciousness provides a metaphysical problem also. If so, then so be it. If there is a genuine metaphysical problem, then it is outside the scope of this book.

The book’s subject is the epistemological problem posed for materialism by phenomenal consciousness: the existence, or apparent existence, of an explanatory gap between the phenomenal and the material. One thing is clear: if consciousness is not an epistemological problem, then it is not a metaphysical problem either.

4 Explaining consciousness

The problem of explaining phenomenal consciousness is the problem of explaining how consciousness can come from what is not conscious. And one can understand the idea of consciousness coming from what is not conscious either causally or constitutively. For various reasons I prefer the constitutive construal. Suppose, for example, we say that phenomenal consciousness is causally produced by brain activity (McGinn 1989, 1991; Searle 1992). Causal relations, as Hume taught us, involve distinct existences. So, if we talk of consciousness being causally produced by neural activity then there is a danger that we have already implicitly bought in to a metaphysical understanding of the problem: we have already implicitly assumed that consciousness is distinct from this neural activity. We can avoid this metaphysical temptation by regarding the causal relations by which consciousness is produced as diachronic, rather than synchronic. But then the production of consciousness by the brain has to be understood in terms of the idea that a phenomenal property instantiated at time t is produced by brain activity occurring at t−1. But this does not seem to be the correct model for understanding the production of consciousness by neural activity. What neural activity occurring at time t−1 actually causally produces is neural activity occurring at t. And then we still have the problem of explaining how consciousness is produced by this neural activity of time t. If we want to insist that this relation of production is a causal relation, then we fall right back into the metaphysical construal of the problem.

Intuitively, the relation of production we require seems to be more like the relation between the observable properties of water and its underlying
structure, and this (pace Searle 1992) is not a causal relation. Rather, the observable properties of water are, in some sense, constituted by the underlying molecular properties. Phenomenal consciousness, on this construal, is somehow constituted by neural activity, and the problem of consciousness is the problem of explaining how this could (possibly) be so. More generally, how can consciousness be constituted by what is not conscious?

While, for these and other reasons, I favour the constitutive rather than causal construal of the claim that consciousness is produced by what is not conscious, nothing much turns on this assumption. The arguments to be developed in the following chapters have, I think, an application broad enough to cover both constitutive and causal senses, and, accordingly, I shall usually employ the more general term production to subsume both constitutive and causal senses of the relation between consciousness and the material.

The focus of this book, then, is whether it is possible to provide an explanation of how phenomenal consciousness is produced by what is not conscious. We know, I shall suppose, that it is, in fact, produced by what is not conscious, that is why we are not concerned with the metaphysical construal of the problem. What we want is an explanation of how it is so produced.

In attempting to provide an explanation of phenomenal consciousness, it is possible to adopt two quite distinct strategies; one, as I shall put it, vertical, the other horizontal. Vertical strategies, roughly speaking, attempt to build consciousness up from what is not conscious. Horizontal strategies, again roughly, attempt the explanatory task by attempting to pull consciousness out into what is not conscious, i.e. the world. The next two sections deal with the former type of explanatory strategy, the one after that deals with the latter.

5 **Vertical strategies I: the mind–body problem**

To build consciousness up from what is not conscious is to show how various non-conscious processes can, collectively, constitute conscious activity. This strategy of, as we might call it, phenomenal tectonics, of constructing the phenomenal from the non-phenomenal, divides into two distinct approaches. On the one hand, we can try to build consciousness up from processes that are neither conscious nor mental. Our appeal, here, is likely to be to the brain, to neural activity broadly construed. On the other, we can try to construct consciousness from processes that are non-conscious but which are mental. The explanation here is likely to involve, quite centrally, higher-order mental states of some sort, states which are identified as not being essentially conscious, or, at the very