

Chapter 5

Mind and Body: What Should We Believe?

Midterm review. On the mind/body problem we have so far done the following:

- Introduced the two basic positions: dualism and materialism
- Briefly sketched the implications of each for immortality and free will
- Examined whether there is a criterion of the mental
- Laid out the main varieties of dualism
- Taken a brief look at idealism
- Sketched the main varieties of materialism, looking in some detail at functionalism and neurophilosophy

The previous chapter gives us a fairly complete survey of dualist and monist positions on the mind. Now we come to the key question: Of the range of different positions, which should we believe? Which has the greatest likelihood of being true? That is the topic of this chapter.

1 Why Has Dualism Had Such a Strong Appeal?

Before we turn to that question, let's stand back for a minute and survey the scene as a whole. For the past fifty years or so, most philosophers of mind, psychologists, and other researchers working on the mind have been materialists of one stripe or another. (Indeed, the vast majority of them have held the variety of materialism we earlier called FUNCTIONALISM, a view to which we will return.) Indeed, materialism is so widely believed nowadays that dualism is often not even taken seriously. This simple faith in materialism makes it easy for those who hold it to forget that things have not always been thus. Indeed, up until about 100 years ago, it seemed clear to most people that some form of dualism had to be true, anyone who thought otherwise was simply ignoring some obvious facts. Any view that seemed so obviously true to so many highly educated and

intelligent philosophers and psychologists is not a mere silly, simple-minded mistake. It may (or may not) be wrong, but it is not silly or simple-minded.

In fact, dualism was so dominant prior to about 1900 that theorists of any enduring influence who were materialists can be counted on one hand: the Greek philosopher Democritus (b.ca. 460 B.C.), the Roman philosopher Lucretius (b.ca. 94 B.C.), the English philosopher Hobbes, a few French philosophers such as de la Mettrie (b. 1709) and Condillac (b. 1715), and that's about it. Until very recently, almost all the most intelligent people in the world believed in dualism. The most significant exception, actually, consisted not of materialists but IDEALISTS. (We discussed them in the last chapter.) A theory that seemed so obviously true to so many intelligent people for so many centuries is not to be lightly dismissed. What is it about the mind that made dualism seem so compelling to so many people for so long?

Two answers are often given, but they are both too simple. The first oversimple answer: until very recently, most philosophers were religious; religious people believe in immortality; but (barring bodily resurrection, which many found too mysterious) it is difficult to believe in immortality unless you believe in dualism; so philosophers believed in dualism. This answer is a bit insulting. It suggests that some of the greatest minds who ever lived let their religious convictions control their theory of mind. Anyway, it is not true. Descartes, the philosopher who invented the dualist picture of the mind as we know it, had doubts about many important aspects of religion, yet dualism still seemed to him to be true. Indeed, he thought that he had a number of absolutely airtight arguments to show that it is true.

The second oversimple answer: INTROSPECTION, not religious faith, is what made dualism seem plausible. When philosophy and psychology replaced introspection with something better about 1900, dualism soon lost its appeal. This answer is a lot better than the first one. Indeed, it has an element of truth to it. One of the reasons why it has seemed to many people that the mind is something different from the brain is indeed the way in which we appear to ourselves in introspection. Introspection is the awareness we have of ourselves "from the inside." It's the awareness you have of your thoughts, feelings, etc., just by having and paying attention to them. By contrast, if you want to become aware of someone else's thoughts, feelings, etc., you have to observe their bodily movements, listen to what they say, etc.; in short, you have to pay attention to how their thoughts and feelings are manifesting themselves on the surface of the

body. This contrast between awareness of one's own thoughts and awareness of the thoughts of others is one of the main sources of the PROBLEM OF OTHER MINDS, and we will explore it in more detail in chapter 7. Now, many philosophers and psychologists have held that when they "look into themselves," the self that they find is utterly different from anything that could be made out of matter, including the matter of the brain. This is what Wittgenstein described as the feeling of an "unbridgeable gulf between consciousness and brain process" (1953, §412). As we appear to ourselves in introspection, we appear to be something quite different from anything made out of nerve cells or anything else made out of matter.

So introspection has been one of the sources of the appeal of dualism. Unfortunately, introspection has turned out to be not a good method for uncovering the nature of the mind. With the creation of laboratories for studying human behavior in Germany, the United States, and Russia toward the end of the nineteenth century and with the (re)discovery by Freud that a great deal of human mental activity is not open to introspection (is *unconscious*, in one of the meanings of the word 'unconscious'), theorists came to see that there is a great deal that introspection cannot tell us about the mind. Introspection is central to self-knowledge in everyday life—without introspection, we would not be aware of ourselves and this would be as good as being dead—but having a crucial role in everyday life is not the same thing as being a good tool for discovering what we are really like. In addition to not giving us any access to the many things in the mind not open to introspection, introspection also suffers from being unverifiable and inconstant from person to person:

- *Unverifiable* How can anyone else check your introspections? Worse, how can even you compare how your mind seems to you with how your mind actually is?
- *Poor intersubject reliability* How one person seems to herself is very often different from how another person seems to himself, even concerning the same issue.

Finally, it turned out that introspection is often easily fooled. As Freud and others showed us, we distort our awareness of ourselves in all sorts of ways. Indeed, as later theorists have shown, we can even make things up in introspection (this is called CONFABULATION). In short, introspection is a thoroughly unreliable tool for studying the mind and it was abandoned in favor of laboratory experiments and other methods around 1900.

Whatever the problems facing introspection, was it the main basis on which people were led to accept dualism? Not in our view. Introspection

was one thing that made dualism seem plausible, but we believe that another thing played an even larger role in its acceptance.

Before we look further into what gave dualism its appeal, let us introduce a second issue to see if we can't deal with the two of them together. If we want to understand what it is about the mind that gives dualism its appeal to some people, we also want to know what it is about the mind that makes materialism seem just as obviously true to other people. What is it about this thing called the mind that makes two such totally different conceptions, dualism and materialism, so attractive to different people?

The distinction between dualism and materialism is just one reflection, in our view, of a vast, deep distinction between two ways of viewing human beings. Let us call these two ways of viewing people two *images* of the person, a term coined by the contemporary American philosopher Wilfrid Sellars. One image of the person is related to relationships, responsibility, social groupings—in short, to everyday life, to how people appear in our ordinary dealings with them. The other image is more scientific. Sellars called the former the *MANIFEST IMAGE* and the latter the *SCIENTIFIC IMAGE*. We can define the two images as follows:

- (1) *Manifest image* The image of the person as a single, unified center of consciousness and decision making.

This is the image of the person that we find in ordinary moral, social, and interpersonal life. Here we view the person, this center of consciousness and decision making, as a being that can focus attention on things; take account of reasons for doing this or believing that; make decisions in a unified, focused way; govern him- or herself by freely chosen standards; and so on. (Sellars choose the term 'manifest image' because this is the way we manifest ourselves to ourselves in everyday life.)

- (2) *Scientific image* The image of the person as a vast assemblage of cells tied together into a complex system.

This is the image of the person as an organism, a middle-sized mammal in a world of similar creatures—a mammal that is, for example, closely related to chimpanzees, gorillas, and other primates. Here we view the person as we view other organisms: as a huge mass of neurons and other cells linked in complex cause-and-effect relationships. (In fact, there is more than one scientific image of the person alive in our culture currently: in addition to the organic, neurological image just sketched, there is also the picture of the person as a complex information-processing system.)

Now, the manifest and the scientific image do not contradict one another. They are simply two alternative ways of looking at one and the same thing, in this case, the human person. However, for some people, the manifest image seems to be the better picture, i.e., the picture that is truest to our real nature. As we said, it emphasizes the conscious, reasoning, unified aspect of people, and for some people, these properties are the heart of the matter. On the other hand, for other people, the scientific image seems to offer the best hope of generating a complete, detailed, deep theory of persons, it being the image of the person as an incredibly complex biological or information-processing system.

People who are most at home in the manifest image will also tend to think that the scientific image can never capture important elements of what a person is. For these people, dualism will be very attractive: it may appear to be the only theory that offers any hope of capturing those aspects of people that, as it seems to them, the scientific image cannot capture. On the other hand, people who expect that the scientific image will eventually explain everything essential about the human person tend to find materialism more appealing.¹

This distinction between the manifest and scientific images also explains the difference between what most people believed prior to 1900 and what most researchers on the mind believe now. Prior to 1900, for most people the *only* image of the person was the manifest image. As Descartes put it, the mind is “a thing which thinks” (1931 [1641]), where by ‘thinking’ he meant all the features of the manifest image: unified consciousness, focused decision making, governing oneself by freely chosen standards, etc. The idea that people might be nothing more than complex organisms did not take hold until after Darwin (b. 1809) late in nineteenth century. Before then, dualism seemed obviously true.

By contrast, for most theorists now, it is some version of the scientific image that seems to tell the fullest story about people: people are complex organisms, or perhaps complex information-processing systems, but at any rate complex *natural* systems, part of the material universe, like other animals and objects in general. People for whom the scientific image is dominant will be drawn to materialism.

We now have a picture of the background against which dualism looked obviously true to most people and materialism looked equally obvious to a few. The vast divide between the manifest and the scientific images of the person is what mainly accounts for these phenomena. So let’s turn to the big question: which view should we believe (if either)? Because the dualist side is often slighted by contemporary theorists, we

will begin by examining arguments for dualism. Some of them are very interesting.

2 Four Arguments for Dualism

Here is the big question put more precisely: is the mind (as delineated in chapter 4) an aspect of the brain (or brain plus), or is it something different from the brain? That is the big question if we are considering substance dualism. If we are considering property dualism, the big question is this: is the mind simply neural properties of the brain, or is it made up of properties that are radically independent of neural properties? Put yet one more way, is the mind a thing separate and apart from the brain, though in close association with it (substance dualism); is it a set of properties that share a single mind/brain unit with other, purely neural properties (property dualism); or is it simply an aspect of the brain (materialism)? Which view should we accept?

As we just saw, some form of dualism seemed obviously true to most thinkers for virtually the whole of human history up until about one hundred years ago. It is unlikely that a view that was believed that widely has nothing going for it. Many arguments have been advanced by philosophers in support of dualism. They generally have the same structure: Find some property of people (or minds) that no material object could have or that is not a material property of any object. Infer that people (or the business part of people, i.e., minds) are neither material objects nor made up of material properties.

We will consider four such arguments. These four are not by any means all of the arguments advanced on behalf of dualism. They just strike us as four of the most interesting ones. They all derive in one way or another from Descartes, though the form in which we will consider them is sometimes quite different from the form they took in Descartes's writings. They can be grouped into two sets of two (see table 5.1). The horizontal distinction is in terms of the kind of dualism argued for. Two of the argu-

Table 5.1
Four arguments for dualism

	Property dualism	Substance dualism
Introspection	Mental state vs. brain state	Conceivability
Nonintrospection	Intentionality	Indivisibility

ments aim to prove property dualism, two of them aim to prove substance dualism. The vertical distinction needs a word of explanation.

Earlier in this chapter, we mentioned that one of the perennial reasons why people believe in dualism is how we appear to ourselves when we introspect ourselves. We just appear to be something very different from any brain process. And it is certainly true that a number of the most influential arguments for dualism have made exactly this kind of appeal to introspection. Consider this statement of Descartes's: "When I consider the mind, that is to say, myself inasmuch as I am a thinking being, . . ." (1931 [1641]). Notice the move: to examine the *mind*, Descartes looks into *himself*—he appeals to introspection. Descartes simply assumes that the best way to find out what the mind is, is to look into his own self. It would be hard to imagine any theorist making that move now. Nowadays theorists would be much more likely to set up an experiment. (We looked briefly at why introspection was abandoned as a research tool in the earlier discussion.)

If introspection has been one basis for arguing for dualism, however, it has certainly not been the only one. The simple-minded picture of why people believed in dualism prior to this century is too simple. Of the four arguments named in table 5.1, two of them appeal to introspection, but, as we will see, two do not, certainly not in any obvious way. Nor should we leave the impression that Descartes used only arguments from introspection. In fact, he mounted arguments for dualism of both the kinds we are distinguishing here.

It is time now to see how these arguments go. Let's first take a look at the two arguments for property dualism. Property dualism, recall, is the idea that mental properties are utterly independent of neural properties, even if I, the thing that has both kinds of properties, am one single thing.

Mental-State-versus-Brain-State Argument

The mental-state-versus-brain-state argument is simple. Its simplicity notwithstanding, it is probably the source of a lot of the appeal of dualism. Imagine a tasty, tempting meal. Form a really clear image of it. Now pay attention to that image. If that image is a brain state, it should have the properties of brain states. It should be made out of millions of tiny gray or white units, it should be soaked in blood and other liquids, it should be in total darkness—in short, it should be like a brain state.

But the image that I am introspecting has none of those features, so the argument goes: it is a single image, not an assembly of millions of units; it is not gray or white like brain cells are—in short, it is nothing like a state

of a brain. Conclusion: mental states like images are totally different from states of the brain.

Argument from Intentionality

Next, our old friend intentionality. As we said in chapter 4 when we considered it as a criterion of the mental, intentionality enters into philosophy of mind in numerous places. There it was offered as a criterion of the mental. We will now examine the use philosophers have made of it to argue for property dualism.

The argument is deceptively simple:

Premise 1 States and events that have intentionality can be false as well as true, inaccurate as well as accurate.

(Example: Compare 'I believe that rocks are hard' and 'I believe that the Tooth Fairy will give Rob \$1 tonight'. The first belief is true, the second is false (or else Rob is in for a *big* surprise!). Yet both beliefs have intentionality, i.e., are about something. Being about something in this way is what makes perception, thought, science, literature, and all the other products of the human mind possible.)

Premise 2 To be false is to represent what is not.

Premise 3 The material world cannot have to do with what is not.

Conclusion Therefore, states and events that have intentionality cannot be states of matter or events in matter.

Many philosophers have accepted some form of this argument, including Russell (1912).

Conceivability Argument

Now substance dualism. The conceivability argument is the first of two arguments that we will examine for substance dualism, the idea that the mind and the brain are two different things. We have derived the conceivability argument from an argument that Descartes (1931 [1641]) mounts in his Second Meditation, but we will give it our own form. It goes like this:

Premise 1 I can conceive of myself without a body.

Premise 2 If I can conceive of myself without a body, then it is possible for me to be without a body.

Conclusion 1 Therefore, it is possible for me to be without a body.

Premise 3 But if C1, then I am something more than a body.

Conclusion 2 Therefore, I am something more than a body.

This is a very nice argument from a number of points of view. First, it looks to be *valid*. It appears to have the following form:

P1 A

P2 If A, then B

C1 B

P3 If B, then C

C2 C

This is clearly a valid form of argument: it is just two applications of *modus ponens*. Of course, we need to make sure that the argument *really does* have this form. If it does, then attention turns to its premises. Is *P1* true? Is *P2* true? Is *P3* true? For if the conceivability argument is valid and the answer to these three questions are all 'Yes', then the argument is also *sound*—and that means that dualism is true. We will try to answer these questions, but before we do, we need to lay out the fourth argument.

Indivisibility Argument

The indivisibility argument, also derived from Descartes, is extremely simple. To our minds, it is also extremely powerful. Indeed, we think that it proves a form of dualism. (Whether it is the form of dualism that Descartes and other dualists wanted is another question.) Here is how it goes:

Premise 1 The mind cannot be divided.

Premise 2 All material objects, including the brain, can be divided.

Conclusion Therefore, the mind is not the brain (or any other material object).

Notice that the term 'divided' could mean a number of different things. We will return to this issue.

3 Assessment of the Arguments

How well do these four arguments for dualism stand up to critical evaluation?

Mental-State-versus-Brain-State Argument

At first glance, the mental-state-versus-brain-state argument looks pretty solid: as we are aware of mental states, they do indeed seem to be nothing at all like brain states. We need to ask, however, whether what mental states *seem to us* to be like is necessarily a good indication of what they

are like. Could appearances be deceiving here? Many things do not appear to be as they are.

Consider a lowly table. Modern science tells us that a table is mostly empty space. Indeed, well over 99 percent of it is empty space. It is made up of atoms and molecules, and atoms and molecules consist of a small number of extremely tiny electrons circling at a vast distance, relative to their size, from an only somewhat larger nucleus. There is nothing in the space between them. If so, tables are mostly empty space.

Yet they do not appear to be mostly empty space. They appear to be completely solid. Perhaps the same is true of mental images: perhaps they appear to be very different from brain states yet are brain states.

The analogy with a table can be put another way. Ask yourself, “How would a table that was mostly empty space appear to us?” Answer: “As a solid table.” For that is how mostly empty space containing a bunch of nuclei and electrons appears to us: as solid. Now ask, “How would a complex circuit in the brain appear to us in introspection?” Perhaps the answer is “As a mental image.” What else would you expect a brain state to appear as?

In short, the mental-state-versus-brain-state argument does not work. We have found an analogous case, a case that seems indeed to be completely analogous in the relevant respect, and it has revealed a flaw in the original argument. The argument depends on an inference principle that we considered in chapter 2: the way things seem \rightarrow way things are. And we saw that this principle is not always reliable. A stick in water *seems* to be bent, but for all that, it is not bent. The inference from how mental states seem to us to how they are may be no more reliable than the chapter 2 example of an inference from how the stick seems to us to how the stick is.

Before we move on, note something important: we have not shown that mental states *are* brain states (or that mental events are brain events). All we have shown is that an argument that they are *not* does not work. But, as we will try to show in the next section, that may be enough. On to the second argument.

Argument from Intentionality

Is having intentionality enough to show that something cannot be made out of matter? Possible counterexamples immediately come to mind: photographs and sentences in a book. They are clearly made out of matter, yet they are about something and can be false and inaccurate as well as true and accurate. If so, having intentionality does not stand in the way of

something being made out of matter. However, as we said in chapter 4, sentences and such may derive their intentionality from elsewhere, so they are not a clearcut counterexample. The real problem is that we do not understand intentionality well enough to say with any confidence *what* it implies for the mind/body problem. For this reason, it cannot be used as an argument for anything.

Again, it is important that we are not arguing that intentional states *are* brain states; we are merely urging that an argument that they are not does not work.

That brings us to the arguments for substance dualism, the idea that the mind and the brain are two different things.

Conceivability Argument

We said that *if* the conceivability argument is valid and *if* the premises are true, the argument is sound and dualism is true. It is now time to examine those two big *ifs*. First, premise 1: I can conceive of myself without a body. Is this true? *Can* I conceive of myself without a body? It is in fact far from clear that I can. Perhaps I can form some vague idea of being without a body, but whether I can form a clear, complete picture of myself without a body is another question. (I can also form a vague idea of a round square, but just try to form a clear image of one.) Indeed, there is a risk that P1 is *begging the question*, already assuming, that is, that dualism is already true—just what the argument is trying to prove. For if I *am* my body—and that possibility has to remain open till the conclusion is reached—then for the argument not to beg the question, to conceive of myself without a body I would have to conceive of myself without *myself*. In short, P1 is far from clearly true and also seems to beg the question.

Next, is the conceivability argument valid? One requirement for an argument to be valid is that all the key words in the argument are used with the same meaning throughout. Changing the meaning of key terms is called the FALLACY OF EQUIVOCATION and ruins an argument. Here is an example of an argument that clearly fails because of equivocation:

Premise 1 Happiness is the end of life.

Premise 2 The end of life is death.

Conclusion Therefore, happiness is death.

(The problem, of course, is with the word ‘end’: in P1 it means goal; in P2 it means termination.)

Compare P1 (I can conceive of myself without a body), and P2 (If I can conceive of myself without a body, then it is possible for me to be without

a body). To have a chance of being true, P1 has to be using 'conceive' in a very weak, "form a vague idea of" sense (see the example of the round square just above). But P2 requires the much stronger sense of 'conceivable' at work in this inference principle:

(3) *The principle of conceivability* What is conceivable is possible.

Showing that something is conceivable in *this* way is enough to show that it is possible. This sense of 'conceive' is stronger than the one used in P1 because conceiving of a round square in that former sense emphatically does *not* show that round squares are possible. If P1 and P2 are using the word 'conceive' in two different senses, then the argument suffers from a fallacy of equivocation, and P1 and P2 together do not entail C1.

Interestingly enough, the move that many people object to immediately, the move made by P3, may be alright. P3 depends on the following inference principle:

(4) *Inference principle* If it is so much as possible for *me* to be without *a body*, then this *me* is something more than the body.

And (4) seems alright: If *I* could exist without *a body*, then it would seem that this *I* does indeed have to be more than the body. Yet because of the problems already found with P1 and P2, it does not matter whether P3 is true or not, and we won't examine it further.

To summarize, even if P3 is all right, the argument as a whole is not. First, P1 may beg the question it is supposed to help settle. Second, the move from P1 to P2 seems to suffer from equivocation. If so, the conceivability argument is not valid, and even it is were, it would not be sound.

Indivisibility Argument

The indivisibility argument goes as follows:

Premise 1 The mind cannot be divided.

Premise 2 All material objects including the brain can be divided.

Conclusion Therefore, the mind is not the brain (or any other material object).

To assess it, the first thing we have to do is to get clearer about what is meant by saying that the mind is indivisible. There are two ways in which it might be impossible to divide something:

- Into two or more whole examples of the thing
- Into parts

When Descartes said that the mind is indivisible, he may have thought that both kinds of dividing are impossible. It now seems likely that he would have been wrong about splitting a mind into two or more whole minds. The evidence is difficult to assess, but consider cases of dissociative identity disorder (what used to be called ‘multiple personality disorder’) and patients who have had brain bisection operations (more properly known as hemispherectomies). In both kinds of cases, it could be argued that one mind has split into two (or more) minds. (See the Suggested Further Readings for more information about multiple personality and hemispherectomy). If so, both kinds of cases at minimum put pressure on the idea that the mind cannot split into two or more whole minds. Because the possibilities concerning the division of minds into two or more *whole* minds are at best unclear, we will focus on P1’:

Premise 1’ The mind cannot be divided into parts.

Here is an argument for P1’: A chair can be divided into parts. When we do so, the various parts are separated from the object (and from one another) but continue to have the properties that they contributed to the object when they were joined to it.

- (5) *A part of x* Something that can be separated from *x* and yet retain its integrity as a distinct unit whose role in the original system remains clear.

Now, the parts of a mind would presumably be things like the thinking apparatus, the emotions, the language-processing unit, perhaps the sense of humor, and so on. And the suggestion is that these cannot be separated from the whole mind and yet retain their integrity as distinct units. Why? If any of the mind’s aspects were to be taken from it, *that aspect would cease to exist; it would have no existence apart from the mind*. Think of such aspects as a sense of humor, an ability to think through hard problems, memory, imagination—none of these could exist apart from the mind of which they are an aspect. By contrast, a leg separated from a chair does continue to be a distinct unit, namely, a chair leg. Therefore, aspects of mind are not parts, not in the sense that legs are parts of a chair.

P2’ is straightforward:

Premise 2’ The brain can be divided into parts.

P2’ also seems to be true. Items in the brain can be separated from it and yet retain their integrity as distinct units with the properties that gave

them their roles in the original system (this happens in neuroanatomy laboratories every day). The conclusion would then follow:

Conclusion Minds are not brains.

So far so good, but perhaps 'so good' is not quite good enough. Rather than looking further into whether P1' and P2' are true and whether C follows from them, let us accept C for the sake of argument and ask a different question: So what if C is true? Would that give substance dualists what they want?

We think not. Here's the problem. Descartes and others like him want not only dualism but also *immaterialism*—that the mind and brain are not just distinct things (which we are granting for the sake of argument) but also that the mind is *not made out of matter*. It is the latter that is not at all established by this argument. Moreover, many dualists want *separability*—they want to show that the mind could go on separated from the body after death. That is to say, they want the mind to come out as not only an object distinct from the brain but also as an immaterial and separable object. This stronger conclusion is needed if there is to be any serious hope of immortality, and grounding at least a possibility of immortality has been an objective of most substance dualists (see section 6). Thus, even if the mind cannot be divided into parts, that would not be enough to prove that it is immaterial or that it is separable from the body.

Indeed, the (limited) kind of dualism proved by the indivisibility argument is found in all sorts of things that are obviously completely material. Think of radios. We cannot break a radio's functions into separate parts existing apart from the rest of the radio. If you try to separate out these functions, they cease to exist: if we remove the ability of a radio to pick out individual stations, for example, that tuning ability does not continue on outside the radio, pulling in our favorite stations. Similarly with computers. If we remove a computer's ability to check spelling, that function does not continue on somewhere outside the computer. If so, a radio and a computer are different from the hardware that makes them up, and we have just proved dualism for radios and computers! However, it is not a form of dualism of any interest to traditional substance dualists, because none of the "objects" in question can exist outside the matter that makes them up. In this sort of dualism, nothing need be either immaterial or separable.

The reason that radios, computers, etc., are more than the hardware out of which they are made is that to be a radio or a computer, a thing must have certain *functions*. That is true of many kinds of things; for

many kinds of objects, to be an object of that kind, something must have certain functions. Such functions cannot be separated from hardware as computer chips, dials, drives, etc., can be, yet they are still crucial to the object's being the kind of object it is. That is what explains our dualism, not any immateriality in the units.

Minds and brains are exactly parallel. Minds are more than the “wet-ware” of the brain; they also consist in certain functions. That is why minds cannot be split into parts, while brains can. It has nothing to do with minds being made out of anything immaterial. Curiously enough, the indivisibility argument, which philosophers accepted for hundreds of years as one of the most profound and convincing arguments for dualism and immaterialism, turns out, when closely examined, to be really a new argument for functionalism!

4 What Should We Believe?

If none of the arguments for dualism works, i.e., if none of them gives us any reason to believe that the mind is different from the brain or that mental states and events are different from brain states and events, what should we believe? There is a way of settling the question that will look very dubious to some people at first but seems to us entirely valid and also very powerful. This is the notion of BURDEN OF PROOF.

In some cases in philosophy it is important to know who has the burden of proof and the dualism/materialism debate is one of those cases. The idea of burden of proof can most easily be illustrated by reference to criminal and civil law. In criminal law, the burden of proof is on the prosecutor. The prosecution must prove the defendant guilty (beyond a reasonable doubt). The defendant, by contrast, does not need to do *anything* to establish his or her innocence. Indeed, if a defense lawyer thinks that the prosecution has not proven its case, the defense can ask for the case to be dismissed without presenting any evidence. The burden of proof is on the prosecution. In a civil case, by contrast, the burden of proof is equally on both parties. A balance of probability is all that is required to win or lose a case. The side most likely to be right on the balance of probabilities wins the case.²

Where does the burden of proof lie in the dualism/materialism debate? To answer this question, we need to bring back OCCAM'S RAZOR, which we first introduced in chapter 3. Occam's razor is the principle that we should not multiply entities needlessly. What this means, in more straightforward

terms, is that you should not believe in something unless you have *some reason* to believe it.

Let us give an example. Suppose that Rob walks into a room and says, "Are those leprechauns ever lively!" Andy says, "I don't see any leprechauns." "Ah," says Rob, "leprechauns are not the kind of thing you can see." So Andy gets a very fine screen and passes it over every cubic inch of the room. He says, "The screen didn't touch anything, so there can't be any leprechauns in the room." Guess what Rob will say: "Ah, but leprechauns can't be detected with a screen." And so on. It won't take you long to conclude that Rob has no evidence of any kind for there being leprechauns in the room. And from this you will quite rightly infer that neither Rob nor anyone else should believe that any leprechauns exist in the room. Why? The operative principle is this:

- (6) Excellent evidence that something does not exist is a complete lack of evidence that it does exist.

Now apply Occam's razor in the same way to dualism. Materialists and dualists agree that matter exists, so we don't need any argument for that. The dualist wants to maintain, however, that something else exists, too: something immaterial or some kind of nonneural property. The burden is on the dualist to produce some reasons for believing that this something else also exists. If we can find no evidence for anything more than matter, the only rational thing to believe is that nothing exists except matter, which is the materialist's position.

So what should we believe? For anyone who accepts Occam's razor and also finds that none of the arguments for dualism gives us any evidence that persons are made up of something more than matter, the answer is obvious: we should believe materialism. We have lots of reason to believe that we are *at least* material objects, so if none of the arguments gives us any reason to believe that we are made up of *anything more* than matter, we should believe that we are simply made up of matter, that we have no immaterial element.

That leaves us with one last question: which form of *materialism* should we believe? The two options remaining from our discussion in chapter 4 were functionalism and neurophilosophy. Obviously, functionalism is the more immediately plausible of the two, but plausibility is not enough. What we need is something that determines which is more likely to be true, not merely which is initially more plausible.

Here Occam's razor is of no use to us. Both parties believe in the same number of kinds of stuff, namely one, matter, and both parties agree that

every mental event is also an event in the brain: there is no multiplying of substances *or* properties in either theory. What separates functionalism from neurophilosophy is that functionalists believe that this mind/brain is a bunch of symbol-manipulating powers that have intentionality, and neurophilosophers deny these claims. Having intentionality is one more property than neurophilosophers want to allow, but it is not clear that Occam's razor counts against positing this property. Anyway, functionalists think they have good reasons to posit the property of intentionality, and Occam's razor gets a start only if there is *no* argument or evidence for something. If so, Occam's razor is not going to help us here. In fact, it may very well be the case (and this is a crucial point) that there is nothing distinctively *philosophical* about the disagreement between the two theories at all. The disagreement may be a straightforwardly EMPIRICAL QUESTION:

- (7) An *empirical question* is a question that requires evidence from observation and experimentation and cannot be answered by argument or analysis alone.

In the case of functionalism and neurophilosophy, the most important "evidence from observation and experimentation" that we need is evidence about which approach will result in the most comprehensive, satisfactory theory of what we currently call the human mind.

Such facts as we currently have available pull in both directions. Cognitive psychology, indeed cognitive science generally, generally uses the language of intentionality. Functionalism advocates using the language of intentionality to do psychology, while neurophilosophy rejects it as useless. So the huge success of cognitive psychology provides some support for functionalism. On the other hand, the neurosciences have leapt forward in the past decade, generating new discoveries by the hundreds every week. To the extent that the success of this work is leading theoreticians away from the traditional conception of the mind as a system best described using the language of intentionality, this success lends support to neurophilosophy. In our view, this is as far as we can go at the present time. Which picture of the mind is the best one is an issue, we think, that cannot be settled by philosophical analysis and argumentation. We will return to this issue in chapter 8.

Just to complicate matters even more, we should recall once again the issue of our ever changing and ever more complicated conception of matter. As we argued at the end of section 3 of chapter 4, a *sufficiently* large change in our understanding of what matter is like could conceivably cause us to rethink the materialist conception of the mind from the

ground up, especially if the new conception of matter turns out to be more like the classical dualist and idealist conception of mind than our current conception of matter. If that happens, who knows whether any of our current ideas about functionalism and neurophilosophy would survive.

5 Is the Mind Fundamentally a Symbol Processor?

One thing a computer can do that most humans can't is be sealed up in a cardboard box and sit in a warehouse.

Jack Handey

Recently a new, deep-running issue has appeared that opens the chasm between functionalism and neurophilosophy even wider. Recall that eliminative materialism started off as essentially a critique of folk psychology and a recommendation to look to neuroscience. The move to the deeper issue that we are about to introduce is a large part of what the transformation of eliminative materialism into neurophilosophy has consisted in.

The new issue is nothing less than the fundamental nature of the human cognitive system. (The COGNITIVE SYSTEM is the system that processes information about the world, itself, one's body, etc. 'Cognitive system' is the successor term in contemporary cognitive research for what used to be called 'the mind'.) As we saw, functionalists ground their model of the mind in information-processing functions described and explained using the concepts of "folk psychology," i.e., concepts that attribute intentionality. Most functionalists hold that these functions that make up the mind are primarily symbolic processes. A SYMBOLIC PROCESS is one that starts with strings of symbols—strings of words structured by the syntax of a language, for example—and transforms these strings of symbols into other strings of symbols according to rules. This is very much the way traditional computers work; the COMPUTER METAPHOR plays a considerable role in contemporary functionalism.

The idea that the mind is a symbol-processing system is not the only possibility, however. Neurophilosophers tend to hold that the mind does its cognitive processing nonsymbolically. Symbols only appear at the end of the process: we translate cognitive results into symbols when we want to communicate them to others in speech or preserve the results in written form.

This new debate, about whether the human cognitive system is fundamentally a symbol-processing system or not, has been added to the old debate about the scientific usefulness of folk psychology. Functionalists

urge that cognitive activity is symbol manipulation a long way down. There is nothing more basic than symbol manipulation (except implementation in assemblies of nerves). Neurophilosophers reject this idea and claim instead that the actual processes that perform cognitive tasks do not use symbols; symbols enter only after the information-processing is over, at the stage, as we said, when we want to write the results of the processing down or when we want to communicate them to others.

This disagreement leads to a deeper form of the old disagreement over whether such a thing as a mind exists. Both parties start from a common conception of what a mind would have to be like: it would have to be a symbol processor. If so, then if human cognitive systems are symbolic almost all the way down, minds exist. If, on the other hand, cognitive activity is fundamentally non-symbolic, with symbols entering only at a late-stage translation process, then the human cognitive system is not at all like what we normally think of as minds, and it would be clearer and more precise simply to say either that minds do not exist or that the mind, in the traditional sense, only enters as a late, superficial aspect of human cognition.

If most of the activities of the human cognitive system are not transformations of strings of symbols into other strings of symbols, then what are they like? We can only touch on this big question. But the basic idea runs as follows (see Churchland and Sejnowski 1992). The fundamental role of cognition, say the neurophilosophers, is to turn the input of the senses into control of motor output: to turn perception of food into actions to get the food, perception of a tiger into actions to avoid the tiger, and so on. Now, when someone sees a ball coming at her head and ducks, it is very unlikely that it is *strings of symbols* or *rules for transforming symbols* that transform the information from her eyes into the contracting and stretching of her muscles. For one thing, simpler animals like mice are very good at ducking, avoiding danger, etc., yet they very likely have no strings of symbols in their brains at all. Next, they point out, language and all other activities that have a *clear* symbolic structure are late arrivals on the evolutionary scene; perception and motor control arrive far earlier. Indeed, few animals even have a structured symbolic system like language. But the bulk of human cognition is likely to be similar to cognition in other animals. If so, it is likely that the bulk of human cognition is nonsymbolic and that symbolic processing is merely a late evolutionary addition stacked on top of an already-existing non-symbolic system. (If this sounds like a replay of the debate in chapter 3 over whether thought is inner speech, well, it is. What neurophilosophers

add is that, because “thought” *isn’t* inner speech, not even inner speech couched in the language of thought, minds as we usually conceive of them do not exist.)

Of course, language and other symbolic processes turned out to be absolutely crucial to human development. If we had no language, then we would have no books, no bridges, no science, no modern medicine—and no philosophy texts. In short, we would have no life as we know it. That would explain why theorists have emphasized language to the virtual exclusion of everything else in their models of the mind. But it does not follow from this that they are right. For all its importance to us, symbolic processes may indeed be merely a superficial layer in our cognitive processing. That is the challenge with which neurophilosophy confronts our orthodoxies, and it is one of the great intellectual challenges facing current research on the mind.

In the mid 1980s, theorists devised a new model for how nonsymbolic processing might work. They called such systems CONNECTIONIST SYSTEMS or NEURAL NETWORKS. Unlike traditional serial computers and traditional models of cognition, connectionist systems do not contain discrete, separate symbols. Nor do they contain explicitly represented rules for transforming symbols. Yet they can do highly significant cognitive work.

The difference between the connectionist picture and the traditional symbolic picture is very deep. For connectionists, thinking is fundamentally a process of associating properties with other properties, much like the old empiricist idea of the mind as a vast system of associations (introduced in chapter 2). On the symbolic picture, the mind works by following out the implications of sentences and sentencelike structures. This is closer to the old rationalist picture (also introduced in chapter 2), though with more room left for the input of the senses. One of the crucial questions for connectionism is, Can connectionism model language? Recall our discussion of the relationship between language and thought (chapter 3). Even if a lot of thought is nonsymbolic, as neurophilosophers maintain, language is still central to some of the most important kinds of thinking. So if connectionism cannot give us an account of language, then connectionism cannot be the whole story about the human mind. (For more on this fascinating new debate, see Rumelhart and McClelland 1986, where it all started, Fodor and Pylyshyn 1988, Clark 1989, and Horgan and Tienson 1996.)

Time to sum up. Like the general question of whether functionalism or neurophilosophy as a whole is a better picture of human cognition, the specific question of whether the symbolic or nonsymbolic picture of

human cognition is the better one is in the end an EMPIRICAL QUESTION, and we do not yet have the facts we need to settle it.

However the symbolic/nonsymbolic debate turns out, it is one of the most interesting examples of how philosophical work on cognition is being penetrated and informed by work from other cognitive disciplines. Left to themselves, it is unlikely that philosophers would ever have invented connectionist models at all or seen more than a tiny fraction of what is built into the debate about the mind as a symbol processor. This new meeting of “minds” (how else can we put it?) between philosophical and other approaches to cognition will be a main theme of chapter 8. When we return to the issue of functionalism versus neurophilosophy there, we will tip the scales very tentatively and very slightly in favor of functionalism.

6 Do We Need to Study the Brain?

I think a good movie would be about a guy who's a brain scientist, but he gets hit on the head and it damages the part of the brain that makes you want to study the brain.

Jack Handey

The debate between functionalism and neurophilosophy raises a further issue that we haven't examined yet: do we need to study the brain to understand the mind? Even if the rational thing to do is to accept some form of the materialist theory of mind, the two leading forms of materialism, functionalist and neurophilosophy, tend to differ strongly over the importance of studying the brain. Here's why.

If, as functionalists maintain, types of mental states and events can never be identified with, or reduced to, types of brain states and events (in the sense of (6) on p. 67), then it would seem that so far as understanding cognitive function is concerned, it should not matter whether we know much about how the brain is built or not. That was indeed what functionalists thought for many years: in principle, they held, we can understand the mind while knowing virtually nothing about the brain. One of the slogans of functionalism is that function does not determine form—a given function can be performed by many different kinds of systems. The job is to understand the functions that make up the mind. The details of the particular apparatus performing them (brain, computer, whatever) do not matter. A cognitive-psychologist friend of ours recently put it this way: “I don't care whether cognitive functioning is done in the brain or the liver!” This view or one like it was held by many theorists

in the heyday of classical ARTIFICIAL INTELLIGENCE, when researchers believed that we would soon have computers that could perform mind functions nearly as well as we can. Such optimism about how quickly we would understand cognition proved to be wildly exaggerated, be it artificial cognition or “natural” cognition (natural cognition, in case you are wondering, is us).

Neurophilosophers and their scientific fellow travelers took exactly the opposite view, of course. Since study of the brain using the language and techniques of neuroscience is the *only* way to understand what we now misleadingly call “the mind,” study of the brain is absolutely crucial. And neurophilosophers follow their own advice. Many of the great centers of cognitive studies in Europe and the United States have affiliated themselves with neuroscience research centers.

More recently, functionalist philosophers of mind have also started to change their minds on the issue of whether the study of cognition (cognitive science) can be done without studying the brain (neuroscience). For one thing, *how* a task is performed tells us a great deal about precisely *what* task is being performed. Nevertheless, a great deal of the work on cognitive functions in cognitive psychology and linguistics still takes place with very little knowledge of the brain.

7 Why Isn't Functionalism Just Good Old Dualism in New Clothes?

Let's close this discussion of functionalism and neurophilosophy with a question that will have occurred to many readers: Why is functionalism a *materialist* theory? Why isn't it just property dualism in a new guise? (If it were and if no arguments for dualism are sound, then we would have to reject functionalism too, of course.) Let us come at this question by once again turning to *psychological explanation*.

Neurophilosophers reject psychological explanation—the search for reasons, goals, purposes, etc.—as a dead end. They want us to focus on mechanistic explanations of how the brain works. Functionalists, by contrast, embrace psychological explanation, as we saw. Indeed, the central figure of contemporary functionalism, Jerry Fodor (1975 and elsewhere), argues that psychological explanation is currently the only game in town. That is to say, the only way we have to explain ourselves and others currently is in terms of beliefs and desires, in terms of our reasons for doing what we do. And the only way to describe our reasons for what we do is in the language of intentionality that neurophilosophy wants to dismiss as scientifically useless.

If this is what functionalists believe, how exactly do functionalists and property dualists differ? The obvious answer is that functionalists hold that there are token-token mind/brain identities. That is to say, even though they deny that any *type* of mental state or event (e.g., all beliefs about the wetness of water) can be identified with, reduced to, any *type* of brain state or event (e.g., circuits of such and such type), they hold that every *token* (or occurrence) of a mental state (or event) is nothing more nor less than some brain state (or event). Thus, even though they do not believe in the identity theory of the mind/brain (as defined in (7) on page 67), they are still materialists. So they say. But what kind of materialism is it that insists that the types of states and events that make up the mind can never be identified with types of states or events in the brain (or even the brain plus)? This looks suspiciously like some kind of mind-brain dualism.

If functionalism looks suspiciously like a form of dualism, what form of dualism? Certainly not *substance* dualism. The substance dualist believes that persons are made up of two very different kinds of things. No functionalist believes that. Functionalists believe that the mind/brain is one single system. At the other end, functionalism *embraces* explanatory dualism, so it definitely is this form of dualism, but explanatory dualism is also compatible with materialism, so there is no risk to functionalism as a form of materialism from that direction. The only form of dualism left is property dualism, so what the question boils down to is this: is functionalism just a new form of property dualism?

Let us come at this question via explanatory dualism. Explanatory dualism does not maintain that the *mind* is a duality of any kind. It maintains only that we have a duality in the *explanations* that we can give of the mind. This is why explanatory dualism is entirely compatible with materialism. Now, it is hard to see why the need for two kinds of *explanation* would force us to postulate that minds have two kinds of *properties*. Consider some analogies. To explain the growth of plants, we need explanations of both how they get water and how they photosynthesize carbon, yet no one advocates a “property dualism” for plants of any philosophically interesting kind. Similarly, many linguists hold that we need a different kind of theory to explain the syntax of a sentence from what we need to explain the semantics, the meaning, of a sentence, yet there is only one language being explained. (These comparisons are a nice example of the way analogies work in philosophy, by the way.) Thus, from the fact that we have two ways of explaining human behavior, it does not automatically follow that human beings have two radically different kinds of *properties*. And similarly with functionalism.

Even though functionalists insist that kinds of mental event cannot be identified with or reduced to kinds of events in the brain, they still insist that every mental event is some brain event. In the jargon we introduced earlier, they reject type-type identities but insist upon token-token identities. Property dualists, by contrast, hold that mental events are not neural events at all; the properties of the mind/brain that are mental are completely independent of physical properties of the mind/brain. Thus functionalism is not property dualism.

This difference between the materialism of functionalists and the dualism of property dualists tends to go with a deep disagreement about the nature of mental properties. Functionalists view them as perfectly straightforward properties of the world, open to scientific exploration and explanation like any other. We may need to use a special kind of explanation to understand them, namely psychological explanation (as described in chapter 4 and discussed above), but many kinds of phenomena that cannot be easily reduced to brain states need their own kind of explanation: color vision and phonetic processing are two examples. By contrast, property dualists tend to think of mental properties as deeply mysterious in nature, certainly not the sort of thing that could be studied scientifically. Indeed, many property dualists think that the only way to observe mental properties is by introspection. As we saw earlier in this chapter, however, most scientific-minded theorists reject introspection as a viable way to study the mind. In short, for all their superficial similarity, the differences between functionalists and property dualists run very deep.

8 Immortality Again: Can We Look Forward to Life after Death?

Sometimes I think I'd be better off dead. No, wait. Not me, you.

Jack Handey

We've spent a lot of time talking about the nature of mind and its relation to matter. We will end this chapter by considering an issue that motivates many people to *care* about whether the mind is or is not simply processes in matter. The issue is immortality; we introduced it near the beginning of chapter 4.

(8) Can (and will) we continue to exist after the death of the body?

More specifically, the issue is PERSONAL IMMORTALITY, which holds that a person will continue to survive *as the person she is now* after the death of her body. So the specific question of interest is (9):

(9) Can (and will) I continue to exist *as me* after the death of my body?

One of our two conclusions will be that there's little evidence that the answer is 'Yes', little evidence for an afterlife. Our other conclusion (you might call it our big conclusion) will be that, surprisingly, having little hope of a life after the death of our body shouldn't be very upsetting.

Here's the plan. We will present two different arguments, each of which tries to show that there is life after death. Neither of these arguments works very well. Since those who endorse life after death have the burden of proof, i.e., they need to give some *reason* for their view, those who don't believe in the afterlife have the most rational position if no other positive arguments are forthcoming. But things are worse than that for the afterlife proponent. For there are several compelling reasons for thinking that there is no life after death. (Each of these anti-immortality arguments harks back to the foregoing discussion of dualism and materialism.) The overall conclusion, then, is that the mind probably ceases to exist when the body dies. Having reached this conclusion, we will consider whether it is really so distressing.

One argument that you sometimes hear in favor of the afterlife is this:

Premise 1 If life truly ends at death, then life ultimately has no meaning.

Premise 2 It would be awful if life ultimately had no meaning.

Conclusion Therefore, life does not end at death.

This argument suffers from two very serious flaws. First off, it's merely *an argument from hope*. Sadly, arguments from hope are, well, hopeless. Specifically, they are invalid: the conclusion can be false while the premises are true. To see this, compare the following: Andy and Rob really wish they had a million dollars each; therefore, Rob and Andy do have a million dollars each. The premise of this parallel argument is true: we certainly *wish* we were millionaires. But, we can assure you, the conclusion is false. So, insofar as this is just an argument from hope, i.e., insofar as the arguer is merely *wishing* that the conclusion were true, it cannot succeed.³ Besides, there's a second flaw in the argument. Though P1 may look plausible at first glance, in the end it's not obvious at all. To see this, ask yourself *why* exactly must life continue to exist after the death of the body if it is to have any meaning. Couldn't life's meaning derive from the 75-odd years that the body is alive? Or couldn't the source of meaning be one's fellows and one's achievements during this life? Unless there's a good answer to these questions, P1 is also questionable. If so, this first argument, though quite familiar, is not very promising.

Here's a less familiar, but more intriguing, argument for life after death.⁴

Premise 1 If you cannot even imagine something, then it doesn't really exist.

Premise 2 You cannot even imagine the death of your own mind.

Conclusion Therefore, your mind never dies.

The argument's form is fairly simple. But the premises need to be explained.

P1 basically joins together two logical principles: first, that if you cannot imagine something, then it's not *possible*, and second, that if something is not possible, then it isn't *actual*, i.e., it doesn't in fact exist. (We saw the converse of this argument above: the existence of mind without body was said to be conceivable, and hence possible.) A parallel example may help here. You cannot even imagine a colorless red ball. That's because it's *impossible* for something to be both red and colorless. Now, given that colorless red balls are impossible, it's no surprise that there aren't any to be found in your closet. In fact, there aren't any to be found *anywhere*: impossible things, since they *cannot* exist, clearly *do not* exist. (Notice, in contrast, that certain things that do not exist are nonetheless possible things: Weeble has no offspring, so "Weeble's black kitten" isn't a real object; it is nevertheless a possible object. That is to say, Weeble might have had a black kitten.) Now P2 says, of your mind's death, that you cannot imagine it. Hence, by P1, there's no such thing as the death of your mind. This is a convoluted way of saying that you must continue to live after the death of your body.

Even if we grant P1, what supports P2? Essentially this: that whenever you try to imagine your absolute nonexistence, you fail. That is, your mind's death is wholly and utterly unfathomable. That's because you have absolutely no idea of—you cannot experience, even in your own imagination—your own total annihilation. To put it metaphorically: each time you try to imagine, from the inside as it were, what it feels like to not be, *you* are there. Hence what you imagine is not a situation in which your mind has really ceased to exist.

But this argument doesn't support P2, as T. Nagel shows. If it did, we could argue that because "you can't conceive of what it would be like to be completely unconscious, even temporarily" (Nagel 1987, 88), therefore you will never be unconscious. And that's absurd. You do become unconscious, every night. Again, you can't experience, even in your own

mind, total unconsciousness. That wouldn't *feel like* anything. But this cannot demonstrate that you will never be completely unconscious.

The argument is built on a fallacy of equivocation. It equivocates on 'imaginable' (recall the related problem with 'conceivable' in the conceivability argument). In one sense, you clearly *can* imagine yourself being unconscious. That is, you can think of yourself, from the outside, lying motionless on the street, having been knocked out. What you cannot do is *simulate* the sensation of unconsciousness. That's because there is no such sensation. Here's a comparison. Suppose that someone says, "Imagine riding a raft down that river." You can do this in two ways: by "experiencing the ride" from a rider's point of view, but without actually taking the plunge, or by picturing yourself, i.e., your body, on the raft, bolting down the rapids. In the case of consciousness, only the second kind of imagining is possible. That's because there is no *sensation* of unconsciousness, no "way that it feels." So in the first sense of 'Imagine yourself being unconscious', you can't do it. But, crucially, unimaginability *in this sense alone* doesn't show that people are never unconscious. To show this, being unconscious would need to be *wholly* unimaginable, which it clearly isn't.

Similarly, it is perfectly possible to imagine the nonexistence of your mind, in the sense of thinking of your lifeless body, from the outside (e.g., it's certainly possible to imagine your own funeral), though there's another sense of imaginable, the one where you "take on" the agent's point of view, in which you can't imagine what it's like for your mind not to exist. The question is, which sense of "imaginability" is at work in P1? Presumably, it's not just the restricted sense of 'imagine' in which 'Imagine that you're *x*' requires "experiencing" what it's like to be *x*, without actually being *x*. 'Imagine' in P1 is being used in the broad sense. But, if we read 'imagine' in this broad sense, P2 is just plain false, however initially appealing and intuitive the premise may be on a weaker reading of 'imagine'. Hence this argument for an afterlife also fails.

Remember the lesson drawn about the invisible leprechauns: someone who posits something needs to provide some reason for believing in it (Occam's razor). Thus, when no reasons are forthcoming, one may rationally reject the proposed entity without argument. So, as we saw, if there's no grounds whatever for believing in both mind and body, then the materialist (who believes in body only) wins. Similarly here: unless there is some reason (however minimal) to believe in the afterlife, the afterlife skeptic triumphs. Put otherwise, if neither side presents arguments, the

result is *not* a tie: the person who makes the claim about the extra kind of existence to the effect that there is an afterlife (who makes the EXISTENTIAL CLAIM, as philosophers call it) loses. Still, winning by appeal to burden of proof isn't very satisfying, so let's see if there are any arguments against an afterlife.

We will consider two arguments for the conclusion that death really is the end. The first is one that we have already introduced, in the first section of chapter 4: in the absence of a being with godlike powers able to reassemble and reanimate long-dead bodies (bodily resurrection), life after death requires, though it is not guaranteed by, dualism. If dualism is true, we can understand how life after death might be possible. On the other hand, if dualism is false and the mind is just the brain, then the mind clearly cannot exist when the body is dead (again, in the absence of miraculous intervention). The trouble is, as was argued above, dualism is likely not true. Even worse, though the truth of dualism seems like a *necessary* condition for life after death, it isn't a *sufficient* condition. Recall, for example, property dualism: here the continued existence of the nonmaterial mental properties clearly *requires* a body. For it is the body that houses the mental properties. Hence, for property dualists, if there's no body, there's no mind, i.e., when the body dies, so does the mind. Nor is it only the property dualist who will think that the continuation of the mind is body-dependent.

Consider this: the mind, even for the substance dualist, gets most of its stimulation from the body. Furthermore, the mind acts via the body. Given that so much of mental activity arises from bodily stimulation and so much of it is designed to contribute to bodily movement, some philosophers have concluded that the human mind is *essentially* embodied. That is, the human mind is radically unlike, say, the mind of pure intellect that some theorists would assign to God. Taking this seriously, however, it seems that the *human* mind could not exist without a body, even if the mind is a distinct substance. An afterlife hopeful might say, "Well, the mind would undoubtedly undergo changes once it lost all sensory input and once it ceased to have behavioral effects in a physical world, but it would still exist." This is a fair point. However, it raises an uncomfortable question: if your mind, considered as a separate substance that merely accompanies the body, changes *very* radically after death, would it be appropriate to say that *you* survive the death of the body? On this hypothesis it's true that *something* would still exist after your body died, something associated with your body. But being so drastically different from who you are now—i.e., an embodied, worldly, sensual being—

would that “something” be you? If the answer is ‘No’, then once again it turns out that the continued existence of you requires a body, *even if dualism is true*.

The argument as a whole may be summarized as follows:

Premise 1 If dualism is false, then there is no life after death.

Premise 2 If dualism is true but the continued existence of the mind requires a body, then there is no life after death.

Premise 3 Either dualism is true, but the continued existence of the mind requires a body, or dualism is false.

Conclusion Therefore, there is no life after death.⁵

The argument is valid. But is it sound? For instance, someone might well say, “There can be an afterlife, even if the existence of the mind requires a body, because we could continue to exist in *another* body.” Indeed, functionalism—as we have said repeatedly, a version of materialism—presumably permits this: if the mind is just a set of functions, a kind of program running on the brain, then your “program” could plausibly be transferred to another brain, or to a silicon-based computer for that matter. So it seems that some varieties of materialism are consistent with life after death. In which case, P1 is just false.

The afterlife skeptic won’t give up this easily, however. First, while functionalism apparently makes the after-death existence of the mind *possible* even without dualism, it doesn’t make it *actual*. In the world as it actually is, death is the end, because our “programs” *aren’t* stored elsewhere, not even by reincarnation; death occurs before the mind is “backed up” onto another physical medium. Materialism requires that there *always* be a physical medium. So the idea of your mind “passing into” another body, far away and years later, just isn’t credible. Besides, suppose your program *were* restarted in another body. Would this really be a matter of *you* surviving? It may seem that the answer would be ‘Yes’. But consider the following sort of case.

Imagine that you’ve gone into the Eternal Life Center to have your body rejuvenated and your mind transferred into the fixed-up body. You climb onto the table, hear some whirring sounds, and then the lights go out. As you descend the table, an embarrassed attendant explains that there’s been a slight glitch. He tells you, “The way the technology normally works is this: a new body is created, the information from your brain is put into its brain, and your old body is then destroyed. The problem is, though a new body was created and your program was put into it, the power unfortunately went out before the old body (i.e., you!)

could be atomized.” Now, they can’t let two of you leave the center. So the attendant makes a simple request: “Please return to the table, so we can destroy the old body.”

Many people would resist this request. And yet purportedly (the rejuvenated) “you” is already in another room, so “you” will still be alive after the troublesome old body is gone. Why not agree, then, if transferring your “program” is sufficient to achieve continued life? The answer seems to be that having your mind “moved” to another body may, despite initial appearances, actually be a way of *dying*, not a way of continuing to live. In that case, P1 really is true: without substance dualism, there can be no life after death. Since substance dualism likely isn’t true, it seems probable that the mind dies when the body does. (Puzzles such as this, about what philosophers call *personal identity* are well discussed in Parfit 1984.)

We end by considering whether this is such an awful conclusion. First, an obvious point: it might be *boring* to live forever. Even if you would like to live longer than the usual 75 years, you probably would not want to *never* die. Also, death sometimes ends great suffering, whether physical pain or mental depression or something else. So death is not always, no matter what the circumstances, bad. On the other hand, is death *as it usually occurs* good? After all, as a matter of fact it does not typically cut off extreme pain or boredom. On the contrary, the deceased typically misses out on good things (or, more precisely, things which *would* have been good). And death ordinarily comes long before extreme boredom sets in. So death isn’t usually a good thing, though there are circumstances in which it might be good. Instead, death, by depriving us of continued enjoyment of life, is usually an unwelcome fate.

Some people go further than that, however. Some people think that death is *really* bad. Indeed, they find the prospect of dying to be absolutely terrifying. Is this a reasonable attitude? By dying, one is likely to miss out on some pleasures, and this may be a bad thing. But death should not be *scary*. This for two reasons. First, “the time after life” is just like the time before birth, on the assumption that there is no afterlife. But the time before birth was not awful. So, because the “beforelife” and the afterlife are basically the same, you should not expect the time after death to be awful. Indeed, death, we are supposing, is the end: there is nothing left after the body dies. But you cannot reasonably be afraid of *nothing*. There is no pleasure, it’s true. But there’s also no pain, no suffering, indeed, no displeasure of any kind. Hence there is nothing to be afraid of.

As Nagel (1987) says, death is something to be afraid of only if we will survive it.

The overall conclusion, then, is that there is insufficient evidence to justify belief in an afterlife. Hence we should not believe in it. But this should not frighten or depress us, because, though death is a little bad (the good times stop), it is not *really* bad.

Study Questions

1. Contrast the manifest image and the scientific image. Why does the manifest image point to dualism and the scientific image to materialism? Could they be combined? If so, how?
2. Are there problems with introspection that make it a poor source of information to use in science? If so, what are they? If not, why not. (It certainly appears that there are problems!)
3. Explain each of the four arguments for dualism given. Which ones use introspection? Which don't?
4. Do you think that any of the arguments for dualism work (that is, do they prove dualism)? If so, which one, and why? (If not, skip this question and go on to the next one.)
5. Outline the responses to these four arguments given in the text. Do you think that the dualist has any effective response to any of the reasons given for rejecting the arguments for dualism?
6. Can you think of any arguments for dualism of your own? How would a materialist argue against you?
7. How does the notion of "burden of proof" play a role in the dualist-materialist debate? What is Occam's razor? How deeply does it cut against certain claims?
8. Why is functionalism a version of materialism, while property dualism is not?
9. What sort of life changes would result in a loss of personal identity over time? Wouldn't death count, if anything does? If that's right, what would it be for anyone to survive death as the same person?
10. Why is it harder to believe in life after death if you are a materialist? Are explanatory dualists and property dualists any better off than materialists in this regard? Explain.

Suggested Further Readings

As we noted at the end of chapter 4, arguments for dualism may be found in Descartes 1931 [1641], property dualism is defended in Strawson 1959, and Hobbes 1996 [1668] and La Mettrie 1994 [1748] are especially important early (and rare) statements of materialism. Sellars 1997 [1956] introduced the distinction between the manifest image and the scientific image. P. M. Churchland 1984, Flanagan 1984, Jacquette 1994, and Kim 1996 are good overviews of the dualism-materialism debates. Jacquette 1994 espouses property dualism.

An interesting paper that both reviews the phenomena associated with brain-bisection operations (hemispherectomies) and discusses the problem of how to think about these phenomena is Nagel 1971. Marks 1980 is a short book entirely devoted to the topic. A long, controversial, but very readable account of dissociative identity disorder (what used to be called 'multiple personality disorder') can be found in Hacking 1995.

The new debate over connectionism can be followed in Rumelhart and McClelland 1986, Fodor and Pylyshyn 1988, Clark 1989, and Horgan and Tienson 1996.

For more recommendations, see the Suggested Further Readings at the end of chapter 4.