

## I Assumptions, distinctions, and a map

---

The nature and aims of my project have already been explained in the Preface. In this opening chapter I shall lay out some of my background assumptions, introduce a number of important distinctions, and outline the direction which the discussions of later chapters will follow.

### I Physicalism and naturalism

In this section I shall briefly explain and defend two default assumptions, which form the background to the problem of phenomenal consciousness. It is these assumptions which appear to be challenged by the very existence of phenomenal consciousness, as we shall see in chapters 2 and 3.

#### 1.1 *Physicalism*

One assumption I shall make is that we should at least *try* to be token-physicalists about the mind. We should maintain that all particular (or ‘token’) mental states and events are at the same time physical (presumably neurophysiological) states and events, if we can do so consistently with our other beliefs. In the present section I shall briefly motivate this assumption, which is shared by almost everyone now working in the philosophy of mind – which is not to say that physicalism itself is mandatory, of course; indeed, many of the arguments against physicalism derive from considerations to do with phenomenal consciousness, as we shall see.<sup>1</sup>

<sup>1</sup> There are many who would deny the claim that mental states and events are *neurophysiological* states and events, not because they reject physicalism, but because they endorse an *externalist* account of the individuation-conditions of mental states with intentional content, such as beliefs and desires (e.g. Burge, 1979, 1986a, 1986b; McDowell, 1986, 1994). On such accounts, the identity of a mental state is tied up with the identity and existence of the worldly objects and properties which that state is *about*. I shall ignore such views here, for simplicity only. The basic argument for physicalism can still go through, only with the complication that the mental cause of a bodily movement is a complex relational entity, involving both the brain state which is the immediate physical cause of the movement and relations to the items in the world which that mental state is about. The distinction between externalist and internalist (or between wide and narrow) accounts of intentional content will become important in chapters 3 and 4.

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

## 2 I Assumptions, distinctions, and a map

Descartes famously held that the mind is non-physical while the brain is physical, and that they interact causally with one another. For example, sensory stimulation causes conscious experience, and decisions cause bodily movements. One of the main objections to dualism ever since has been the difficulty of making sense of such a causal connection. Not that there is any problem of *principle* in understanding causal connections between physical and non-physical, in my view. For there is nothing in the concept of causation, as such, which requires all causes to be mediated by physical mechanisms. The real problem is to understand *how* such causation can occur, given what we already know about the physical world, and about causation in the brain.

Consider, first, the physical world in general. Most scientists now believe that physics is *closed*, in the sense of permitting no interference from, or causation by, events at higher levels of description (e.g. chemical or biological). On this view all atomic and sub-atomic events happen in accordance with physical laws (albeit probabilistic ones), and all events at higher, more abstract, levels of description must be realised in, or constituted by, those physical processes, in such a way as to allow no independent point of causal leverage. So while there may be chemical and biological laws, the events which figure in these laws must always, at the same time, fall under the laws of physics. On this picture there is simply no *room* for a distinct and independent psychological level, whose events are not physically constituted, but which can have an impact upon the physical behaviour of the body.

Consider, now, what is known about the brain. There is much still to learn, no doubt – about the functions and interactions of its parts, for example. But much is already known. It is known that the brain consists of nerve cells, of various known types. And much is known about how such cells function, and the physical causes which lead to their activity. Certainly there would appear to be no ‘inverse causal black-holes’ in the brain, such as would seem to be required by the interactionist picture (that is, there are no places from which brain activity emerges *for no physical reason*). Indeed, enough is already known about the brain to justify the claim that each event in the brain has a sufficient physical cause. So, again, the moral would appear to be that there is no room, here, for mental events to cause physical ones, unless those mental events are themselves physically constituted – that is to say, unless physicalism is true.

What are the alternatives to physicalism? One possibility would be to go for some sort of *panpsychism* (Nagel, 1979), believing that current descriptions of physical reality are inadequate, and that all physical events are in some sense already mental ones, or possess mental properties.

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

Another possibility would be to exploit the indeterminacies left open by physical theory at the sub-atomic level, to find a place for mental–physical interaction. It could be maintained, for example, that the mind somehow resolves all the sub-atomic indeterminacies which exist within the neurological events in our brains in one direction or another, thereby having an influence upon the overall patterns of activity in the brain (Penrose, 1994). Yet another alternative would be to embrace *epiphenomenalism* about the mental in general, or about phenomenal consciousness in particular, believing that conscious experiences are not physically constituted, and that while being caused by physical events in our brains, they can have no further physical effects in their turn (Jackson, 1982).

None of these alternatives to physicalism is at all attractive. For example, in connection with the last, there are real problems in explaining how we can know that we ourselves are phenomenally conscious, at least if it is allowed that intentional mental states like beliefs have a physical constitution (as does Chalmers, 1996). For then, by hypothesis, our belief that we enjoy experiences with *feel* will not be a product of those experiences themselves (but rather, at best, will be caused by the physical events which themselves cause such experiences), and would have occurred just the same even if brain events had *not* caused experiences.<sup>2</sup>

Alternatively, if the thesis of non-physicality is extended to intentional mental events as well as to phenomenally conscious ones, then our problem is to explain our knowledge of the mental states of others. For while our belief that we ourselves are phenomenally conscious may be caused by the presence of conscious experience, the mental states of others (and of ourselves) can have no causal impact upon behaviour. So even while someone is describing in technicolour detail how it feels to them to be undergoing a certain sort of experience, their behaviour provides no real evidence of the presence of such experience; for by hypothesis, they would have behaved just the same even if brain events hadn't given rise to mental events at all.

This is not the place to develop these and other objections to the various alternatives to physicalism in any detail. For my purpose here is just to remind the reader of the considerations which make physicalism the default option in the philosophy of mind. Unless there are very powerful arguments to the contrary, we should believe that all mental states and events are physically constituted. Most philosophers think that the strongest challenge in this regard is provided by phenomenal consciousness

<sup>2</sup> See section 3 below and chapter 5:3 for discussion of the distinction between intentional states such as beliefs and thoughts, on the one hand, and experiences on the other.

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

## 4            1 Assumptions, distinctions, and a map

itself. Just how powerful this challenge really is will be considered in chapters 2 and 3, where it will be suggested that all the main anti-physicalist arguments commit fallacies of one sort or another.

1.2        *Naturalism*

Naturalism is the belief that all of the events and processes which occur in the world are natural ones, happening in accordance with causal laws. So there are no miracles, and everything which happens can in principle be provided with a causal explanation, or is subsumable under laws (albeit probabilistic ones). In addition, naturalism is normally construed as involving the idea that the different levels of causation in nature are *ordered*, in such a way that processes at higher levels are always realised in, and reductively explicable in terms of, those at the lower levels. This need not mean that all *properties*, or *types* of phenomena, are identical with types identifiable in terms of physics, since higher-level types (e.g. wings) may be multiply-realised in lower-level processes or structures (as in the differences between the wings of birds and the wings of bats). But it does mean that all higher-level properties should be physically *constituted*, in such a way that each instantiation of such a high-level property admits of reductive explanation into lower-level (ultimately physical) terms.<sup>3</sup>

These have been the guiding methodological assumptions of science. When puzzling events occur, scientists do not just accept them, and postulate a miracle. Rather they continue to probe and investigate, working on the assumption that there must be a causal explanation, if only they could discover it. And when scientists discover laws and law-like relationships in nature, they do not rest content with a heterogeneity of such laws. Rather, they assume that nature constitutes a *unity*, and they seek to understand the operations of some laws in terms of others.

Since these naturalistic assumptions have received ample – albeit not conclusive – vindication through the advancement of science, it should require some powerful considerations to overturn them in connection

<sup>3</sup> Not *everyone* accepts that naturalism must involve a commitment to the reductive explicability of higher-level phenomena into lower-level terms. Thus Chalmers (1996), for example, describes his dualist account of consciousness as ‘naturalistic’ – since he believes that the properties involved in consciousness are subject to natural law, and are linked with brain-events by basic natural laws – although he thinks that phenomenal consciousness cannot be reductively explained. However, this position, if correct, would be highly revisionary of our scientific world-view. The conception of nature as *unified* – in a way that requires commitment to the possibility of reductive explanation – is so deeply built into scientific methodology that it surely deserves incorporation into our understanding of naturalism. At any rate, this is what I shall assume in what follows (nothing substantive hangs on it – the point is merely terminological).

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

with the mind and mental phenomena. Our default assumption should therefore be that all mental events occur in accordance with causal laws, and that we may hope to explain both the operation of, and the properties involved in, those laws in lower-level (ultimately physical) terms. However, precisely what naturalism commits us to is important to get right; especially since the project of this book is to naturalise phenomenal consciousness. I shall return to the issue in more detail in the chapters which follow, especially in chapter 4. (But even in chapter 4 my discussion will be relatively superficial, digging just deep enough into the issues to serve my own explanatory purposes. For a much more extensive and sophisticated treatment, see Papineau, 1993.)

## 2 Functionalism and theory-theory

The assumptions in section 1 above relate to the metaphysics of the mind. In this section I shall say something about how I take the mind to be conceptualised, or conceived of. I shall be assuming that some form of functionalism provides the best account of the way in which we conceptualise mental states. Again the position is not entirely mandatory, and again some of the main challenges come from considerations having to do with phenomenal consciousness, as we shall see. But the advantages of functionalism as an account of the mind (*viz.* its metaphysical neutrality – hence allowing interactive dualism to be a conceptual possibility – and its solution to the problem of other minds) mean that it should not be given up lightly.

### 2.1 *From Cartesian concepts to analytic functionalism*

As I have just noted, the thesis under discussion in section 1 was metaphysical – it concerned what mental states themselves really *are*. But what of our mental-state *concepts*? Even if mental states turn out to be physical, that does not seem to be how we conceptualise them – Cartesian dualism is a conceptual possibility, at least, even if it is actually false.

The thesis often attributed to Descartes is that mental-state concepts are (at least at bottom) bare recognitional capacities – capacities to recognise the distinctive *feel* which our mental states have. More recently, Goldman (1993) has defended a version of this view – claiming that we know of our own mental states by direct recognition, attributing the feelings in question to others by a generally-reliable process of *simulation*.<sup>4</sup>

<sup>4</sup> For an extended critique of the simulationist position, see my 1996b, and Botterill and Carruthers, 1999, ch. 4.

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

## 6 I Assumptions, distinctions, and a map

While I shall accept (indeed urge) that *some* of our mental-state concepts are Cartesian in this sense – consisting in bare recognitional capacities for the subjective *feels* of experience – I shall argue in chapter 9:3 that such concepts are parasitic upon those which are more theoretically embedded. And there are a number of powerful arguments against any attempt to extend the Cartesian view to *all* mental-state concepts. The main ones are as follows:

- (1) The Cartesian view makes it difficult to see how the idea of *non-conscious* mental states – or states which would lack any distinctive subjective feel – is even so much as a conceptual possibility. (See section 3 below, and also chapter 6.)
- (2) There are many conscious mental states which seem to lack distinctive feels – for example, beliefs and abstract (as opposed to bodily) desires. Perhaps it may be replied that these states are *dispositions* – dispositions to engage in acts of *thinking*, which have felt properties. But even if (many) acts of thinking do have felt properties (by figuring in ‘inner speech’, say), they do not seem to be *conceptualised in terms of* those properties. And the idea of ‘purely propositional’ (unfelt) thinking does seem to be a conceptual possibility; indeed many people believe it to be actual.<sup>5</sup>
- (3) The Cartesian view makes it difficult to see how we could ever acquire the rich causal knowledge which we manifestly do have concerning the operations of minds. Compare sense-data theory as an account of vision, which is the idea that we begin with capacities to recognise unstructured sense-data (such as colours and textures) and then build up to a complex causal representation of the world by a process of learning. No one thinks that this is a viable developmental story any longer in the case of vision; nor should they in the case of our common-sense understanding of the mind.

I shall assume, therefore, that the Cartesian view of mental-state concepts, when put forward as the basis of all mental-state understanding or

<sup>5</sup> See the results of Hurlburt’s (1990, 1993) introspection-sampling studies. Subjects wore a modified paging device through the day, which delivered a beep via an ear-phone at irregular intervals. Subjects were instructed to ‘freeze’ the contents of their conscious awareness at the moment of the beep, and to make brief notes to be reported to the experimenters later. All normal (as opposed to schizophrenic) subjects reported instances of ‘inner speech’, in varying proportions; with most also reporting visual images and emotional feelings. Many also reported the occurrence of ‘purely propositional’ (wordless) thoughts. In my 1998c I argue that there may actually be no such thing as purely propositional conscious thought, and that these reports may really be the result of swift self-interpretation. But the argument is empirical, not conceptual. There seems no doubt that the *idea* of purely propositional thinking makes perfectly good sense.

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

as applying to all such concepts, should be rejected in favour of some alternative.

Most philosophers of mind over recent decades have claimed that we conceptualise mental states in terms of their distinctive causal roles, or *functions* (Lewis, 1966; Putnam, 1967; Stich, 1983). So for example, beliefs are states which are caused either by perception or inference or testimony, and which in turn interact with desires to generate intentions and actions. Pains are states which are caused by bodily damage or disturbance, which in turn cause the subject to have a desire to cry out, rub the offending part, and so on.

On this account, there is no problem in allowing for non-conscious as well as conscious mental states, provided that the difference between the two can be accounted for in terms of causal role. Nor is there any problem in allowing for mental states which lack feels. Moreover, it remains explicable that metaphysical dualism should ever have seemed an option. For although we conceptualise mental states in terms of causal roles, it can be a contingent matter what actually *occupies* those causal roles; and it was a conceptual possibility that the role-occupiers might have been some sort of *soul-stuff*. However, there are two main problems with analytical functionalism:

- (1) It is committed to the analytic–synthetic distinction, which many philosophers think (after Quine’s ‘Two dogmas of empiricism’ – 1951) to be unviable. And it is certainly hard to decide quite which truisms concerning the causal role of a mental state should count as analytic, rather than just obviously true.
- (2) Some mental states seem to be conceptualised purely in terms of subjective feel, or with beliefs about causal role taking a secondary position, at least. For example, it seems to be the feel of pain which is essential to it (Kripke, 1972). We seem to be able to imagine pains which occupy some other causal role; and we can conceive of states having the causal role of pain which are not pains (which lack subjective feel).

These problems seem sufficient to motivate rejection of analytic functionalism, in favour of the so-called ‘theory-theory’.

## 2.2 *Theory-theory*

A better variant on functionalism about mental-state concepts is to say that such concepts (like theoretical concepts in science) get part of their life and sense from their position in a substantive *theory* of the causal structure and functioning of the mind. (The other part they get from their

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

## 8 1 Assumptions, distinctions, and a map

causal-referential relations to the items which they concern.) On this view, to know what a belief is (to grasp the concept of belief) is to know sufficiently much of the theory of mind within which that concept is embedded. All the benefits of analytic functionalism are preserved. But there need be no commitment to the viability of an analytic–synthetic distinction, if only because of the indeterminacy of ‘sufficiently much’.

(Some of us also believe that our theory of mind – generally called ‘folk-psychology’ – is largely implicit and substantially innate, emerging in normal human children by means of maturation-in-a-normal-environment, rather than by a process of learning. It is very hard indeed to see how the theory could be acquired so early – by the age of three or four – by ordinary learning; and there are just the same patterns of genetically-caused breakdown which one would expect if it *were* innate – i.e. autism, widely thought to be a kind of *mind-blindness*. See Fodor, 1987, 1992; Leslie, 1991, 1994b; Carruthers, 1992a, 1996c; Happé, 1994; Baron-Cohen, 1995; Botterill and Carruthers, 1999, chs.3–4; Hughes and Plomin, 2000.)

What of the point that some mental states seem to be conceptualised purely or primarily in terms of feel? A theory-theorist can allow that we have *recognition capacities* for some of the theoretical entities characterised by the theory. (Compare the diagnostician who can recognise a cancer – immediately and without inference – in the blur of an X-ray photograph.) But it can be claimed that the concepts employed in such capacities are also partly characterised by their place in the theory – it is a *recognition application* of a *theoretical* concept. Moreover, once someone possesses a recognition concept, there can be nothing to stop them prizing it apart from its surrounding beliefs and theories, to form a concept which is *barely* recognition. Our hypothesis can be that this is what takes place when people say that it is conceptually possible that there should be pains with quite different causal roles.<sup>6</sup>

The only real competitors to a theory-theory account of our folk-psychological concepts are some combination of Cartesianism with simulation, on the one hand, or some sort of *interpretationalism* or *quasi-behaviourism*, on the other, of the sort defended by Davidson (1970, 1974, 1975), Dennett (1971, 1981, 1987), Gordon (1995) and various Wittgensteinians.

The former position has already been criticised briefly above. Here just let me mention, in addition, that since this position takes phenomenal consciousness for granted, its adoption would cut us off from the possibility of reductively explaining such forms of consciousness in cognitive

<sup>6</sup> It will be a consequence of the position to be defended in chapter 9:3 that purely recognition (or ‘Cartesian’) concepts of experience, while perfectly possible, are actually parasitic upon a theoretical understanding of the subjectivity of experience.



Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

## 3 Some distinctions: kinds of consciousness

9

terms – for if our concepts of the cognitive are grounded in awareness of *feel*, then we cannot use the former in reductively explaining the latter. The fruitfulness of cognitivist approaches to the problem of phenomenal consciousness – to be defended at length in this book – will therefore be a further, if somewhat back-handed, argument against the introspectionist–simulationist position.

As for interpretationalism, I believe that this sort of view is unacceptable in its *anti-realism*, failing to do justice to the realistic commitments of the folk (see Fodor, 1987, ch. 1; Davies, 1991; Botterill and Carruthers, 1999, ch. 2). So I shall assume that theory-theory is the default position to adopt, unless considerations to do with phenomenal consciousness can convince us otherwise.

### 3 Some distinctions: kinds of consciousness

There are a number of different notions of consciousness and/or a number of different kinds of use of the term ‘conscious’ which need to be distinguished carefully from one another. Failure to draw the right distinctions, and/or failure to keep the different notions apart, has vitiated much work in the area. What follows draws heavily on the work of Rosenthal (1986), Block (1995) and Lycan (1996).

#### 3.1 *Creature-consciousness 1 – intransitive*

Sometimes we treat consciousness as an intransitive, non-relational, property of a creature. Here the *subject* of consciousness is the person (or animal); and consciousness is treated as a simple property of that person. So we speak of someone ‘losing consciousness’ and ‘regaining consciousness’; we say of the coma-victim that he has not been conscious since his accident; we say ‘I want you to make sure that my cat is not conscious during the operation’; and ‘I was conscious all the while’; and so on.

Here ‘conscious’ seems to be more-or-less equivalent to ‘awake’. Roughly speaking, to say of an organism that it is conscious (intransitive) is just to say that it is awake, as opposed to asleep or comatose. At any rate, it seems to be a *sufficient* condition for a creature to count as conscious at *t*, that the creature should be awake at *t*. It is perhaps more debatable whether wakefulness is also a *necessary* condition of intransitive creature-consciousness. For we might wonder whether or not we should say that people are conscious during periods of dreaming, even though we are quite clear that they remain *asleep* during dreams.

I suspect that what may be going on here is that we think that being a subject of conscious mental states – state-consciousness; see sections 3.4

Cambridge University Press

0521781736 - Phenomenal Consciousness: A Naturalistic Theory

Peter Carruthers

Excerpt

[More information](#)

10      1 Assumptions, distinctions, and a map

to 3.6 below – is a sufficient condition for intransitive creature-consciousness; and what we are wondering is whether dream-experiences should count as conscious ones. I am more inclined, myself, to say that state-consciousness need not imply creature-consciousness. So I would be inclined to say that the dreaming subject is not conscious (hence requiring that wakefulness be necessary *and* sufficient for creature-consciousness), although the dreamer may be undergoing mental states which are conscious. But this point will not matter to what follows.

There does not seem to be anything especially philosophically problematic about intransitive creature-consciousness, as such. At any rate, the awake–asleep distinction, while no doubt interesting, does not seem to hold any particular difficulties for physicalist and theory-theory conceptions of the mental. And in so far as there *is* anything problematic about this form of consciousness, the problems derive from its putative conceptual connections with *state-consciousness*. The latter notion will be discussed below.

### 3.2      *Creature-consciousness 2 – transitive*

Besides saying of an organism that it is conscious (*simpliciter*) we also say of it that it is conscious *of such-and-such* (transitive), or aware of such-and-such. To say this is normally to say at least that the organism is *perceiving* such-and-such. So we say of the mouse that it is conscious of the cat outside its hole in explaining why it does not come out; meaning that it *perceives* the cat's presence. To provide an account of transitive creature-consciousness would thus be to attempt a theory of perception. No doubt there are many philosophical problems lurking here; but I propose to proceed as if I had the solution to them.

Two points about perception are worth making in this context, however. The first is that perceptual contents can be (and often are, to some degree) *non-conceptual*. While perception often presents us with a world of objects categorised into kinds (tables, chairs, cats, and people, for example) sometimes it can – and in the case of young children and many species of animal, presumably it often does – present a world which is largely unconceptualised, but rather presented as *regions of filled space* (Peacocke, 1992). Perception presents us with a complex array of surfaces and filled spaces, even when we have no idea *what* we are perceiving, and/or have no concepts appropriate to what we perceive. Imagine a hunter–gatherer transported to some high-tech scientific laboratory, for example – she may have literally no idea what anything that she is seeing *is*; but for all that she will see the distribution of surfaces, shapes and masses; she will have an idea which are distinct objects; which are liftable; and so on.