3
Materialism

Although Cartesian dualism is today a minority view in the philosophy of mind, that should not blind us to the enormous influence Descartes has on contemporary thinking about the mind-body problem, and particularly on materialism. I say this not only because materialists are explicitly guided by an animus against Descartes's dualistic metaphysics, but also, and just as significantly, because they are at least implicitly guided by a commitment to certain other, distinctly Cartesian, assumptions. Descartes believed that the world consisted of two basic kinds of substance: thinking substance and extended substance, *res cogitans* and *res extensa*. The modern materialist rejects the former, but endorses the latter. Descartes was, it is thought, at least half right: his *res cogitans* is, by the materialist's reckoning, a fiction, but his *res extensa* most assuredly is not – indeed, it constitutes the whole of what a human being is.

To be sure, Descartes's concept of matter as essentially "extended" cannot be maintained without qualification given developments in modern physics, which hold that certain fundamental physical particles are best conceived on the model of unextended mathematical points. Nevertheless, his notion that the physical world constitutes a vast "machine," with material objects – including the human body – being but smaller machines operating within it, has come to dominate the thinking of modern philosophers and scientists alike. It has become a hallmark of intellectual life in the post-Cartesian period that *understanding* something is thought paradigmatically to involve taking it apart and seeing how it works, the way one would understand any

mechanism. A physical thing, on this model, is like a clock, the operation of which can be grasped by determining how each part interacts mechanically so as to generate the behavior of the whole. Nowadays, this approach to inquiry may seem to be just obviously correct, the epitome of "thinking scientifically." Yet, as we will see later on, it constituted a dramatic departure, both scientifically and, more significantly for our purposes, metaphysically, from the assumptions that prevailed in most ancient and medieval thought – a departure that in many respects can be said to have created the mind-body problem as we know it today. That problem is thus as much an artefact of the points on which materialists and dualists agree as of those on which they do not. We will in due course be examining more carefully the nature – and the ultimate plausibility – of this approach to understanding the material world, shared by Cartesian dualists and materialists alike. The question at hand is whether, where the mind-body problem is concerned, that approach favors its materialist advocates over its dualist ones.

Tables, chairs, rocks, and trees

It is certainly no mystery *why* the approach in question has come to seem obviously correct. Modern science has, to all appearances, been one long success story, a success made possible in large part because of its commitment to the mechanistic model of the world. The behavior and properties of the ordinary middle-sized objects of everyday experience – tables, chairs, rocks, trees, water, metal, as they burn, melt, freeze, reflect light, exhibit magnetism, conduct electricity – have been explained in great detail via physical and chemical theories of extraordinary predictive power, whose application has made possible the breathtaking technologies of the modern world, technologies that would have seemed magical to earlier generations. These theories have revealed the existence of a micro-level of physical reality – a realm of molecules, atoms,

electrons, protons, quarks, etc. – which our ancestors would have found equally marvelous, and they have also proved themselves applicable to, and revealed the unexpected vastness of, the macro-level of the universe – solar systems, galaxies, galaxy clusters, and the very fabric of time and space. Most relevant to our present concerns, they have proved successful in explaining the operation of the human body and its various subsystems, opening the way to the healing of diseases that have plagued humankind for millennia, the extension of longevity through medicine and the use of artificial organs, and even the assisted or artificial reproduction, through laboratory means (*in vitro* fertilization and cloning), of life itself.

It is no surprise, then, that many philosophers have taken the view that the human mind ought also to be explicable in terms of the same sort of mechanical account to which the rest of the universe has apparently yielded. This view is more or less what is meant by "materialism" – the theory that reality, or (when the term is used specifically to denote a position in the philosophy of mind) at least human reality, consists of purely material or physical objects, processes, and properties, operating according to the same basic physical laws and thereby susceptible of explanation via physical science. There is, in short, no such thing as immaterial substance, or soul, or spirit, nor any aspect of human nature which, in principle, elude explanation in purely physical terms. The mind is, paradoxical as it may sound, entirely material. (It is material, that is to say, *if* it exists at all, and there are a few radical materialists who are of a mind to deny that it does. But more on them later.)

Materialism is also sometimes referred to as *physicalism* or *naturalism*, though these terms are occasionally used by philosophers to denote views which are intended to be distinguished from materialism. This confusion in terminology is, in a way, entirely appropriate, for the materialist thesis is by no means as evident or clear-cut as it might at first appear.

Modern physical science's success in explaining the tables, chairs, rocks, and trees of everyday experience is not the only source of materialism's intuitive appeal. There is also the fact that such ordinary physical objects seem to be paradigms of what counts as *real* in the first place. If we can see, hear, taste, touch, and smell something, we know for sure (barring Cartesian evil spirits and dreams) that it exists. Conversely, our failure to provide observational evidence for something typically leads us to doubt its existence. But then, it seems that we ought to be suspicious of any claim that something other than the objects, processes, and properties of everyday experience really exists, at least if the very existence of these everyday objects, processes, and properties themselves doesn't point to the existence of some other kind of thing. Modern science has given us good reason to believe that these everyday objects, processes, and properties are constituted of the micro-phenomena described by physics and chemistry, and that they in turn constitute the macro-phenomena described by astronomy and cosmology. So we are justified in holding that such micro- and macrophenomena also exist, even though they are, in general, not directly observable. But science gives us no reason to believe that entities such as ghosts and poltergeists are real; the evidence for such things is weak, and easily explicable in more mundane terms (hallucinations, delusions, tall tales and the like). It also appears to give us no reason to believe in such things as souls or Cartesian immaterial substances. The reasonable conclusion would thus seem to be that there just are no such things. At the very least, materialists hold, we have every reason to act on the assumption that there are not, and to expect to be able to explain mental phenomena entirely in terms of the operation of physical processes and properties.

But while such considerations may give the appearance that materialism represents (as dualism claims to do) nothing more than the drawing out of the unavoidable implications of some homespun common sense, appearances are in this case deceiving. For

scientific explanations have a way of not only *explaining* what we observe in everyday experience, but also, to a very great extent, explaining it *away*. The table in front of you seems absolutely solid and impenetrable, as unlike a cloud as anything possibly could be. Yet physics tells us that a cloud, of sorts, is *exactly* what it is – a cloud of unobservable particles, each occupying less space than exists between them, so that the apparently solid and impenetrable table is mostly empty space. We take our senses to give us as much certainty as it is possible to have, and so we base our science on them. But science then informs us that our senses are largely wrong. The world revealed to us by sight, hearing, taste, touch, and smell – the world of tables, chairs, rocks, and trees – is not the touchstone of reality; that honor goes to the strange world of unobservable entities postulated by physics – the world of molecules, atoms, electrons, and quarks. What becomes, then, of the commonsensical idea that the physical objects of everyday experience are the paradigms of reality? (And if what the table *really* is is something we *don't* directly observe – a cloud of particles – then why ought we to be so suspicious of claims to the effect that certain other unobservable phenomena – souls or Cartesian immaterial substances – exist as well?)

Reduction and supervenience

As the example above illustrates, modern science also tends, in the view of many materialists, toward what is often called *reductionism*: the table is sometimes said to be "reducible to" or in reality "nothing but" a collection of particles, with the appearance of it being something other than that dismissed as an illusion. The various *properties* of the table are also reduced: what solidity it does have is said to be nothing but the state its molecules happen to be in when the field of force they generate repels those fields of force associated with other collections of particles (your hands, or the

book lying on the table). Similarly, the solidity of an ice cube is nothing but the state water molecules are in when at freezing point, while the liquidity exhibited by water at higher temperatures is nothing but another state of its molecules. The temptation is to suppose that *everything* real – not just tables and ice cubes, but planets and galaxies, animals and human minds – must in some way be entirely reducible to the basic categories of physics: in some sense a planet and a mind are nothing but different kinds of configurations of molecules or atoms. The sort of materialism that makes this boldly reductionist claim is often labeled *physicalism*, the idea that basic physics reveals to us what is truly real.

The trouble is that there are things it is *very* hard to reduce down to the categories of physics in this strong sense, as most physicalists themselves will acknowledge. Cultural artefacts provide obvious examples: what makes a dollar bill the kind of currency it is seems to have little to do with the specific physical properties involved – a silver dollar is just as much a dollar as a paper one – and everything to do with social conventions, which are themselves hard to reduce to the properties of molecules in motion. Of course, all such cultural and social phenomena are ultimately mind-dependent; and the mind itself is the most notorious (and, for our purposes, relevant) example of something it seems hard to reduce to the physical, for reasons sketched in chapter 2, which we will be exploring in greater detail in the next few chapters. Moreover, physics is by no means a finished project, with the basic constituents of the material universe, and the laws governing them, all accounted for and neatly catalogued. The physics of Einstein and Heisenberg differs radically from that of Galileo and Newton, and the physics of the future may differ from both in radical ways. So in *which* physics exactly is everything real supposed to be reducible? Physicalists often reply that it is the categories of a *completed* physics – whatever body of theory future scientists will develop to solve all the problems current physics has yet to solve – that will do the job. But what if this future physics ends up having to postulate

immaterial or non-physical properties to account for mental phenomena, as some dualists have argued it will (for reasons we will be exploring later)? In that circumstance, physicalism would turn out to differ not at all from dualism – in which case, it would not be a version of materialism at all.

Such problems with physicalism have led other materialistically inclined philosophers to reject strict reduction as essential to their position and to opt instead for the notion of *supervenience*. One thing "supervenes" on another just in case there could not be a difference in the first without there being a difference in the second. Materialism can accordingly be understood as the claim that all *real* objects, properties, and processes, including those of the mind, supervene on purely *physical* objects, properties, and processes: nothing that happens, and in particular nothing mental, can happen at all unless something happens at the purely physical level, and ultimately at the level of the most fundamental entities postulated by physics. Unlike reductionism, this need not entail that the basic entities are, in some sense, all that "really" exist: perhaps there is a sense in which tables, chairs, rocks, trees, bodies, brains, and even minds are every bit as real as fundamental physical particles. It entails only that everything that happens at the level of tables, rocks, minds, etc. ultimately happens only *because* something happens at the level of fundamental particles. Some philosophers who are committed to the idea of the supervenience of the mental on the physical prefer the label *naturalism* to physicalism, the idea being that it isn't necessarily just the basic entities postulated by physics that constitute reality, but rather the natural world of material phenomena in general (as distinguished from purported *supernatural* phenomena, for example, Cartesian substances, angels, or God).

Of course, as it stands, this is all pretty vague; and one of the things that needs to be clarified is what exactly is meant by the claim that there *could not* be a difference in the thing that supervenes without a difference in the thing supervened on. Is it that it

is metaphysically impossible for a difference in the first to occur without a difference in the second (to use the terminology introduced in the last chapter), or only that it is physically impossible? If the claim is understood in the first way, then many of the problems that afflict reductionism turn out also to afflict the suggestion that the mental supervenes on the physical (for reasons we'll be exploring later). But if the claim is understood in the second way, then it isn't clear that the position that results genuinely counts as a form of materialism. For to claim that it is physically impossible for there to be a difference at the mental level without some difference at the physical level is just to claim that there can be no such difference *given the way the actual world happens to work*; it is not to claim that it is metaphysically impossible; that is, impossible in any possible world, not just in the actual one – and thus it is not to claim anything that rules out the dualist's basic idea that it is metaphysically possible for the mind to exist apart from the brain and body.

The advocate of supervenience has, no less than the reductionist, the problem of giving a useful account of exactly *what* the basic entities and laws of physics are on which everything is claimed to supervene. The response that a "completed physics" will someday give the answer leaves open the possibility that the hypothetical physicists of the future will see fit to add non-physical or immaterial phenomena to their list. Indeed, at least one self-described naturalist, David Chalmers, has predicted that this is precisely what the physics of the future will require – which is why he counts himself not only as a naturalist, but also as a dualist, thereby explicitly rejecting any essential link between naturalism and materialism!

This last point should caution us to keep in mind that, as I indicated earlier, the terms "naturalism," "materialism," and "physicalism" – and I should now add the terms "reductionism" and "supervenience" – are used by philosophers in a bewildering variety of ways. For our purposes it will suffice to reiterate that

"materialism" essentially conveys a general commitment to the idea that physical reality is all the reality there is. Attempts to spell this basic idea out in greater detail tend either to take current physics (or something like it) as the touchstone of what counts as "physical reality" (and thus frequently adopt the label "physicalism"), or instead to leave the concept of the physical somewhat open-ended (and thus sometimes opt for the label "naturalism"). Predictably, the former sort of approach, being bolder and more determinate, is harder to defend, while the latter, though easier to defend, is often less determinate, and in some cases even less clearly "materialistic" in substance. Either way, the intuitive and commonsense feel of materialism seems to last only as long as one keeps one's statement of it vague.

Cause and effect

So far it might seem that the initial plausibility of materialism is so vitiated by its indeterminacy that, while it is understandable how some might find it attractive, it is hard to see why it has become the mainstream position in the philosophy of mind. But we must not forget the interaction problem that, as we saw in the previous chapter, serves as the main objection to dualism and the chief philosophical motivation for materialism. Modern physics, as usually interpreted, teaches us that the material universe, to which dualists, no less than materialists, take the human body to belong, is causally closed. Accordingly, nothing outside it – nothing non-physical – would seem capable of having any causal influence on what happens in that universe. But then the mind, if it were a Cartesian non-physical substance, would be incapable of having any effect on the body; and yet it seems just obvious that it does. The materialist thus concludes, and surely not unreasonably, that if the mind interacts with the body, it can't be a Cartesian non-physical substance, but must be purely material or physical.

This argument appeals to general facts about the nature of cause and effect relations in the physical world. But there are also quite specific facts about mind-body interaction that give further support to the materialist thesis. We know from everyday experience that changes in the body can have drastic effects on the mind – for instance, ingesting too much alcohol or suffering head trauma can radically impair one's ability to think clearly, or even to think at all. How could this be, if the mind is as utterly distinct from the body and brain as Descartes held it to be? We also know from modern neuroscience that various specific mental functions – vision, hearing, the understanding of language, and so on – are associated with specific regions of the brain. Again, how likely would this be, if the mind and the brain were distinct things? Nor is neuroscience the only source of scientific objections to dualism. Modern biology tells us that human beings are the products of the same, purely material, process – evolution – which operates according to the same physical laws that govern the rest of the physical universe and, beginning in the purely material environment of the early history of the Earth produced cows, houseflies, and bacteria, all of which seem obviously to be purely physical entities. So how can human beings, one outcome of this material process, be anything other than purely physical entities? The theory of relativity postulates that space and time form a single continuum – space-time – so that anything existing in time must exist also in space. Yet mental processes seem clearly to exist in time, as even Descartes acknowledged, in which case they would surely have to exist in space as well. How then could they fail to be physical or material processes?

The appeal to the success of modern science in applying the mechanistic model of explanation to every other phenomenon in the universe is thus by no means the only arrow in the materialist quiver. Both the general nature of physical causality and the specific details of the causal relations between mind and body also confer considerable plausibility on materialism. Given (a) that the

nature of cause and effect relations seems to require that the causes and effects of physical processes be themselves physical, (b) that application of this idea has led to a general mechanistic model of the universe that has been enormously successful in explaining every other aspect of reality, and (c) that we already know of certain specific causal links between the mind and the brain, the materialist can argue that the most reasonable conclusion is to suppose that the mind will, eventually, yield *completely* to a purely physical explanation.

None of this exhibits by itself any fallacy in the arguments for dualism – such as the conceivability argument – that we considered in the previous chapter. But some materialists have suggested that they can even present a conceivability argument of their own, to counter that of the dualist. Imagine that in the far future, teleportation devices of the sort described in science-fiction stories become possible. A person steps into a chamber here on Earth, and a supercomputer scans his or her body and brain, recording all the information gleaned, down to the last molecule. As the person's body is destroyed, this information is beamed to another chamber on Mars and an exactly similar body appears in the Martian chamber. This sort of scenario raises all sorts of interesting philosophical questions, such as whether the person who appears in the chamber on Mars is the same as the one who stepped into the chamber on Earth, or a mere duplicate. We will address such questions in chapter 8. What we want to take note of here is that it certainly seems conceivable, and thus metaphysically possible, that the person who appears in the Martian chamber will, whether or not he or she is identical to the original, exhibit exactly the same sort of behavior, and thus appear, no less than the original did, to have a *mind*. But what caused this person to exist was the storage and transmission of purely *physical* information – the information the computer scanned from the body and brain on Earth – and the use of that information to produce the person who appeared in the chamber on Mars. It would seem, then, that purely

physical factors can generate a mind, in which case there is reason to believe that the mind is purely physical.

This argument is not exactly parallel to the dualist's conceivability argument. That argument was intended to prove that the mind and brain are not identical, while this one is intended to support the claim that they are or at least that the former supervenes on the latter. But conceivability arguments, if they prove anything, seem unable to prove positive claims about identity or supervenience. If you really can conceive of the mind existing apart from the body or brain, it is at least plausible that this would provide evidence that they are not identical, for if they were, how could you have one without the other? But to conceive of them existing together hardly proves that they are identical – after all, even the dualist supposes that they normally do exist together, and insists only that they nevertheless could, in principle, come apart. To imagine that all creatures with kidneys also have hearts doesn't prove that hearts and kidneys are the same type of organ; similarly, to imagine minds existing wherever brains exist hardly demonstrates that the mind and the brain are the same thing. So the materialist conceivability argument cannot, in the nature of the case, prove its conclusion. Nevertheless, it vividly illustrates, and provides intuitive support for, the conclusion the materialist draws from the other considerations we've examined: that it seems at least possible that purely material processes could entirely account for the existence and nature of the mind.

Behaviorism

Suppose we grant the strength of the materialist's case so far. As it stands, it supports at most the claim that it is possible to give a purely physical account of the mind. But how is this possibility to be made actual? Can the materialist tell us specifically how entirely material processes in the body and brain produce all the rich

mental phenomena we've described in the previous two chapters – consciousness and thought, qualia and intentionality, and a robust sense of selfhood? Materialists have proposed several possible answers to this question, and the first to gain currency in the mid-twentieth century – the era in which materialism became the majority position within the philosophy of mind – was *behaviorism* (sometimes called "philosophical behaviorism" to distinguish it from the "methodological behaviorism" associated with B. F. Skinner and other psychologists, which is a different, though related, idea).

Behaviorism holds that to attribute a mind to something is to attribute to it certain *behavioral dispositions*; to have the relevant dispositions *just is* to have a mind. To experience pain, for example, is nothing more than to be disposed to exhibit such behaviors as moaning, wincing, crying, or saying "Ouch!" when one's body has been injured. To believe that it's raining outside is to be disposed to look for an umbrella, or put on galoshes whenever the weather forecast predicts rain. To feel fear is just to have a tendency to tremble and/or run away when in the presence of wild animals, or knife-wielding strangers in dark alleys. In general, to have any sort of mental state is just to have a propensity to produce certain behavioral *outputs* in response to given environmental *inputs*, and in particular in response to the effects one's surroundings typically have on one's sensory organs. If behaviorism is true, then the explanation of the mind in entirely material terms would be relatively easy, simply a matter of showing that a purely physical system is capable of exhibiting the behavior associated with having a mind – something the human body obviously *is* capable of.

Behaviorism isn't true, though. It is sometimes said that no philosophical theory has ever been decisively refuted, although probably not by anyone familiar with this account of the mind, which appears not to have a single defender today. To be fair, it is clear that behaviorism has certain advantages. It makes the mind

every bit as observable and accessible to scientific study as tables, chairs, rocks, and trees, and it can seem to reflect common sense, in so far as the way we normally do have access to minds, or at least to the minds of other people, is precisely through their behavior. What you *observe* in observing someone's grief seems, strictly speaking, not to be something going on *inside* him or her, but rather just certain outward behaviors: sobbing, grimacing, and the like. Moreover, this fact, together with a certain theory of meaning prominent in mid-twentieth century philosophy – the "verifiability theory," which held that the meaning of a statement is its method of verification – seemed to make behaviorism almost unavoidable: if the only evidence you could have for verifying claims about what other people are thinking is the behavior they exhibit, then to say that they are thinking must be nothing more than to say that they tend to exhibit certain behaviors.

The verifiability theory has long since been abandoned, for a number of reasons, not the least of which is that, since it is hard to see how the theory itself could be verified, it is also hard to see how it could fail to imply its own meaninglessness; and with the verifiability theory goes the strongest argument that could possibly be given for behaviorism, in the absence of which its problems seem overwhelming. For one thing, it is notoriously difficult to see how talk about minds could ever be completely reduced to talk about behavior. To say that to believe it is raining is just to be disposed to put on galoshes or look for an umbrella is obviously not quite the whole story. Someone who believes that it is raining will do these things only if he or she desires not to get wet, and a desire is itself a kind of mental state. So the behaviorist now has to analyze *the desire not to get wet* in terms of behavior, in order to complete the analysis of *the belief that it is raining* in the same terms. But someone will desire not to get wet only if, for example, he or she also fear catching cold, and *the fear that one will catch cold* is thus yet another mental state that must be analyzed in terms of

behavior – a mental state that will in turn be present only if a further mental state, *the belief that getting wet causes colds* is also present, and which will also have to be given a behaviorist analysis. And so on *ad infinitum*. There seems, accordingly, no way for the behaviorist ever to cash out all talk about mental states and processes in terms of nothing but behavior.

More fundamentally, the theory leaves out the subjectivity that, as we saw in chapter 2, seems essential to the mind. Whether or not I know about other people's minds from behavior alone, that is surely not how I know about my own: it's not as if I have to catch myself in a mirror screaming and crying before I can conclude "Hey, look at that! I must be in pain!" The subject of thoughts and experiences appears to have an access to them that others do not have, an access that does not rest on the observation of behavior. Indeed, given this subjectivity, behavior of any sort seems inessential to the mind. A good actor could convincingly exhibit all the behavior associated with the most excruciating pain, and yet not be in pain at all; an even better actor could really be suffering excruciating pain and yet, to all appearances, be feeling nothing. The mental facts – the presence or lack of the "qualia" associated with pain – would in either case consist entirely of what was going on from the "inner," subjective point of view of the actor, and be knowable only from that point of view, the behavior being irrelevant.

The issue of *causation* is also relevant here, as it was in the discussion of dualism. One of the materialist's objections to dualism is that it allegedly fails fully to account for the fact that mental states are the causes of behavior. But behaviorism also fails to take account of this. For if mental states are identical to behavior, they can't be the causes of it: your belief that it's raining doesn't *cause* you to get your umbrella, according to the behaviorist; it is your getting your umbrella. To take seriously the materialist's commitment to the causal efficacy of the mental requires the rejection of behaviorism.

The identity theory

Inspired by the fact that mental states and processes seem clearly to be inner processes of some sort, and states and processes that cause outward behavior, materialists turned away from behaviorism in the 1950s and 1960s and tended to favor instead the *identity theory*. If mental states and processes are the causes of behavior, but causes that are in some way inside the one exhibiting the behavior and thus unobservable, then there seems to be an obvious candidate from the materialist point of view for where exactly such inner causes might be found: the brain. In this view, mental states and processes are just neurological states and processes; that is, they are states and processes of the brain and central nervous system. The mind is identical to the brain.

Here again we have a claim that seems simple and obvious, but which in reality is neither. The idea is that any given mental state – your thought about your grandmother, the sensation of pain in your lower back, your memory of your last trip to London – is *the exact same thing as* the firing of such-and-such a clump of neurons in your brain. It is important to understand precisely what this means. It is *not* the claim that what happens in the mind is *affected* by what happens in the brain – that the feelings and sensations you have, your abilities to remember and think clearly, and so forth, depend on various neural structures and processes. No-one denies that – certainly not the dualist, who insists, as we've seen, that the mind and brain interact with one another (even if he has a hard time explaining how). If that were all the identity theory were saying, it wouldn't be very interesting or controversial. The theory is, rather, not that your thought is *caused* by such-and-such neurons firing, but that it *is* such-and-such neurons firing. There is *nothing more* to the thought than that. Certain electrochemical signals are sent from one part of the brain to another; and that, and only that, is what constitutes a thought, feeling, or sensation. If you were able to peer inside someone's skull and somehow see

the neurons firing, you would, literally, be looking at his or her thoughts.

If that doesn't sound strange to you, you probably haven't understood the theory correctly. It is meant to sound strange; or at least, it is not meant to sound obvious. Identity theorists took themselves to be putting forward a bold scientific hypothesis, not a common sense truism. The idea was that the identification of the mind with the brain ought to be accepted as the latest in a long series of scientific reductionist explanations. As noted earlier, everyday physical things like tables and chairs, though they seem to be utterly impenetrable objects with features like color, taste, and odor, are really nothing but swarms of colorless, odorless, and tasteless microscopic particles. Physical objects have been "reduced" to collections of molecules and atoms by contemporary physics. Similarly, properties like heat, cold, liquidity, or luminance have been reduced to properties of aggregates of molecules, or atoms. So water turns out to be nothing other than a particular chemical compound, a composite of hydrogen and oxygen: H_2O. Heat, to use another typical example, is nothing but the motion of molecules – high mean molecular kinetic energy, to be slightly more precise. Such reductions reveal the true nature of everyday commonsense phenomena, and allow us to understand them and predict their behavior with greater precision than common sense makes possible.

Reductions sometimes take place *within* science: the biological concept of the gene, for instance, turns out to be reducible to the more fundamental concept of DNA. This sort of example is called an "intertheoretic reduction": the reduction, that is, of the laws and ontology of one scientific theory to those of another. The ontology of a theory is just the list of the basic entities it postulates, such as the molecules, atoms, and sub-atomic particles of modern physics; the laws of the theory are the principles it says govern the activities of the entities in its ontology, such as the principles of quantum mechanics that are said to govern the basic entities

postulated by physics. In the case of an inter-theoretic reduction, the entities of the theory that gets reduced turn out to be identical to, or "nothing but," the entities spoken of by the reducing theory: genes, to over-simplify again, turn out to be reducible to, or are in reality nothing but, aspects of DNA. There is, accordingly, a law-like connection between the entities of the reduced and reducing theories: in every case where such-and-such a gene is present, such-and-such an aspect of DNA is also present.

The identity theory is sometimes formulated as a kind of inter-theoretic reduction. Our ordinary, commonsense way of talking about our minds and of explaining our behavior in terms of what is happening in our minds – speaking of beliefs and desires, for example, or of a person's behavior as being caused by certain specific beliefs and desires – is claimed to be a quasi-scientific "theory." It is, to be sure, not a sophisticated theory, stated with mathematical precision, created by an eccentric academic or graduate student, proffered in the lecture hall or tested in the laboratory. But it does, or so it is argued, have certain features of a scientific theory. It has a complex ontology – it talks not only of beliefs and desires, but also of hopes, fears, experiences, feelings, emotions, sensations – and it appeals to certain quasi-law-like generalizations: that a desire for a cheeseburger will tend to cause one to eat a cheeseburger, that the sensation of pain will tend to cause moaning and complaining, or that the belief that danger is near will tend to cause fleeing the scene. Since this "theory" is a theory about the mind, and since it is a theory that is held by the "common people" as much as by the educated, it is typically referred to by philosophers as *folk psychology*. The identity theory can thus be expressed as the hypothesis that folk psychology can be reduced to neuroscience, the science of the brain. Just as the theory that spoke of genes and the like turned out to be reducible to a theory that speaks instead in terms of DNA, so too should we reduce beliefs, desires, experiences, sensations, and emotions, to brain states and processes.

Identity theorists appeal, in defense of their theory, to the sorts of considerations adduced earlier in favor of materialism in general, and to the dependence of various specific kinds of mental functions (language, vision, etc.) on various specific regions of the brain in particular. They acknowledge that their theory might seem counter-intuitive: how, it might be asked, can subjective thoughts and sensations be nothing but electrochemical signals passing between nerve cells? But they also note that a table, for instance, does not seem much like a collection of particles, even though that is what it is. Common sense has often been challenged by the advance of science. If the identity theory too challenges common sense, that can, by itself, be no objection.

There are, however, more serious problems with the identity theory, which materialists themselves have pointed out. The first has to do with a technical distinction made by philosophers between *types* and *tokens*. Consider the sentence: "The cat is on the mat." How many words are in that sentence? The answer depends on whether we count "the" once or twice. If we count "the" as one word, we are counting it by *type*; if we count it twice (since it appears twice in the sentence) we are counting its *tokens*. There are *five* different words in the sentence if we count word types, and *six* if we count word tokens. What is true of words is also true of mental states and brain states (and pretty much everything else, for that matter). We can, for instance, distinguish between a general type of mental state (for example, the belief that it is raining) and particular tokens of that type (for example, the belief that it is raining that I had earlier this summer, the belief that it is raining that I had last April 16, the belief that it is raining that you had on May 1, and so on). The identity theory was originally intended as what might be called a "type-identity" theory: it claimed that for each *type* of mental state (the belief that it is raining, the belief that it is sunny, the desire for a cheeseburger, the desire for a cookie, and so on and on) there could ultimately be matched, one-to-one, a specific *type* of brain state (neuronal

firing pattern of type A, neuronal firing pattern of type B, and so forth).

The trouble is that it seems clear that there *can't* be such a neat matching, because there can't be such a thing as a law-like correlation between mental states and brain states. Recall a point made above in response to behaviorism: a person will typically desire not to get wet only if he or she has other mental states, such as a fear of catching cold and a belief that getting wet tends to cause colds; moreover, he or she will have *those* mental states only if he or she also believes that catching a cold will be unpleasant, and desires to avoid this unpleasantness more than desiring to frolic in the rain, etc. Any given mental state, then, is never had individually, but involves the having of other mental states as well; and it typically also involves there being rational connections between the mental states one has. It is *because* one believes that catching cold is unpleasant and that getting wet tends to cause colds that one *infers* that one had better not get wet, and then draws the further inference that since going out in the rain, however pleasant, will cause getting wet, one had better not go out in the rain.

So there are *logical relations* between mental states that partially determine precisely which mental states one will have, if one has any at all. But there seem just obviously to be no such relations between neurons firing in the brain. It would be absurd to say – indeed, it isn't clear what it could even *mean* to say – that "neuronal firing pattern of type A logically entails neuronal firing pattern of type B," or that "the secretion of luteinizing hormone is logically inconsistent with the firing of neurons 6,092 through 8,887." Neurons and hormone secretions have *causal* relations between them; but *logical* relations – the sort of relations between propositions like "It is raining outside" and "It is wet outside" – are not causal. There seems to be no way to match up sets of logically interrelated mental states with sets of merely causally interrelated brain states, and thus no way to reduce the mental to the physical. The best we can hope for is a kind of "token-identity" theory:

particular
tokens – y
some neur
state and b
relationshi
This sort
coined by
closely ass
events, th
"monism"
connecting
"anomalou

A relate
sible that t
they lack b
– somethin
those com
obvious ex
beings are
exist in the
acteristics
similar to
artificial br
would also
to think ar
then, if mi
brains, how

Functio

The multip
torically le
been domi

particular mental state tokens are identical to particular brain state tokens – your belief that it's raining is identical to the firing of some neurons or other – but there is no way to correlate mental state and brain state types in a law-like way, no way to describe the relationship between them in terms of a rigorous scientific theory. This sort of view is sometimes called *anomalous monism*, a label coined by Donald Davidson (1917–2003), the philosopher most closely associated with it: mental events are identical to physical events, the physical being all that ultimately exists (hence "monism"); but there is no way to formulate any scientific laws connecting the mental and the physical (hence the adjective "anomalous").

A related problem with the identity theory is that it seems possible that there could be creatures that have minds even though they lack brains; the mind, that is to say, seems "multiply realizable" – something that could be "realized," or exist in, systems other than those composed of neurons. Divine beings and angels would be obvious examples, and even most atheists would admit that such beings are at least metaphysically possible, whether or not they exist in the actual world. Extraterrestrials with physiological characteristics utterly different from our own – with nothing remotely similar to human brains or nervous systems – and androids with artificial brains composed of silicon, plastic, and copper wiring, would also seem potential candidates for creatures that can be said to think and feel despite lacking our neurological makeup. But then, if minds could possibly exist in physical systems other than brains, how can they be *identical* to brains?

Functionalism

The multiple realizability objection leads us naturally – as it historically led most materialists – to the form of materialism that has been dominant in the philosophy of mind since the 1970s.

Functionalism takes as its starting point the observation that many things are properly characterized not in terms of the stuff out of which they are made, but rather by reference to the functions they perform. A knife is defined by its ability to cut, not its material composition; whether the knife is made of steel or plastic is irrelevant to its status as a knife. The game pieces of checkers are defined in terms of the functions each piece plays in the course of the game: usually the pieces are made out of plastic and moved about on a cardboard surface, but in principle one could draw a checkers board on the beach, and play the game using crushed beer cans and dead crabs. Of course, not just *any* sort of material composition will do: it would be difficult to play checkers with game pieces made of shaving cream, and a knife made out of shaving cream wouldn't truly be a knife at all. But the point is that there is still no *specific* kind of physical stuff that knives or checkers pieces have to be made out of; lots of things could do the job, as long as they have the right sort of structure to perform the requisite functions.

The functionalist claims that something similar is true of mental states and processes. It is not the stuff of which it is made that makes a particular mental state the kind it is – whether the firing of neurons or otherwise – but rather what it does, and, in particular, what sorts of causes and effects it has. What makes a sensation of pain the kind of thing it is, is that it is caused by damage to the body and tends to cause in turn certain other mental states, like anxiety, as well as behaviors like screaming and crying. What makes the belief that it is raining the sort of thing it is, is that it tends to be caused by light reflected from raindrops striking the retinas, tends in turn, and when a desire to stay dry is also present, to cause certain other mental states such as the intention to get an umbrella, and tends, in tandem with these other mental states, to cause bodily behavior like going to the closet to get an umbrella. Mental states are to be defined, then, in terms of their causal relations to other mental states, and ultimately this system of mental states is itself to be defined in terms of its causal relations to the inputs

provided b
the output
manifests t
each eleme
makes the
tiated in a
the silicon
vant. Just a
knife, whet
anything n
have a min
ours or so
android, or

One of
for an anal
materialism
mental stat
does not ru
instantiated
something
something
is complex
the human
surely fulfil
sible an exp
together wi
ism. Moreo
tiated in sys
functionalis
ability argu
apart from t
multiply rea
than brains
any materia

provided by environmental influences on the sensory organs and the outputs manifested in bodily behavior. That the whole system manifests the *specific kinds* of causal relations it does is what makes each element within it a distinctly *mental* state or process, and what makes the system as a whole a *mind*; whether this system is instantiated in a human brain, the slimy innards of an extra-terrestrial, or the silicon central processing unit of a sophisticated robot is irrelevant. Just as anything performing the right sort of function is a knife, whether made of plastic, steel, or something else, so too can anything manifesting the right sort of causal relations be said to have a mind, whether it is a creature with a nervous system like ours or some very different sort of being altogether: an ET, an android, or an angel.

One of the advantages claimed for this view is that it allows for an analysis of the mind that is, in principle, neutral between materialism and dualism. Functionalism *per se* holds only that mental states are to be defined in terms of their causal relations; it does not rule out the possibility that these causal relations might be instantiated in a Cartesian immaterial substance rather than in something physical. But of course the theory also allows that something that *is* entirely material could have a mind, as long as it is complex enough to manifest the relevant causal relations, and the human brain, being the most complex object known to us, surely fulfills this requirement. Functionalism thereby makes possible an explanation of the mind in purely physical terms, and this, together with Occam's razor, seems to favor materialism over dualism. Moreover, since the theory holds that minds could be instantiated in systems other than brains, it is sometimes suggested that functionalism allows the materialist to rebut the dualist's conceivability argument: if it seems conceivable that the mind could exist apart from the brain, this might simply be because mental states are multiply realizable – possibly instantiated in physical systems other than brains – and not because they can exist totally independent of *any* material substrate. Thus functionalism, even if in principle

consistent with dualism, has in practice become the favored theory of materialists.

Some might question whether the idea of multiple realizability, on which functionalism rests, is really all that plausible in the first place. Should we accept so readily the suggestion that a sophisticated robot, of the sort described in the science-fiction novels of Isaac Asimov, in the *Terminator* movies, or the character Data on *Star Trek*, can be said *literally* to think and feel as we do? If we accept that such creations of fiction are at least conceivable – that we can coherently imagine a creature constructed of nothing but steel and plastic, yet which has a mind – then this would seem to give some support to the functionalist. After all, if you could really meet Data or the Terminator and engage in a conversation with them, would you really have any doubt that they were as intelligent as you? If Data asked you what time it was, wouldn't this be reason to think he *desired* to know the time? If the Terminator told you he had come from the far future, wouldn't this be evidence that he *believed* that that's where he came from? Beliefs and desires are kinds of mental states; so anything that possessed them could surely be said to have a mind. One might, nevertheless, object that such creatures wouldn't have the *feelings* and *emotions* we have. But why couldn't they? Doesn't this objection reflect merely the bias of science-fiction writers for the stereotype of the cold, unfeeling machine rather than any objective limits on the kind of robots that might in theory be constructed? The functionalist, it must be remembered, holds that feelings and emotions too are nothing but states having certain kinds of causal relations. Why couldn't such states be built into a robot? If a robot had an internal state that was caused by damage to its body, that caused it to scream and cry out and look frantically for ways to repair the damage, why wouldn't this count as pain? If you saw Data flailing on the ground, shrieking and sobbing and holding his side after having been shot with a ray gun, wouldn't you try to help him? Would you say to him "Cut it out, you're just a robot – you don't *really* feel anything!" (And what if he

told you th
wonder at

The fun
there coul
dismissed
claim that
erence to
after all, n
chips in th
gest that s
can think
between c
states than

A single
from some
a compute
rons were
and sent si
any doubt
you were
receiving i
still a hear
should arti
and feelin
ther that th
nerve end
mechanism
exactly as
presence o
that you'd
you were
function p
shouldn't
other neu

told you that it hurt his feelings to hear you say that? Mightn't you wonder at least a *little* whether he really did have feelings after all?)

The functionalist would argue further that the suggestion that there could be thinking and feeling robots cannot in any event be dismissed by anyone who takes seriously the general materialist claim that mental states and processes are entirely explicable by reference to states and processes of the brain. A clump of neurons is, after all, no less purely physical than a cluster of silicon computer chips in the head of a robot. Why should it be so outrageous to suggest that something whose "brain" is made of such computer chips can think and feel as we do? Why should electrical current passing between computer chips be any less capable of producing mental states than electrochemical signals sent between neurons?

A single neuron performs a relatively simple task: it gets signals from some neurons and then sends signals to others. Why couldn't a computer chip do that? Suppose a very small clump of your neurons were replaced by tiny computer chips, and that they received and sent signals in just the way the original neurons did. Is there any doubt that you'd be just as conscious and capable of thought as you were before? An artificial heart doesn't make the person receiving it any less capable of pumping blood: an artificial heart is still a heart, because it performs the functions of a heart. So why should artificial neurons be any less capable of supporting thought and feeling, if they do exactly what real neurons do? Suppose further that the nerve endings in your hand were replaced by artificial nerve endings – made of microscopic wires, or the sorts of tiny mechanisms familiar from nanotechnology – that functioned exactly as the originals did, registering damage to the body, the presence of heat and cold, and so forth. Is there any reason to doubt that you'd be just as capable of feeling pain, warmth, or coolness as you were before? If so, why exactly? The artificial nerve endings function physically in exactly the same way as the originals; so why shouldn't their ultimate effects be the same? Now imagine that other neurons and nerve endings are gradually replaced in a

similar fashion, and also that various organs – a liver, a kidney, a lung – are replaced by extremely complex and sophisticated duplicates, constructed of plastic, steel, and silicon but which exactly mimic the functioning of the originals. Is there any reason to doubt that you would be able to think and feel just as well as you ever did? The new organs and neurons function physically *exactly* as the originals did; so why wouldn't their end results be identical as well? (And if you *do* somehow lose the ability to think and feel as before, exactly *when* does this happen? Replacing *one* clump of neurons or nerve endings had no such effect – so why should replacing two, three, two thousand, or two million?) Finally, imagine that eventually your *entire* body and nervous system is replaced by these artificial duplicates. Is there any doubt that you'd be just as conscious as you were before? Again, if so, why exactly? Your new parts are entirely physical, but so were your original neurons and organs, and the new parts function exactly as the originals did. So what reason could there be for doubting that you still have a mind? Notice, however, that you would in effect have become a robot. But if you, having been transformed gradually into a robot, could nevertheless think and feel, why deny that *other* robots – the kind made in a factory or laboratory – might also?

As this argument indicates, functionalism is closely tied to the idea that the brain is a kind of *computer*, with the mind a kind of *program*: the software that runs on the hardware of the brain. We will explore this in greater detail in chapter 6. Suffice it for now to note that this suggestion provides the materialist with a way of elucidating the functionalist thesis, and of arguing that it eliminates the mystery of how something purely material could have a mind. A computer program is something abstract – a mathematical structure that can be understood and specified, on paper or in the programmer's mind, long before anyone implements it in a machine. Yet for the program to become "real" – for it to have any impact on the physical world and be usable by us – it must be so implemented. Unless you can download it on to an actual piece

of computer hardware, it remains purely abstract and inefficacious. It needn't be any *particular* computer that does the job – some programs could be run on almost *any* computer – but there must be *some* computer or other that does it. This may serve as a fitting analogy for the mind: we can understand the mind in functionalist terms, by abstracting away from it any of the physical details of its implementation in human brains and focusing only on its causal structure. This may give the illusion that it is capable of existing apart from some implementation; but in fact, just like a computer program, it must be implemented in some physical system or other – and if not necessarily in a human brain, then perhaps in a robotic or extraterrestrial brain. Furthermore, despite a program's abstract character, there is no mystery about how it can be run on a piece of computer hardware. But then, by analogy, there need be no mystery about how the mind can be instantiated in the brain: like computer software, it is merely an instance of a complex abstract structure being realized in a complex piece of matter.

The burden of proof

Despite the ambiguities that plague attempts to give the materialist thesis a precise formulation, then, it remains powerful. If the commonsense, down to earth character of materialism is sometimes overstated by its advocates, it nevertheless seems to get strong support from general trends in modern science. Moreover, in functionalism, materialists have a promising general philosophical theory of how the mind might be realized in something purely material, and there is compelling evidence from neuroscience that mental states and processes are indeed inextricably tied to states and processes of the brain.

What implications does all of this have for the dispute between materialism and dualism? Many materialists are of the opinion that the considerations adduced so far are sufficient by themselves to

establish the rational superiority of their creed. Materialism is, in their estimation, fully capable in principle of explaining the mind. The work remaining is little more than a mopping up operation, the mere filling in of details. Dualists have effectively been refuted; at the very least, the burden of proof lies with them, not with the materialists. Given the overall evidence, materialism has a presumption in its favor. It is innocent until proven guilty.

So it might seem. Dualists could reply, however, that the philosophical advantage claimed by materialism may be illusory, with the current consensus in its favor a reflection more of intellectual fashion than of objective, dispassionate evaluation of the relevant arguments. In particular, dualists might argue that there is no good reason to take seriously the suggestion that, in the debate between materialism and dualism, it is materialism which must get the benefit of the doubt. The purported historical justification for such an attitude is familiar enough: for centuries, it is said, materialists and their opponents did philosophical battle, with neither side gaining the advantage; but then along came modern science, and phenomena which previously seemed inexplicable except in terms of supernatural forces increasingly succumbed to materialistic explanation. The mind is merely the last holdout, and that circumstance is only temporary; for with the rise of neuroscience, we now stand on the threshold of finally explaining the mental in entirely physical terms, and the materialist worldview will thereafter be completely vindicated. But however influential it has been, this historical-philosophical case has, arguably, been overstated.

First, the advance of science, far from settling the mind-body problem in favor of materialism, seems to have made it more acute. Modern science has, as noted in chapter 2, revealed that physical objects are composed of intrinsically colorless, tasteless, and odorless particles. Colors, tastes, and odors thus, in some sense, exist only in the mind of the observer. But then it is mysterious how they are related to the brain, which, like other material objects, is composed of nothing more than colorless, tasteless, and odorless

particles. Science also tells us that the appearance of purpose in nature is an illusion: strictly speaking, fins, for example, don't have the purpose of propelling fish through the water, for they have in fact *no purpose at all*, being the products of the same meaningless and impersonal causal processes that are supposed to have brought about all complex phenomena, including organic phenomena. Rather, fins merely operate *as if* they had such a purpose, because the creatures that first developed them, as a result of a random genetic mutation, *just happened* thereby to have a competitive advantage over those that did not. The result mimicked the products of purposeful design; in reality, it is said, there was no design at all. But if purposes are thus "mind dependent" – not truly present in the physical world but only projected on to it by us – then this makes that act of projection, and the intentionality of which it is an instance (as are human purposes, for that matter) at least difficult to explain in terms of processes occurring in the brain, which seem intrinsically as brutely meaningless and purposeless as are all other purely physical processes. In short, science has "explained" the sensible qualities and meaning that seem to common sense to exist in reality only by sweeping them under the rug of the mind; that is, it hasn't really explained them at all, but merely put off any explanation by relocating them out of the physical and into the mental realm. There they remain, however, forming a considerable bump under the rug – one that seemingly cannot be removed by further scientific sweeping.

Second, the debate over materialism has arguably never been more than tangentially concerned with how best to explain physical phenomena – the motions of the planets, the nature of chemical reactions, or even the origins of life. That is to say, straightforwardly scientific issues seem never to have been the crucial ones. Rather, the debate has, for two and a half millennia, focused primarily on three fundamental metaphysical issues: the nature of the mind and its relation to the body, the ontological and epistemological status of mathematical and other apparently

abstract objects, and the question of the existence of God. For materialism now genuinely to have the upper hand would require that materialist arguments have been victorious, or have at least been shown to be considerably more plausible, in each of these subject domains. Has this happened? No one familiar with the recent history of philosophy can honestly think so.

This is obviously so in the case of the first domain, which is the very subject presently at issue. Materialism may be the majority position in contemporary philosophy of mind, but not because anyone has proved it true. Indeed, as we will see in succeeding chapters, virtually all the work done today by materialist philosophers of mind consists, at bottom, of trying to defend their favored brands of materialism against various objections, which are implicitly or explicitly anti-materialist in character, that is, to the effect that the brand of materialism in question fails genuinely to explain some given mental phenomenon (intentionality, qualia, etc.) in entirely physical terms. Moreover, these objections are typically variations on the same criticisms of materialism that have been given for 2,500 years, with modern materialists no closer to answering them decisively than were their intellectual forebears. Dualists might argue that the fact that the project of naturalizing the mind – of attempting to show it to be explicable without resorting to non-physical properties – is as popular as it is a sign of the weakness of materialist philosophy of mind, rather than of strength; for if there were no serious doubt that the mind is explicable in purely material terms, the naturalization project should have been largely accomplished long ago. Again, the dominance of materialism in the philosophy of mind would seem to rest largely on the belief that materialism has been established everywhere *else*, so that it is reasonable to expect it to succeed where the mind is concerned.

But it seems clear that materialism has not been established everywhere else, at least if we keep in mind that it is metaphysical disputes, not scientific ones, which have historically been at issue.

Consider the second domain of debate between materialists and their opponents, namely, the debate over abstract objects. Among philosophers, mathematics has long been the paradigm of knowledge that is absolutely certain, and that is because the truths of mathematics are *necessary* truths, true in all possible worlds. For this reason, it seems clear that these truths cannot be truths about anything either mental or material: facts about the mental are facts about a subjective realm, but mathematics is objectively true, utterly independent of human interests; facts about the material world are facts about a realm that is constantly in flux, a domain of contingency, but mathematical facts are unchanging and eternal. Mathematics thus seems to describe a third realm, a domain of abstract entities – numbers, geometrical forms – that cannot be reduced to either the mental or the physical; that is, it seems to lead to what is called *Platonism* (after Plato, the philosopher most widely associated with this sort of view). Many philosophers have of course attempted to disprove this conception of mathematics, and to show that mathematical truth can, despite appearances, be naturalized. The point is that such attempts have, at best, consistently proven to be highly controversial, and, more commonly, rejected by most philosophers as ultimately implausible. The dialectic is familiar to philosophers of mathematics: the nature of mathematical truth seems inevitably to lead to Platonism; naturalistically inclined philosophers try to show otherwise; their attempts then prove to be riddled with insuperable difficulties, or even subtly to entail Platonism of a different kind. This pattern seems to be the same today as it has been for the whole history of philosophy. And if anything, it is not naturalism but Platonism – appearing as it does to follow inevitably from the nature of mathematics, and having withstood every attempt to disprove it – which ought to get the benefit of the doubt, especially given that many mathematicians themselves, in their philosophical moments, tend to be Platonists.

What holds for mathematical objects holds no less for other apparently abstract entities. When we understand a truth of

mathematics, we grasp a *proposition* – the proposition that 2 + 2 = 4, say. But we also grasp propositions when we understand any other kind of truth, and, as in mathematics, the objects of our understanding seem clearly to be neither mental nor physical. In understanding the Pythagorean theorem, or that Caesar was assassinated on the Ides of March, you and I understand the *same thing* in each case. It is not that I understand my own subjective Pythagorean theorem and you understand yours; what we understand is something objective, something that holds true independently of either of our minds. So it cannot be something mental. But neither can it be something material, for the fact the theorem describes would hold true whatever occurs in the physical world, and even if there were no physical world. This, again, is no less true of propositions about physical things: the proposition that Caesar was assassinated on the Ides of March would remain true even if the entire physical universe disappeared tomorrow; in grasping it, you can't be grasping something material. This way of putting the argument for propositions as abstract, immaterial entities is associated with Gottlob Frege (1848–1925), but the basic idea goes back a long way in the history of philosophy, and ultimately, to Plato. Plato is also associated, of course, with the idea that our words for the properties of things – redness, roundness, or goodness – refer to *universals* or *forms* which exist in some sense abstractly, independently of particular concrete objects (that is, particular red, round, or good things). Nominalists famously deny this, but equally famously, their attempts to make sense of properties without appealing to abstract universals tend either to be implausible or to entail a subtle commitment to universals after all.

All of this is controversial; indeed, that is precisely the point. The debate over these matters is simply no closer today to being *settled*, much less settled in favor of materialism or naturalism, than it ever was. There have always been critics of Platonism about mathematics, propositions, and properties, and they have always failed decisively to make their case. For all that, they may turn out to be

correct. But if so, no one has yet shown that they will. If naturalism about these purportedly abstract entities is favored by many today, that may be only because, as in the philosophy of mind, philosophers assume that naturalism or materialism has been somehow established in other contexts, and so must be the correct view to take in this one. But then the state of things in the debate over abstract objects cannot be appealed to as independent evidence of there being a reasonable presumption in favor of materialism generally.

The same thing appears to be true where the debate over the existence of God is concerned. There are, of course, a number of standard objections to the traditional arguments for God's existence. But there has also been in recent decades a great revival of interest among philosophers in the philosophy of religion in general and in the traditional theistic arguments in particular. Many contemporary philosophers of religion hold that the traditional arguments can be reformulated in a way that makes them immune to the usual objections, and that many of those objections rest in the first place on misunderstandings or even caricatures. So philosophically sophisticated is the work of these recent defenders of traditional religious belief, and so significant is the challenge it poses to atheistic naturalism, that the prominent atheist philosopher Quentin Smith has gone as far as to concede that "the great majority of naturalist philosophers have an unjustified belief that naturalism is true and an unjustified belief that theism (or supernaturalism) is false." Smith's view is not that these naturalistic philosophers are mistaken – as an atheist, he shares their naturalism – but rather that most of his fellow naturalists and atheists have not made a serious attempt to grapple with the powerful arguments that can and have been made for the other side, so that the level of confidence they have in the truth of their own position is unwarranted. The question of whether God exists is, in short, as live a philosophical issue as it ever was, and cannot reasonably be assumed to have been settled in a way that would provide support for a presumption in favor of naturalism and materialism.

A materialist could accept these points about the debate over mathematics, propositions, properties, and God (as Smith appears to do) – nothing said in this section shows, or is intended to show, that materialism is false. But to accept them would be to acknowledge that there is no basis for a *presumption* in favor of a materialist account of the mind. Such an account may have to stand or fall entirely on its own merits. Of course, if one can independently argue for a broadly naturalistic account of mathematics, propositions, properties, and the origins of the universe, then one could reasonably hold materialism to be the natural default position to take in the philosophy of mind. But by the same token, if one has instead independent reasons to endorse Platonism and/or theism, one would thereby have strong grounds for giving dualism the benefit of the doubt. The a priori plausibility of either side in the debate between materialism and dualism depends largely on the background metaphysical assumptions brought to bear in evaluating that debate. If those metaphysical issues have not been settled in favor of materialism, then there are no grounds for putting the burden of proof on the dualist.

Materialism, then, whatever its merits, may not be in quite as overwhelmingly strong a position as is often assumed. This is especially so when one considers that nothing said so far has really undermined the arguments for dualism discussed in the previous chapter. Even the claim made by some materialists that the mind's multiple realizability suffices to explain away the dualist conceivability argument is dubious: for the point of that argument is not that it is conceivable that the mind could exist in physical systems other than the brain, but rather that it is conceivable that it could exist apart from *anything physical at all*. So far we have seen no reason for doubting this.

Yet to *give* a reason for doubting it would seem necessary if materialism is to be established; and accomplishing this – showing that it is *not even conceivable* that the mind could exist apart from the physical world – is surely a tall order. If the interaction problem poses a difficulty for dualism, the dualistic arguments we've

examined pos
Accordingly, th
appeals to the
materialist ont
brain, etc., ulti
this, and show
qualia and con
appearances to
cannot conceivab
is in the details,
ability to accou
at last turn.

Further r

Materialism o
defended by
(Oxford: Black
is defended by
Mind, revised e
and J. D. Trout
Routledge, 19
Objections to P
contains essays

Reductionis
Kathleen Lenn
Clarendon Pre
Jaegwon Kim'
University Pres
in *The Conscio*
The "materiali
developed in
Francisco: West

examined pose an equally daunting challenge to materialism. Accordingly, the materialist has so far achieved stalemate at most, and appeals to the advance of science, the greater parsimony of a materialist ontology, general correlations between the mind and brain, etc., ultimately cannot break it. Materialists must go beyond this, and show that all the various specific aspects of the mind – qualia and consciousness, thought and intentionality – are, despite appearances to the contrary, purely material properties, features that *cannot conceivably* exist apart from some physical substrate. The devil is in the details, and materialism and dualism stand or fall with their ability to account for those details. It is to those details that we now at last turn.

Further reading

Materialism or naturalism as a general metaphysical position is defended by David Papineau in his *Philosophical Naturalism* (Oxford: Blackwell, 1993); as a theory of the mind in particular, it is defended by D. M. Armstrong in his *A Materialist Theory of the Mind*, revised edition (London: Routledge, 1993). Paul K. Moser and J. D. Trout, eds. *Contemporary Materialism: A Reader* (London: Routledge, 1995) is a useful anthology, as is Howard Robinson, ed. *Objections to Physicalism* (Oxford: Clarendon Press, 1993), which contains essays critical of materialism.

Reductionism is the subject of the articles in David Charles and Kathleen Lennon, eds., *Reduction, Explanation, and Realism* (Oxford: Clarendon Press, 1992). An influential work on supervenience is Jaegwon Kim's *Supervenience and Mind* (Cambridge: Cambridge University Press, 1993). Chalmers' "naturalistic dualism" is defended in *The Conscious Mind* (New York: Oxford University Press, 1996). The "materialist conceivability argument" outlined in the text is developed in chapter 10 of Peter van Inwagen, *Metaphysics* (San Francisco: Westview, 1993).

Behaviorism is most widely associated with Gilbert Ryle's *The Concept of Mind* (London: Hutchinson, 1949). The identity theory is famously presented in J. J. C. Smart's "Sensations and Brain Processes," anomalous monism in Donald Davidson's "Mental Events," and functionalism in D. M. Armstrong's "The Causal Theory of the Mind" and Hilary Putnam's "The Nature of Mental States." These classic essays are widely anthologized, and all four can be found (alongside other important related articles) in either David M. Rosenthal, ed. *The Nature of Mind* (New York: Oxford University Press, 1991) or David J. Chalmers, *Philosophy of Mind: Classical and Contemporary Readings* (New York: Oxford University Press, 2002).

The debate over the metaphysical status of numbers, propositions, and properties is surveyed in Michael Jubien, *Contemporary Metaphysics: An Introduction* (Oxford: Blackwell, 1997). Frege's argument for propositions as abstract entities can be found in his famous essay "Thought," reprinted in Michael Beaney, ed. *The Frege Reader* (Oxford: Blackwell, 1997). A more recent defense of the same idea (along with a response to a common epistemic objection to belief in abstract objects) is in chapter 6 of Alvin Plantinga, *Warrant and Proper Function* (New York: Oxford University Press, 1993). J. J. C. Smart and J. J. Haldane, *Atheism and Theism*, second edition (Oxford: Blackwell, 2003) contains a good overview of the recent debate over the existence of God and an excellent bibliography of recent work in the philosophy of religion. Quentin Smith discusses the current state of atheistic naturalism in "The Metaphilosophy of Naturalism," *Philo: A Journal of Philosophy*, vol. 4, no. 2 (Fall 2001).