Chapter 1 – The Reality of Consciousness

One world or two?

Mind and matter don’t seem to fit in the same world; this is the essence of the mind-body problem. Space-filling solid stuff doesn’t seem to belong with invisible inner-experiencing. The neural processing of the brain is best known through third-person scientific investigation, whilst the subjective first-person perspective of the mind is best known through introspection. How are we to make sense of these seemingly incongruous things being unified aspects of a single reality?

One option is to avoid the clash altogether by denying the existence of matter: this is the idealist solution to the mind-body problem. A second option is to accept that mind and matter are utterly distinct but equally fundamental features of reality: this is the dualist
solution to the mind-body problem. I don’t think we should rule out either idealism or dualism as possibilities. But it is prima facie desirable to have a unified picture of nature that doesn’t deny either the reality of the conscious mind or the observer-independent existence of the external world; such is the aspiration of the anti-idealist monist.

The aim of this book is to make sense of anti-idealist monism. I focus on two forms: physicalism, one of the most popular views in contemporary philosophy of mind, and Russellian monism, a new and radical alternative to physicalism. The first half of the book is spent arguing against physicalism. The latter of half is spent exploring and defending Russellian monism.

In this first chapter I will briefly describe both of these views, and outline the main claims I make about them. Before doing that, however, I will outline my fundamental methodology and ‘big picture’ perspective on the issues we will be dealing with.

The Big Picture

The datum of consciousness

A thing is conscious just in case there’s something that it’s like to be it, if it has an ‘inner life’ of some kind.¹ There’s something that it’s like for a rabbit to be cold, or to be kicked, or to have a knife stuck in it. There’s nothing that it’s like, in contrast, for a table to be cold, or to be kicked, or to have a knife stuck in it (or so we ordinarily suppose).² There’s nothing that it’s like from the inside, as it were, to be a table. We mark this difference by saying that the rabbit but not the table is conscious.

It is important to note that consciousness so defined – sometimes called ‘phenomenal consciousness’ to distinguish it from other notions – is not something cognitively sophisticated, such that we might be reluctant to ascribe it to non-human animals such as sheep or hamsters. In this respect the standard philosophical definition of ‘consciousness’

¹ Most people trace this way of defining consciousness back to Nagel 1975, although it appears earlier in Sprigge and Montefiore 1971.
² As will become clear in later chapters, I am in fact open to the possibility that tables are conscious. Nonetheless, my purpose at this point is simply to clearly articulate the concept of consciousness, and I do this with reference to commonsense views about which things do and which things do not have consciousness.
differs from a common meaning of that word in science and ordinary life, where it is often used to mean something like an awareness of self or even an ability to reason. In some sense a rabbit is aware of the world around it, but it is doubtful that it is able to think reflectively about itself as an occupant of that world. The lack of self-awareness does not bar the rabbit from enjoying a rich inner experience. Thus the rabbit is ‘conscious’ in the phenomenal sense.

Whilst ‘phenomenal consciousness’ is a technical term, the property it refers to is part of our commonsense picture of the world. We all believe that there is something that it’s like to be hamster, whilst there is nothing that it’s like to be a rock or a planet. Common sense also leaves a grey area. Is there something that it’s like to be a fly? When a fly bangs repeatedly against a window is it experiencing some very crude form of suffering and frustration; or is a fly just a mechanism with no kind of inner experience whatsoever? Such questions are not easy to answer, but the ease with which we grasp the questions illustrates that having an inner life – having phenomenal consciousness – is an essential part of our commonsense conceptual scheme.

Phenomenal consciousness is a general property that comes in specific forms: pain, anxiety, and the forms the experience involved in seeing red, or smelling gasoline, or tasting coffee. Specific forms of phenomenal consciousness are variously called ‘conscious states’, ‘experiential properties’ or ‘phenomenal properties’; I will use these terms interchangeably. And the concepts we deploy when we think about conscious states as such, that is when we think of them in terms of how they feel, or what it’s like to have them, are known as ‘phenomenal concepts.’ For example, if I am in pain, and I attend to my pain and think about in terms of how it feels, I thereby form a phenomenal concept of my pain.

My methodological starting point is that phenomenal consciousness is a hard datum which any adequate theory of reality must accommodate. Moreover, consciousness must be accommodated unrevised, in the following sense:

_The Consciousness Constraint_ – Any adequate theory of reality must entail that at least some phenomenal concepts are satisfied (A concept is satisfied when it truly corresponds to reality, e.g. the concept of God is satisfied if and only if God exists).
This principal does not entail that we never make mistakes about consciousness, even in one’s own case. Consider the following example. You are extremely anxious during a visit to the dentist. The dentist applies pressure to the inside of your mouth. The combination of anxiety and pressure leads you to think that you are feeling pain, when in fact this is not the case. This could be a case in which you mischaracterise your own conscious experience: you think you are feeling pain when in fact you are feeling anxiety and a sensation of pressure. The Consciousness Constraint tells us merely that some phenomenal concepts are satisfied, and so is consistent with the fact that we apply the wrong phenomenal concepts in certain cases.

Nonetheless, the Consciousness Constraint is a very significant constraint on metaphysical enquiry, as it tells us that one of our ordinary pre-theoretical concepts gets the world exactly right. In general we should be happy to revise our ordinary notions to make them fit with our best scientific theory of the world. We have a commonsense notion of solidity according to which solid objects are all filled in. Science has shown us that nothing is solid in that sense: the chair you are sitting on is mostly empty space. This doesn’t mean nothing is solid. It just means we have to revise our pre-theoretical notion of what solidity is. Einstein told us some weird things about the nature of time, and we revised our concept of time to accommodate them. In some cases we are even happy to eliminate from our ontology certain commonsense entities. Perhaps our scientific world view is incompatible with the existence of free will, and as a result philosophers ought to conclude that there is no such thing.

Matters are different when it comes to consciousness. I cannot rule out for certain the possibility that I am in the Matrix and my entire experience of the external world is an illusion. But it seems that I can rule out the possibility that I have no inner life whatsoever; the evil computers couldn’t be making me think I’m conscious when in fact I’m not. And it’s not simply that I know that there’s something that vaguely resembles our commonsense notion of ‘consciousness’, as might be the case with our commonsense notion of ‘solidity’.

---

3 I will argue in chapter 5 that certain judgments about our experience do involve something close to certainty, but this is not true across the board.
4 As Tom Winfield pointed out to me, in many Matrix style scenarios, the computers are beneficent rather than evil. I maintain my epistemic right, however, to feel ill at ease with them.
When I entertain the proposition <there is something that it’s like to be me> I know with a very high degree of justification, greater than my justification that the external world exists, that *that very proposition* is true – not merely that proposition or some revision of it containing slightly different concepts. The Consciousness Constraint has profound implications for scientific theorising, setting strict limits on its potential for rewriting our conceptual scheme. If it turns out that our scientific world view is inconsistent with the existence of inner experiencing – according to our ordinary notion of inner experiencing – then it is the scientific story rather than consciousness that must budge.\(^5\)

The Consciousness Constraint is my fundamental axiom; the Archimedian fixed point around which all else revolves. Readers who do not share my non-negotiable commitment to the reality of phenomenal consciousness may nonetheless be interested to see what picture of the world is compatible with that commitment. I am inclined to think that appeal to the Consciousness Constraint is a much neglected tool in metaphysics, and in the final chapter I outline a way forward for metaphysics rooted in it.

With the Consciousness Constraint understood, we are now in a position to reformulate the aim of this book: to find an anti-idealist monistic conception of the world that respects the Consciousness Constraint.

**Science and metaphysics**

Philosophers, including myself, pour much time and effort into technical arguments. But most of us are motivated by a ‘big picture’: a fundamental commitment that the philosopher is not going to dispense with unless she *really* has to (and even then she probably won’t!).

Here’s one quite popular ‘big picture’ approach:

*Methodological naturalism* – The lesson we should draw from the success of natural science is that we should look to, and only to, the third person scientific method (i.e.

---

\(^5\) Accepting the Consciousness Constraint is compatible with accepting the fallibility of our philosophical intuitions concerning what is required for phenomenal concepts to be satisfied.
rigorous empirical investigation of what is publically observable) to tell us what reality is like.

Let us spell out this ‘big picture’ in a little more detail. In the last five hundred years or so the project of natural science has gone extremely well. From the movement of the planets, to the evolution of life, to the fundamental constituents of the matter, natural science seems to be an unstoppable juggernaut of explanation. For the methodological naturalist, what this shows is that we’ve finally found something that works, something we can put our ontological faith in. For thousands of years before the scientific revolution philosophers struggled to find out what reality is like and got nowhere. Since the scientific revolution natural science has enjoyed success after success after success.

From this perspective, philosophers who try to look somewhere other than third person empirical science to try to work out the nature of the brain, or of matter in general, are ‘old school’, trying to drag us back to the dark ages. They are to be equated with folk who believe in magic, or deny climate change, or think that the world was created in six days. One of the most fervent defenders of methodological naturalism is Patricia Churchland. In her recent book *Touching a Nerve* she dismisses those who doubt the potential of wholly empirical methods to explain consciousness as ‘nay-sayers’:

Well into the nineteenth century, the scientific consensus was that light is a fundamental feature of the universe, never to be explained by anything more fundamental. What happened? By the end of the nineteenth century, Clerk Maxwell had explained light as a form of electromagnetic radiation.... Imagine a prediction made in the year 2 CE saying that no one will ever understand the nature of fire.... Or imagine a prediction in year 1300 that science will never understand how a fertilized egg can end up as a baby animal. Or a prediction in 1800 that no one will ever succeed in making something that control infections. Suppose someone predicted in 1970 that science could never find a way to record levels of activity in a normal human brain without opening the skull. Wrong. This technical achievement flowered in the 1990s as functional resonance imaging (fMRI) was developed. To the degree
that the nay-saying rests on an unsubstantiated prediction, it need not deter us from moving forward.\textsuperscript{6}

The picture is clear. Methodological naturalists are pioneering into the future. Their ‘nay-saying’ opponents are dragging us back to the past.

There are a number of problems with the picture Churchland paints of her opponents. Firstly, opponents of methodological naturalism are not necessarily giving up on trying to explain consciousness, they simply find it plausible to go beyond exclusive appeal to empirical methods of investigation.\textsuperscript{7} There are some philosophers, perhaps most famously Colin McGinn, who argue that human beings are by their very nature unable to explain consciousness.\textsuperscript{8} But this is by no means an orthodox position amongst the opponents of methodological naturalism. The charge of ‘nay-saying’ is a caricature.

Churchland implies that the opponents of methodological naturalism hold that consciousness is fundamental, by equating this with the mistaken view that light is fundamental. Again, this is a caricature. In later chapters we will explore the views of constitutive panpsychists, who try to explain human consciousness in terms of more fundamental forms of consciousness involved in simpler systems in the brain, and constitutive panprotopsychists, who try to explain consciousness in terms of more fundamental qualities of matter. Far from being ‘nay-saying’, these are systematic attempts to explain.

Of course some opponents of methodological naturalism do take human consciousness to be fundamental; in later chapters we will examine arguments in favour of this view. However, it is not unknown for science to postulate new fundamental properties. Indeed, whilst Maxwell did not take light to be a fundamental kind in its own right, he did introduce new fundamental physical ontology – electromagnetic charge and electromagnetic forces – to account for electromagnetism, and to this extent Churchland’s appeal to Maxwell is misplaced.\textsuperscript{9} All things being equal a sparser ontology is better, but there are many and

\textsuperscript{6} Churchland 2013: 57-8
\textsuperscript{7} By ‘empirical’ I mean empirical investigation of what is publically observable. Hence, I am not counting investigation using first person introspection as ‘empirical’.
\textsuperscript{8} McGinn 1989.
\textsuperscript{9} Chalmers 1995 makes this point.
various criteria a good scientific theory must satisfy, and sometimes expanding fundamental ontology is on balance called for. What we are looking for is the most simple, elegant, unified theory that accounts for the data; and there is no obvious a priori reason to think the best theory won’t postulate consciousness as a fundamental feature of the universe.

In the context of explaining consciousness, methodological naturalism often leads to a kind of *neuro-fundamentalism*: the view that the only way to make progress on explaining consciousness is to do more neuroscience. Naturally, Churchland is a passionate evangelist for neuro-fundamentalism. She ridicules the methods of David Chalmers, one of most well-known opponents of neuro-fundamentalism, with the following piece of cutting rhetoric:

No equipment had to be designed and maintained, no animals trained and observed, no steaming jungle or frozen tundra braved. The great advantage of nay-saying is that it leaves lots of time for golf.\(^{10}\)

Putting on one side the caricatured charge of ‘nay-saying’, this quotation suggests that Churchland has an exceedingly limited view of how science operates, as though it’s simply a matter of doing the experiments and recording the data. In actual fact, many significant developments in science have arisen not from experimental findings in the lab, but from a radical reconceptualization of our picture of the universe done in the comfort of an armchair. Think of the move in the Minkowski interpretation of special relativity from thinking of space and time as distinct things to the postulation of a single unified entity: spacetime; or the move in general relativity from thinking of gravity as a force to thinking of it as the result of curvature in spacetime. My hunch is that progress on consciousness will involve this kind of radical reconceptualisation of the mind, the brain, and the relationship between them; and the Russellian approach explored in the second section of this book looks to be a promising way of doing this.

The neuro-fundamentalist may reply that the conceptual innovations referred to above were ultimately tested empirically; general relativity, for example, was eventually shown to fit the data better than its Newtonian predecessor. But developments in science are not always a matter of fitting the data better; sometimes they involve improvement in the

\(^{10}\) Churchland 2013: 60.
internal virtue of the theory. Special relativity is a case in point. Special relativity and its Lorenzian predecessor are *empirically equivalent*: both predict, for example, the Michelson-Morley finding the speed of light is measured to be the same in all frames of reference. Einstein didn’t account for some data that Lorenz was unable to account for; it’s simply that he found in his imagination a more elegant way of accounting for the data.

Indeed, in formulating special relativity, Einstein wasn’t out doing experiments; he wasn’t ‘designing equipment’ or ‘braving jungles’ – the kind of activities Churchland maintains are required for ‘proper science.’ He was sitting in an armchair wondering whether it would be possible to ride alongside a beam of light, and figuring out what would follow if it were possible. Einstein’s fundamental motivation in formulating special relativity was to bring greater internal unity to physics, by ensuring that a principle of relativity held both in the older science of mechanics and in the new science of electromagnetism. In our enthusiasm for empirical methods, we must not forget that *deep thought* has a crucial role to play in science, and that at this point the boundaries between science and philosophy start to blur.

Moreover, if we accept the Consciousness Constraint outlined in the previous section, then there is a non-empirical datum that any adequate theory of reality must accommodate: the reality of phenomenal consciousness. Suppose we have a theory that accounts for all of the data of third person observation and experiment, including human behaviour and cognitive processing, but which cannot account for the fact that some phenomenal concepts are satisfied. We should reject such a theory as falsified, just as we would a theory which is inconsistent with empirical data. We know that there is phenomenal consciousness, and hence any theory inconsistent with its reality cannot possible be true.

For this reason there is an indispensable role for philosophers in the project of finding out the nature of fundamental reality. There is no experiment that will tell us what is required from reality in order for phenomenal concepts to be satisfied. Working this out is a philosophical not a scientific task, and so we must turn to our best philosophers to tell us how to interpret the Consciousness Constraint. This is not ideal, as philosophers presently profoundly disagree on what is required from reality for phenomenal concepts to be

---

11 At the very least we need philosophical reflection to tell us what experiment or observational data will help shed light on this.
satisfied. But unfortunately we have no other option. We just have to hope that the slow and careful process of rigorous academic activity eventually leads to consensus.

I am not presuming here that satisfying the Consciousness Constraint requires *adding* to the reality which is postulated to account for the data of observation and experiment. It could turn out – although I will argue against this – that the kind of things postulated to account for empirical data happen to be sufficient to also account for the reality of consciousness. But that can’t be assumed from the start; part of the task of working out what will satisfy phenomenal concepts is working out whether empirical postulates will suffice.

It has not been properly articulated in the public mind, nor in the minds of many philosophers, that the reality of consciousness is a datum in its own right, a datum distinct from the data of third person observation and experiment. I think this in part explains the dismissive reaction by people who consider themselves to be ‘scientifically minded’ or ‘hard-nosed naturalists’, to some of the theories which are taken seriously in this book – in particular to panpsychist views according to which the fundamental constituents of the world have conscious experience. Third person observation and experiment does not give us grounds for thinking that electrons have consciousness, and on this basis many conclude that we have no reason to accept such a view. Properly internalising the imperative to account for phenomenal consciousness – as a datum in its own right – is epistemically transformative: one becomes open to peculiar theories of reality that have the potential to account for the reality of consciousness, in the same way that most people are open to the peculiarities of relativity and quantum mechanics given their potential to account for the empirical data.

I do not mean to say that most people are happy to deny the reality of phenomenal consciousness; this is a radical and extremely rare position. I just mean that most people – most philosophers and the vast majority of those outside of philosophical academia – do not explicitly recognize the unrevised reality of phenomenal consciousness as a datum in its own right, nor have they thought through how this impacts upon the project of trying to work out what reality is like. As a result, the general public assumes that it is the job of scientists to tell us what reality is like, and the job of philosophers to make us feel a bit better about the world that science has resigned us to.
I am cautiously optimistic that this is a phase of history, which we will one day move beyond. It results I think from the visceral effect of technology on metaphysical belief. The wonders of modern technology incline educated women and men to place their metaphysical trust in, and only in, the scientific method. As Woody Allen’s lead character in *Deconstructing Harry*, put it: ‘If it’s between the Pope and air conditioning, I choose air conditioning.’ It has come to be thought that respect for the empirical sciences demands, not only that we concede them authority on their own terms, but that we grant them authority on the complete metaphysical – sometimes even the complete ethical\(^{12}\) – truth.

However, so evident is the reality of consciousness, that I can’t believe that consciousness-denial, or even consciousness-revision, will ever get much of a grip on those serious about truth. And sooner or later it will surely become generally appreciated that this entails taking consciousness seriously as a datum in its own right.

Things are moving in the right direction. Consciousness has gone from being a taboo topic for the ‘scientifically minded’, to being seen as a ‘hard problem’ for science. However, despite the subtitle of the book in which David Chalmers coined the phrase, the ‘hard problem of consciousness’ has generally been interpreted as a tricky puzzle which will one day go away if we just do a bit more neuroscience.\(^ {13}\) The next stage is for people to see consciousness not as something to be squeezed into the world we already know about from science, but as an *epistemic starting point*, on a par with the *epistemic starting points* we get from observations and experiments. Consciousness is not a ‘mystery’; nothing is more familiar. What is mysterious is *reality*, and our knowledge of consciousness is one of the best clues we have to working out what that mysterious thing is really like.

**The Philosophical Foundations of Physics**

I finished the last section with some brave hopes for the future. But for the moment, the methodological naturalist’s narrative about the success of science, and the accompanying claims about the morals we should draw from it, is a powerful motivating force for many. I

---

\(^{12}\) See for example Harris 2010.

\(^{13}\) Chalmers 1996. Actually, the phrase ‘hard problem of consciousness’ was coined in his 1995 article.
want to suggest an alternative way of thinking about the history of science, which may motivate us in a different way.

One of Galileo’s most important contributions to the foundations of the scientific revolution was his assertion that the language of natural philosophy should be exclusively mathematical, a view he expressed in this well-known passage from *The Assayer*:

> Philosophy [by which he means ‘natural philosophy’, which became physical science] is written in this grand book, the universe, which stands continually open to our gaze, but it cannot be understood unless one first learns to comprehend the language and read the letters in which it is composed. It is written in the language of mathematics, and its characters are triangles, circles, and other geometrical figures, without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth.\(^\text{14}\)

What is less often focused on is the *metaphysical picture of reality* which underwrote the Galilean research programme. Before Galileo philosophers took the world to be full of *sensory qualities*: colours, smells, tastes, sounds. And intuitively one cannot capture the redness of a tomato, or the spicy taste of paprika, or the sweet smell of flowers, in the austere, abstract language of mathematics. Galileo got around this problem by stripping the world of such qualities and locating them in the soul.\(^\text{15}\)

For Galileo, the spiciness of the paprika isn’t really in the paprika; rather it’s in the soul of the person tasting the paprika. The sweet smell of the flowers isn’t really in the flowers; rather it’s in the soul of the person smelling them. Even colours for Galileo resided not ‘out there’ on the surface of objects but within the human soul. By stripping external objects of any qualities other than shape, Galileo created a metaphysical picture of the material world in which it could be exhaustively described in mathematical geometry:

> Now I say that whenever I conceive any material or corporeal substance, I immediately feel the need to think of it as bounded, and as having this or that shape; as being large or small in relation to other things, and in some specific place at any

\(^{14}\) Galileo 1623: 237–238.

\(^{15}\) In contrast to Descartes, Galileo conceived of the soul in Aristotelian terms, as the principle of animation in the body (*corpo sensitivo*). See Ben-Yami 2015, chapter 3 for more discussion of this.
given time; as being in motion or at rest; as touching or not touching some other body; and as being one in number, or few, or many. From these conditions I cannot separate such a substance by any stretch of my imagination.

But that it must be white or red, bitter or sweet, noisy or silent, and of sweet or foul odour, my mind does not feel compelled to bring in as necessary accompaniments. Without the senses as our guides, reason or imagination unaided would probably never arrive at qualities like these.

Hence I think that tastes, odours and colours, and so on are no more than mere names as far as the object in which we place them is concerned, and that they reside only in the consciousness. Hence, if the living creature were removed, all these qualities would be wiped away and annihilated.¹⁶

We can see then that it was a change in our philosophical conception of the world that made mathematical physics possible, and that change was a matter of placing the sensory qualities we encounter in conscious experience outside of the material domain studied by physics.

This is the start of mathematical physics, which subsequently proved to be a great success. Once we can capture nature in mathematics, we can start to frame laws of nature in mathematical language. A short while later we had Newton’s laws of motion and gravity. And five centuries of developing more and more accurate mathematical models of the world’s causal structure has enabled us to manipulate it in all sorts of extraordinary ways, giving us lasers and microwave ovens and flights to the moon.¹⁷

It is tempting to take this success as evidence for methodological naturalism, or even a kind of physics-fundamentalism, according to which the great success of physics shows that physics is on its way to giving us a complete account of fundamental reality. But this is the wrong conclusion to draw. The success of mathematical physics resulted from limiting the scope of enquiry; by putting the sensory qualities we encounter in conscious experience –

¹⁷ By ‘causal structure’ I mean features of reality which can be captured in a mathematical or causal vocabulary. This notion is spelt out at much greater length in the account of ‘pure physicalism’ in the next chapter.
colours, smells, tastes, sounds – outside of the domain of the physical sciences, we are able to give a purely mathematical description of what’s left over. But those qualities that Galileo took out of the material world still exist *somewhere* and must still be accounted for *somehow*. If the spicy taste of the paprika isn’t really in the paprika then where is it? Galileo thought it was in the soul, but if we don’t want to believe in souls then we need to find a place in the natural world for the sensory qualities.

But do these reflections on the *origins* of science really tells us very much about its *potential*? Isn’t the great success of Galilean physics evidence that it’s fundamentally on the right lines, and will one day explain even the sensory qualities which Galileo thought were beyond its reach? This line of reasoning is confused, as can be seen in the following analogy.

In my first year as an Assistant Professor, the head of department kindly let me off administration duties, allowing me to focus exclusively on teaching and research. The fact that I did well when I didn’t have to do administration obviously gives us no reason to think that I would be good at administration. Similar, the fact that physical science has done well since sensory qualities were put outside its domain of enquiry, gives us no grounds for thinking it has the potential to adequately deal with sensory qualities themselves.

And so it is just wrong to think of the anti-physicalism I shall defend in this book as pushing against the onward march of science. The anti-physicalist can perfectly well appreciate that physics is sovereign in its own domain, i.e. the causal structure of fundamental reality, whilst maintaining that there is more to fundamental reality than mere causal structure.

Galileo took the sensory qualities out of the first material in order to make mathematical physics possible. But we know with something close to certainty that these sensory qualities exist; their reality is evident in our experience. If redness doesn’t exist ‘out there’ on the surfaces of tomatoes, then at least it exists in my mind as a form of conscious experience. At some point those qualities need to be put back into our metaphysical picture of the world. That is precisely the aim of this book.19

---

18 Sadly I wasn’t.
19 I am inclined to think that if Galileo were to time-travel to the present day and be told about the ‘hard problem of consciousness’, his response would be: ‘Of course physical science can’t explain the qualities we find in experience. I created physical science by taking them out of matter!’
Having hopefully given the reader a sense of where I’m coming from, I turn now to the view we shall be exploring.

**Physicalism and Russellian monism**

**Physicalism**

Physicalism is the view that fundamental reality is entirely physical. Despite the popularity of this view, it is actually quite tricky to spell out exactly what it amounts to, as we will explore in detail in the next chapter. In order to simplify matters somewhat, the view I focus on in the first half of the book is an unadulterated form of the physicalist position which I call ‘pure physicalism’. Pure physicalism involves the thesis that the complete nature of fundamental reality can in principle be captured in the vocabulary of the physical sciences. I go on in later chapters to show how the argument can be applied to more diluted forms of physicalism.

My charge against physicalism is a familiar one: physicalism is unable to account for the reality of phenomenal consciousness. In other words, physicalism is unable to respect the Consciousness Constraint. Although familiar, my argument differs significantly in detail to those which have been given previously.

Probably the two most discussed arguments against physicalism in recent times are Frank Jackson’s version of the knowledge argument and David Chalmers’ version of the conceivability argument. In Jackson’s knowledge argument, we hear about Mary, the brilliant neuroscientist raised in a black and white room from where she learns everything there is to know about colour experience. One day Mary sees red for the first time and learns something new: what it’s like to see red. The moral of the story is that there’s more to the conscious experience of colour than physical science can ever convey. In Chalmers’ conceivability argument he tries to demonstrate that ‘zombies’ – physical duplicates of human beings that lack any kind of conscious experience – are conceivable, in the sense that

---

no contradiction or incoherence can be found in the idea of such a thing. He goes on to argue that if zombies are conceivable, then they’re possible, and if they’re possible then the physical facts cannot account for consciousness.

I have concerns with both of these arguments. The knowledge argument plausibly demonstrates an *epistemic gap* between the physical facts and the facts about experience, but more needs to be done to transform it into a metaphysical gap. My worry with Chalmers’ conceivable argument is that it rests on highly contentious semantic assumptions – those assumed by the ‘two-dimensional’ semantic framework in terms of which the argument is posed – which in my view Chalmers has never adequately justified.

In my discussion of these arguments I try to show that what is really needed to make these kinds of arguments work is a commitment to *Phenomenal Transparency*: the thesis that phenomenal concepts are transparent. To understand this thesis we need to understand a distinction that is central to this book: the distinction between transparent and opaque concepts.\(^{21}\)

A concept is *transparent* just in case it reveals the nature of the entity it refers to, in the sense that it is a priori (for someone possessing the concept and in virtue of possessing the concept) *what it is for that entity to be part of reality.* A plausible example is the concept <sphericity>. For the property of sphericity to be instantiated is for there to be something with all points on its surface equidistant from its centre; if you possess the concept of sphericity, and you’re clever enough, you can work that out from the armchair.

A concept is *opaque* just in case it reveals little or nothing of the nature of the entity it refers to, i.e. very little of what it is for that entity to be part of reality is a priori accessible (for someone possessing the concept and in virtue of possessing the concept). The concept <water> is a plausible example of an opaque concept. For there to be water is for there to be stuff composed of H\(_2\)O molecules; but this cannot be found out from the armchair.\(^{23}\) In

---

\(^{21}\) I have developed this notion in a number of papers, beginning with Goff 2011. Martine Nida-Rümelin (2007) has a very similar notion of *grasping a property.*

\(^{22}\) The kind of reality in question will depend on the kind of entity. For an individual to be party of reality is for it to exist. For a property to be part of reality is for it to be instantiated (Platonists might want to distinguish a property’s existing from its having concrete reality). For an event to be part of reality is for it to take place.

\(^{23}\) For reasons I’ll explain chapter 6, I think the concept <H\(_2\)O> is also opaque, but I ignore this complication for now.
contrast to sphericity, you have to actually do some science to find out the essential nature of water. If <pain> is a transparent concept, then we can know the nature of pain simply by having the concept, without having to leave the armchair. According to Phenomenal Transparency, the phenomenal concept of pain is, in this respect, more like <sphericity> than it is like <water>.

Suppose that Phenomenal Transparency is true. According to physicalism, a physical state like pain has a wholly physical nature; perhaps – to continue with the hackneyed and empirically implausible example favoured by philosophers – for something to feel pain is for its c-fibres to fire. If this were the case, then – assuming that phenomenal concepts are transparent – the physical nature of pain would be a priori accessible; it would be a priori that for something to feel pain is for its c-fibres to fire (or whatever the specific physical nature is). But it is clearly not a priori that to be in pain is to have your c-fibres fire – neuroscience can’t be done from the armchair! – and hence physicalism must be false. If we can secure Phenomenal Transparency, we have a strong case against physicalism.

Now suppose that phenomenal concepts are opaque: that we have no a priori access to the nature of the states they refer to. It follows, or so I try to argue, that we can’t possibly have an a priori reason to deny that the referents of phenomenal concepts are neurophysiological states. If we have no grasp of what on earth conscious states are, then how on earth could we rule out that they are brain states? Without Phenomenal Transparency, the epistemic gap between the physical and the experiential has no metaphysical significance.

Thus I try to show that the charge that physicalism can’t account for consciousness hangs and falls on Phenomenal Transparency. In the final chapter of the first half of the book I present an argument for Phenomenal Transparency, and on the basis of this I reject physicalism.

**Russellian monism**

In 1927, in *The Analysis of Matter*, Bertrand Russell defended a couple of theses that amounted to a novel approach the mind-body problem. By coincidence similar claims were defended by Arthur Eddington in his Gifford lectures of the same year. This approach was
forgotten about in the latter half of the twentieth century, perhaps because it didn’t fit with the physicalist zeitgeist of the time. However, it has recently been rediscovered, leading to a view – or better a school of views – known as ‘Russellian monism.’

There are a number of slight variants on the definition of Russellian monism, but essentially it involves a negative and a positive components:

The negative component: The negative component of Russellian monism consists in the claim that physical science provides only a limited account of the nature of the physical world. This is variously put as the claim that it captures only the relational, dispositional, extrinsic, or structural properties of matter. Russellian monists argue that there must be some further nature to matter underlying these characteristics – a non-relational, categorical, intrinsic, or non-structural nature – about which physics remains silent. I call this the ‘deep nature’ of matter.

The positive component: The positive component of Russellian monism consists in the claim that the deep nature of matter transparently explains consciousness, in the sense that there is an a priori entailment from facts about the deep nature of matter to facts about consciousness. If you were able to access the deep nature of my brain, you would in principle be able to deduce the nature of my consciousness. The mystery of consciousness, then, results from our ignorance concerning the deep nature of matter.

Russellian monism is increasingly being seen as an attractive middle way between the extremes of physicalism and dualism, avoiding many of the problems associated with each of these views. Consider the difficulty outlined in the previous section of reconciling mind-brain identities with the transparency of phenomenal concepts. If the concept of pain reveals what pain essentially is, and pain just is c-fibres firing, why is it not a priori that pain is c-fibres firing? The most common position among contemporary physicalists is to deny Phenomenal Transparency: it is because phenomenal concepts are opaque, and do not reveal the physical nature of pain, that the identity <pain=c-fibre firing> is not knowable a priori.

The Russellian monist turns this on its head: it is the physical concept, not the phenomenal concept, which is opaque. According to Russellian monism, when we think about c-fibre
firing as ‘c-fibre firing’ we characterise it extrinsically, primarily in terms of its functional role in the brain. It is only when we think about c-fibres firing as ‘the feeling of pain’ that we understand the essential nature of pain. For the physicalist pain is essentially c-fibre firing, whilst for the Russellian monist c-fibre firing is essentially pain. If am right that the classical challenges to physicalism, expressed in the knowledge and conceivability argument, all ultimately reduce to the difficulty of reconciling physicalism with Phenomenal Transparency, then there is good reason to think that the Russellian monist can avoid these difficulties.

Dualism on the other hand has profound problems accounting for mental causation. If – as many philosophers believe – the physical forms a causally closed system, it is hard to see how a non-physical mind could causally impact on the physical world. But for the Russellian monist (at least in paradigmatic forms) the mind is part of the causally closed physical world, in virtue of being grounded in the deep nature of the physical.

Russellian monism comes in a variety of forms. Broadly speaking there are three dimensions of difference:

Panpsychist versus panprotopsychist – Panpsychist Russellian monists believe that the deep nature of the physical world is experiential in nature. The typical view is that basic physical entities have a very simple experiential nature, and that the complex experience of humans and animals is in some sense derived from the more simple experience of basic physical entities. Panprotopsychist Russellian monists in contrast hold that the deep nature of the physical is not itself experiential, but somehow intrinsically suited for realising, or bringing about, experience.

Constitutive versus emergentist – Constitutive Russellian monists believe that human and animal experience is *grounded in, or constituted of*, the deep nature of the physical. Emergentist Russellian monists in contrast hold that human and animal experience is *causally brought about and sustained by* the deep nature of the physical.

Smallist versus priority monists – Smallists believe that all fundamental facts are facts concerning micro-level entities and properties; all things exist and are the way they are in virtue of how micro-level entities are. Priority monists in contrast hold that the universe is
the one and only fundamental entity; all other things exist and are the way they are in virtue of how the universe is. Russellian monism can be combined with either of these views.

In chapter 6, and in more detail in chapter 9, I raise empirical worries for emergentist forms of Russellian monism. Very roughly, the idea is that if human and animal conscious subjects were fundamental entities, then they would presumably have distinctive causal powers different from the causal powers of basic material entities. It would seem to follow that there would be all sorts of happenings in the brain or body which could not be explained in terms of the kind of causal principles which operate in nature more generally, in both the animate and inanimate realm. The fact that we don’t seem to find such events in the brain constitutes an inductive argument against emergentism.

On the other hand, there are significant a priori difficulties facing constitutive forms of Russellian monism. I spend chapters 7 and 8 discussing various forms of the ‘combination problem’ for Russellian monism. In its paradigmatic form the combination problem is a challenge for smallest panpsychism, and goes as follows: how on earth do we make sense of the idea of little conscious subjects, such as, electrons and quarks, combining to make a big conscious subject, such as a human brain? However, there are many related concerns to do with making sense of conscious subjects as composite objects, and many of them apply to both panpsychist and panprotopsychism forms of Russellian monism.

At the end of chapter 8 I present what I take to be the hardest combination problem, which I call the ‘Subject Irreducibility Problem’. The problem is that conscious subjects seem to be in a certain sense irreducible: it doesn’t seem that we can specify what it is for there to be a conscious subject in more fundamental terms. I ultimately argue that the problem of Subject Irreducibility has no solution so long as we are working with smallest models of Russellian monism. In chapter 9 I propose, however, a model combining priority monism and panpsychism which I hope is able to make sense – paradoxical as it may sound – of conscious subjects being irreducible and non-fundamental.