DUALISM

Dualism is the theory that two and only two kinds of substance exist: minds and physical objects. A mind is a purely mental, non-material or spiritual substance, and a physical object is a purely material, non-mental, spatially extended substance. It logically follows that no mind is a physical object and no physical object is a mind. A person, on the dualist account, comprises both a mind and a body, but most dualists maintain that a person is essentially his or her mind but only contingently his or her body or, to put it another way, a person is his or her mind but a person has or owns his or her body. It follows that if a person's body should cease to exist it is logically possible that that person should continue to exist; but if a person's mind should cease to exist, then that person necessarily ceases to exist. In principle, minds may exist without bodies and bodies may exist without minds.

Mind-body dualism of this kind has been advocated repeatedly in the history of Western philosophy, and in this chapter we shall concentrate on two of its clearest and most celebrated exponents: Plato, the Athenian philosopher who lived and wrote in the fourth century BC and who arguably did most to initiate the problems and concerns of the Western intellectual tradition, and René Descartes, the seventeenth-century French philosopher and mathematician whose work is part of the transition from the theocentric world-picture of the Middle Ages to the more scientific ways of thinking of modern Europe.

The fact that I have chosen to examine the arguments of two figures drawn from the history of philosophy should not be taken to suggest that dualism is in some way out of date. It would, for example, be a serious mistake to think that the
progress of modern science has disproved the existence of the mind or the soul. On the contrary, some of our most sophisticated scientific and philosophical thinkers are mind–body dualists who take the reality of the mental most seriously. I have in mind in particular the Nobel Prize-winning neurobiologist Sir John Eccles, the philosopher of science Sir Karl Popper, and the Oxford philosopher of science and religion Richard Swinburne.

Nor should we be misled by the relation between dualism and religious, especially Christian, belief. The two are logically quite independent. Dualism might be true even if it is false that God exists, and God might exist even if dualism is false. In other words, the conjunction of dualism and atheism is not inconsistent, nor is the conjunction of the rejection of dualism with theism. It is true that Eccles and Swinburne are Christians, but Popper is an agnostic. Descartes was a Christian, but Plato's writings predate the life of Christ by four hundred years. Indeed, it is important historically to note that the immortality of the soul formed no part of Christianity before Paul and was only fully introduced into that religion by St Augustine of Hippo in the fourth century AD. The doctrine that an immaterial mind survives death was not part of Christ's teaching. He taught the resurrection of the body. It follows that mind–body dualism and orthodox Christianity are not only logically independent, they are incompatible.

There is a large variety of mind–body dualisms. Different taxonomies of dualisms may be generated, for example, by asking the following two questions: what are minds? and what is the relation between minds and bodies?

In their answers to the first question, all dualists agree that minds are both numerically and qualitatively distinct from those solid, spatio-temporally extended, publicly observable entities composed of matter called 'physical objects'. However, dualists disagree among themselves as to what minds are. For example, in their book *The Self and Its Brain*, Popper and Eccles argue that there exists a 'self-conscious mind', the states of which are not identical with any brain states, nor indeed with any sum of mental states.

There is nothing 'other-worldly' about their self-conscious mind. It is what you are, at least from when you wake up in the morning until you go to sleep at night. For Plato, Descartes and

Swinburne, however, the mind is identical with the immortal soul. In the writings of the fourth-century Greek philosopher-scientist, Aristotle, the soul is the 'form' of the body and so does not survive death. Nevertheless Aristotle holds that the mind, in the sense of 'intellect', is immortal. The Aristotelian philosopher-theologian St Thomas Aquinas holds that each person's immaterial soul survives that person's bodily death. However, a person does not thereby survive death because the person that one is is necessarily embodied. These points about surviving death bear closely on dualism because, if it may be proved that the mind is the soul and also proved that the soul is immortal, then it is proved that the mind is something distinct from the body because the body is mortal.

It is clear then that there are several dualist concepts of mind. They range from the commonsensical notion of the centre of conscious awareness which each of us seems to be, to the immaterial soul. It is important to note this variety because it is sometimes assumed that the claim, that consciousness exists, implies that immaterial minds exist. However, on the face of it it might be true that consciousness exists but false that immaterial minds exist. Both claims need argument.

The second way of classifying dualisms is according to the relations which allegedly hold between mind and body. Some dualists think causal relations hold between mind and body; other dualists deny this. Some dualists hold that events in minds are simultaneous with events in bodies; other dualists deny this. There is one relation which all dualists agree is not the relation between mind and body, and that is identity. Clearly, if the mind is identical with the body, or with any part of the body, dualism is false.

Descartes, Popper and Eccles maintain that events in minds cause events in bodies and events in bodies cause events in minds. So, on their dualistic theories there may be mental causes with physical effects and physical causes with mental effects. This view, that causal relations between mind and body are two-way, is known as 'psycho-physical interactionism' or just 'interactionism'.

Incompatible with that view is 'epiphenomenalism', the theory that there are physical causes of mental events but no mental
causes of physical events. The mental is thus causally dependent
on the physical. It is an ‘epiphenomenon’ of the physical just as,
in T. H. Huxley’s memorable image, the smoke above the factory
is causally dependent upon the workings of machinery. It is often
mistakenly said that epiphenomenalism is a kind of materialism.
When some people are asked whether dualism is true, they reply:
o, of course mental states are caused by states of the brain, or
some such. A widespread confusion needs to be dispelled here.
Suppose it is true, as it might well be, that mental states are
correlated by physical states of the brain. That fact, if it is a fact, by
no means refutes dualism. A non-material mind could pre-date
and post-date the states it is caused to be in by a brain. Also, if A
and B are causally related then it logically follows that A and B
are numerically distinct entities: A is not B and B is not A. If that
is right, then if mental events causally depend upon physical
events, that confirms dualism and does not refute it. Causal rela-
tions obtain only between distinct entities so if mental and physi-
cal are causally related, as epiphenomenalism has it, mental and
physical are distinct entities. Mental events are not physical
events: they are their mental effects.

Yet a different claim is that mental events are properties of
physical events, or properties of physical objects. Even then, we
still have a dualism of kinds: a dualism of physical objects and
their mental properties. This sort of dualism is inconsistent with
epiphenomenalism so long as it is impossible for something to be
the cause of its properties.

Several dualists maintain the view that there are no causal
relations between minds and bodies. The seventeenth-century
French philosopher Nicolas de Malebranche, for example, believed
that bodies cannot affect minds and minds cannot affect bodies,
yet both minds and bodies exist and no body is a mind and no
mind is a body. It is his view that the mind is not in the body and
the body is not in the mind, but that both exist only in God. It is
God who causes and sustains a parallel conjunction of mental
and physical events. Notice that it is not necessary to accept the
thesis that only God causes mental and physical events in order
to accept the thesis that mental and physical events do not inter-
act causally. Indeed, there is some intuitive or pre-philosophical
plausibility in the view that there cannot be mental causes of

physical effects or physical causes of mental effects. Consider the
fact that the brain, as a physical object, is only a collection of
atoms in motion. No matter how extraordinarily numerous and
no matter how tremendously complex the motions of the atoms of
the brain, it is hard to see how they could cause events like
hopes, fears, regrets, depressions, memories of yesterday and
images in the mind’s eye. Indeed this fact, if it is a fact, also
makes it hard to see how mental events could be identical with
physical events in the brain, as the materialist wishes to argue.

The dualist is struck by the intense qualitative differences
between the occurrences in one’s consciousness and mere matter
in motion which, however complex, is mere matter in motion.
Dualists agree that minds and physical objects are qualitatively
so different that the mental could not possibly be physical. Some
dualists, for example Malebranche, conclude from this that mental
and physical cannot causally interact. Other dualists, Descartes
and Popper for example, do not conclude this even though they
agree that nothing is both mental and physical.

The theory, that minds and bodies do not causally interact but
that mental events and physical events in a person are correlated,
is called ‘psycho-physical parallelism’ or sometimes just ‘par-
allelism’. Apart from Malebranche, this has been maintained by
another seventeenth-century philosopher: the German, G. W. Leib-
niz. Leibniz accepts that mental events may have mental effects
and physical events may have physical effects but denies that
mental events may have physical effects or physical events may
have mental effects. Leibniz invites us to compare the operations
of the mind and the operations of the body to two clocks, each of
which keeps perfect time. God has initiated the causal chain
which is the mental events in the mind and God has initiated the
causal chain which is the physical events in the body by a ‘pre-
established harmony’. This ensures an orderliness in mental and
physical events which may lead us to believe that the two kinds
of event are causally related. (As with Malebranche’s view we
may, if we wish, disagree that God causes mental and physical
chain of events and yet affirm that mental and physical do not
interact causally.) The difference between Malebranche’s par-
allelism and Leibniz’s is this: Leibniz thinks that God initiates
the two causal chains but does not subsequently intervene;
Malebranche thinks that God both initiates the two causal chains and intervenes as the real cause of every effect. The events which appear to us humans as genuine causes are only apparent causes or the 'occasions' for divine interventions. Hence Malebranche's particular brand of psycho-physical parallelism is sometimes known as 'occasionalism'. Leibniz and Malebranche may be thought of as trying to solve a problem bequeathed by Descartes: suppose dualism is true, how do mental and physical substance interact causally? Their solution is that they do not interact at all.

The dualisms of Descartes, Malebranche and Leibniz are consistent with the view that each mental event happens at the same time as some physical event with which (for Descartes) it may be causally related, and with which (for Malebranche and Leibniz) it may be caused by God to be simultaneous with. (I leave aside here the trivial point that if there are mental and physical events, and many of them, then it is likely that some events of each kind will happen simultaneously.) Eccles, by contrast, concludes from his neurological investigations that it is false that a given mental event is simultaneous with any physical (brain) event with which it is putatively identical or by which it is putatively caused. On the contrary, mental events may pre-date and/or post-date such brain events. It is in any case difficult to correlate mental events with brain events in order to establish either identities or causal dependencies because similar mental states are associated with qualitatively dissimilar neural patterns in different brains and in the same brain at different times. Eccles thinks the point about time is important for dualism. In this he might be right, because if A is B then A and B exist at all and only the same time, and if A wholly or partially pre-dates B or if B wholly or partially pre-dates A then it logically follows that A is not B. If mental events do not happen at the same time as physical events, then those mental events are not physical events and dualism is true.

Those, then, are the main sorts of dualism. I hope I have said enough to show that the truth or falsity of dualism does not depend upon the existence or non-existence of God, even if many dualists are also theists. Our starting point for discussion should be, so far as this is possible, some shared set of assumptions about mental and physical. We may wish to reject these assumptions later but, pre-philosophically, clods of earth that you turn over in digging the garden are physical but moods and emotions are mental. Lumps of metal are physical but images in the mind's eye are mental. Telephones, motor cars and houses are physical but perceptions of telephones and memories of motor cars are mental. The house is physical but as you stand and look at it your conscious awareness of it is mental.

To give some initial plausibility to dualism, or to show that on the face of it there is a distinction between mental and physical, consider the Battle of Waterloo first described just physically and then with a mental description added. In the purely physical description the trajectory and velocity of cannon balls and musket balls are mentioned, their striking human bodies or the earth. Sound waves make contact with the auditory nerves in these bodies and light waves with the retinas. Electro-chemical activity in those physical objects called 'brains' causes the movement of human bodies to and fro across the spatio-temporal region described. When a metal ball with a certain mass and velocity comes into contact with a human body, that body is damaged and gravity pulls it to the earth.

What is missing from this description? Many things, no doubt, but in particular everything mental is missing. There is no terror, desperate hope or deep-felt relief. There are no audible shouts of command, no fearful sights of masses of soldiers, no colours, no crash of musketry, no boom of cannon. There is no imagination of imminent death or injury, no prayers or pleas for survival, no acrid stench of blood and smoke; only silent, invisible matter in motion.

The point is this: no matter how complex and no matter how detailed the physical description, it does not capture what is mentioned in the mental description. No mental facts about the world logically follow from any list of purely physical facts, no matter how long. If that is right then the mental cannot be 'reduced' to the physical. The mental cannot be 'explained away'. The mental is all too real. The difference between the two Battles of Waterloo is that one is only physical and the other is both physical and mental. For the dualist it amounts to this: the Battle of Waterloo as fought by complicated robots, and the Battle of Waterloo with conscious minds participating. Physical objects seem to be spatio-temporal, publicly observable, tangible, solid
and objective. Minds seem to be non-spatial, private to the owner, intangible, ethereal and subjective.

We may turn now to some of the arguments for mind–body dualism presented by Plato and Descartes.

**PLATO**

As is well known, Plato wrote philosophy in the form of dialogues. These take place between Socrates, Plato's great partial contemporary in Athens, and a series of interlocutors. So far as we know, Socrates never wrote any philosophy but questioned people he met about philosophical problems. Plato wrote down the ensuing dialogues and produced the most outstanding works of Western philosophy. I shall not enter here into the vexed historical questions as to whether Plato's Socrates is the historical Socrates, or the extent to which Plato's dialogues are expressions of Plato's philosophy rather than Socrates'. We just need to extract the mind–body dualism from Plato's work, decide exactly what it is and see whether it might be true. It finds its most sustained expression in the dialogue called the *Phaedo*.

**The Cyclical Argument**

The topic of the *Phaedo* is Socrates' arguments for the immortality of the soul, and his defence of them against the criticisms of his two philosophical friends, Cebes and Semmias. The issue has an especially poignant importance because it supposedly takes place while Socrates is awaiting execution for allegedly corrupting the youth of Athens by teaching philosophy with them. The first argument he deploys concerns the coming to be and ceasing to be into and out of opposites. For this reason the argument is usually known as the Cyclical Argument. Socrates next asks Cebes whether he thinks living has an opposite and Cebes naturally replies that it has: being dead. Applying the general principle that opposites arise out of each other, Socrates concludes that living comes to be out of that which is dead and being dead comes to be out of that which is living: 'living people are born from the dead no less than dead people from the living' (*Phaedo*, 18).

Socrates takes this conclusion as evidence for reincarnation. As he puts it: 'the souls of the dead must exist somewhere whence they are born again' (*Phaedo*, 18), and the possibility of reincarnation clearly implies mind–body dualism. If one wished to question the soundness of this argument, two criticisms might be made. It seems relatively uncontroversial that, if something grows, then, it is larger than it was, or that, if something increases in strength, it was weaker than it is. But death might just be the cessation of life and that is all. It does not follow that being dead is a state in some ways rather like being alive but in other ways very different. Being dead might be not being at all. Secondly, it might be objected that the relevant opposite of 'alive' is not 'dead' but 'not alive' or 'inaniment'. Then, if a person is alive, what he has 'come to be out of' might be merely some inanimate matter and not some pre-natal, non-physical realm. It would seem to require further argument to show that if a person is not alive then he or she still exists, but as a non-material soul.

**The Recollection Argument**

However, Socrates deploys a second argument designed to show just that. The so-called Recollection Argument has as its conclusion that our souls existed before we were born. This thesis also clearly implies dualism because, if one's soul existed before one's body, then clearly one's soul cannot be identical with one's body. The argument appears in many places in Plato's dialogues, most notably in the *Meno*. There Socrates teaches some geometry to Meno's slave boy, but the tuition has a peculiar form. It consists in eliciting from the slave boy what he already knows but has forgotten. Socrates questions him about the mathematical properties of certain geometrical figures and receives true replies which the slave boy had not been taught. It is concluded on this basis that the knowledge we have is in fact all recollection. It was not acquired in the physical world we currently experience, so it must have been acquired in some non-physical world which we previously experienced. In the *Phaedo* this conclusion is used directly to maintain that the soul is immortal. Socrates asserts that 'our learning is nothing but recollection' (*Phaedo*, 19), and that this would be impossible unless the soul existed previously, independently of the human body, so the conclusion is: 'the soul is something immortal' (*Phaedo*, 19).

Cebes and Simmias present a fairly obvious objection to this
soul, am compounded or intermingled with my body? If souls are not spatial, it is hard to see how they could have any spatial relationship to a body – like a pilot in a ship or not like a pilot in a ship. The image of the pilot suggests that the soul is located in the brain, or perhaps behind one’s eyes. Am I what looks through my eyes? The idea of the soul’s not just being thus located suggests its pervading the whole body. Do I pervade my body?

In the sixth Meditation Descartes uses the word ‘union’ to denote the relation between mind and body (Meditations, 159). Being a soul is what my existence essentially consists in; but I am, at least at present, united with this particular body. I think perhaps the best way to make sense of this is to think of one’s own body in objective, third-person terms – to think that there should be a functioning human body, just like one’s own, which is different from one’s own in just one respect: it is not yours. Now think of one’s consciousness as existing in this world, and of this body thereby being your body. This body is then no longer just one body among others. It is, for example, the one you look out of and the one you are, in some peculiar sense, co-extensive with. This is an interpretation of what Descartes means by ‘union’ which takes seriously his couching his argument in first-person-singular terms.

Ever since Descartes wrote, the relationship between mind and body has been thought the least satisfactory aspect of his dualism. As we have seen, even Descartes himself did not think this relation could be made fully intelligible. Some philosophers – idealists, materialists and others – have thought the problem fatal to the plausibility of dualism. It seems to me that the problem may not be insuperable and is perhaps part of a wider question about the self. There is a deep sense in which, ironically, we do not know what we are, and this is an area for further philosophical investigation. A large part of Descartes’ greatness as a philosopher lies in his contributions to that area and we may learn from his insights and arguments, even if in the end we should wish to repudiate his strict mind–body dualism.

Logical behaviourism is the theory that being in a mental state is being in a behavioural state. Thinking, hoping, perceiving, remembering and so on are all to be understood as either behaving or else possessing a complex disposition or propensity to behave. The mind is nothing over and above behaviour, whereby ‘behaviour’ is meant publicly observable bodily behaviour. This reduction of the mental to the behavioural is advocated by the logical behaviourist as a linguistic thesis: a thesis about how it is possible for psychological concepts like ‘image’, ‘perception’, ‘thought’ or ‘memory’ to have a use in our language. This is possible, according to the logical behaviourist, because any sentence or set of sentences about minds may be translated, without loss of meaning, into a sentence or set of sentences about publicly observable behaviour. This is the essence of logical behaviourism. Unless our psychological vocabulary referred to overt behaviour, it could not be meaningful. Logical behaviourists differ among themselves as to why this should be so. Some maintain that there would be no way of deciding the truth or falsity of psychological claims unless these were behavioural claims. Others hold that psychological concepts could not have a role in our public language unless there exist publicly available criteria for their use. All logical behaviourists are agreed that, unless our psychological language is about behaviour, it is about nothing.

Logical behaviourism needs to be carefully distinguished from behaviourism in psychology. Behaviourism in psychology is a method for studying human beings. It is not a doctrine about the meanings of psychological concepts, nor a putative solution to the mind–body problem. It is the view common to the American psychologists J.B. Watson and B.F. Skinner that all human
behaviour can be explained as a set of responses to stimuli to which a person is subjected. Neurological facts are not invoked by behaviourists any more than the findings of introspection. It is maintained that knowing the causes of human behaviour—which stimuli cause which responses—is sufficient for explaining that behaviour. It is certainly true that the psychological behaviourists sometimes make pronouncements of a quasi-philosophical kind: Watson, for example, thinks that consciousness does not exist. These pronouncements, however, are no part of their behavioural method. That is the attempt to predict and control human behaviour through the study of its environmental causes.

Indeed, the merits and demerits of behaviourism as a method in psychology are logically independent of putative solutions to the mind–body problem. I mean, for example, that even if mind–body dualism is true, behaviourism might be the best method for explaining behaviour and, even if materialism is true, behaviourism might not be the best method for explaining behaviour.

Even though psychological behaviourism and logical behaviourism are quite distinct, and even though the practice of psychological behaviourism is logically consistent with various ontologies of the mental, logical behaviourism may be construed as a philosophical legitimation of psychological behaviourism. This is because, if all meaningful psychological language is really behavioural language, then behavioural psychology is the only meaningful kind of psychology. Putative rivals to behavioural psychology may then be ruled out a priori. Also, logical behaviourism may partly justify psychological behaviourism’s claim to be genuinely scientific. Watson and Skinner think one of the hallmarks of genuine science is the study of some publicly observable subject matter. If the subject matter of psychology is private and subjective, psychology’s being scientific in that sense is impossible. But if it may be proved, as the logical behaviourists argue, that the mental is really behavioural then psychology is guaranteed a publicly observable subject matter. Logical behaviourism seems to open the way for a scientific psychology.

The two logical behaviourists I have selected for examination in this chapter are the North American philosopher of science, Carl Hempel, and the English philosopher, Gilbert Ryle. Each is an important exponent of a movement in twentieth-century philosophy: Hempel is a logical positivist and Ryle is a linguistic philosopher. These two philosophies are mutually distinct. Logical positivism is essentially the view that every genuine problem may be solved scientifically and that putative non-scientific ways of finding out about the universe are literally meaningless. Linguistic philosophy is the view that philosophical problems like the mind–body problem arise from the misuse of our ordinary non-scientific language. I shall say something more about logical positivism and linguistic philosophy in the sections on Hempel and Ryle, below.

I also include a discussion of some of the massively influential work of the later Wittgenstein. Wittgenstein cannot correctly be called a ‘logical behaviourist’ in any clear or straightforward sense; his thoughts are too complex and subtle for such an easy taxonomy. However, his anti-Cartesianism bears a closer resemblance to logical behaviourism than any of the other views discussed in this book.

HEMPEL

To understand what logical behaviourism is, it is useful to compare it with two other important developments in modern philosophy. Logical behaviourism grew out of and, in a sense, is a philosophical extrapolation of logical positivism; and in its strategy it bears a relationship to certain remarks made by the Austrian philosopher, Wittgenstein, in his later work. I shall say something about logical positivism now but postpone discussion of Wittgenstein until the end of this chapter.

The Vienna Circle

Positivism is the doctrine that any phenomenon may in principle be explained using the techniques of the natural sciences. The logical positivists who met in Vienna in the 1930s to form the so-called ‘Vienna Circle’ tried to reformulate philosophical problems so that they could be solved scientifically. To achieve this, they employed a specific criterion for what was to count as meaningful, and any sentence in philosophy which did not meet this criterion was categorized as meaningless. This criterion for distinguishing
Idealism is the theory that only minds exist. It is clearly incompatible with dualism because that is the theory that both minds and physical objects exist. The implication of idealism, that physical objects do not exist, is potentially misleading and usually misunderstood. No idealist denies that there exist tables, bricks, pieces of wood and all other items we commonly take to be physical objects. Idealists merely deny two alleged facts about physical objects: they deny that they exist unperceived or unthought; and they deny that they are material, that is, made out of a substance called 'matter'. This is, perhaps, denying the existence of the physical objects in the sense of denying their essential properties, but it is clearly not the claim that there are voids or gaps in our visual or tactile field where most people take there to be physical objects. So far as I know, no philosopher has held that.

A central claim of the idealist is that physical objects do not exist independently of minds. Unless there were minds, so-called physical objects could not exist. What exists exists only within consciousness, so that what we commonsensically or pre-philosophically take to be physical is in fact mental. The physical is mental.

Many people, when they hear about idealism for the first time, think it is a foolish theory: I mean so grossly and manifestly false and so hard to believe that only a philosopher divorced from the concerns of everyday life or one with dubious theological commitments could have formulated it. Sometimes, too, idealism is thought to be unscientific in the sense of incompatible with the findings of modern science.

None of these charges has any foundation. The theory that the
physical is mental is consistent with any everyday, practical commitment, consistent with atheism, and consistent with modern science. Contrary to a widespread misconception, nothing at all in modern science rules out idealism. All the sentences of physics, biology and even neurology could be true, and all the sentences of idealism true too; their conjunction forms a self-consistent set. Obviously, idealism is incompatible with the old Newtonian idea of matter as a kind of material or stuff that the universe is made of, but physics gave up that idea long ago. Indeed, the idealist attack on matter anticipates in important respects the demonstration of the limitations of Newtonian physics by Mach and Einstein. Ironically, it might well turn out that, while idealism is compatible with science, materialism is incompatible with science.

In what follows I shall try to provide some intuitive plausibility for idealism, simply because it seems to so many people intuitively false. First, though I shall say something about the varieties of idealism. I have been talking as though idealism were a single homogeneous theory, but this is not really so. All idealists agree that in some sense reality is ultimately spiritual or mental, but there are, nevertheless, radically different kinds of idealism.

Idealism is essentially a British and German movement in modern philosophy. Idealism was not a doctrine in ancient Greek philosophy or in medieval Christian philosophy. An exception to this generalization is the dialectical and spiritualist metaphysics of the neo-Platonist Greek-Egyptian philosopher, Plotinus, who lived and wrote in the third century AD. Plotinus argued that individual human consciousnesses are perspectives or points of view that the One unifying, cosmic consciousness has on the universe which, in a sense, it constitutes. In Plotinus's thought there are massive anticipations of the central tenets of German idealism. Indeed, in my view it is no exaggeration to say that German idealism is essentially neo-Platonism. Plato himself was not an idealist. Plato was a mind-body dualist. The Platonic theory of the Forms, for example, does not commit Plato to idealism because the Forms are not mental or spiritual, even though material particulars depend upon the Forms for their existence. Apart from Plotinus, then, there are no ancient idealists in the Western intellectual tradition.

The two idealists I have selected for study are the eighteenth-century Irish philosopher George Berkeley, and the nineteenth-century German philosopher G. W. F. Hegel. Berkeley thinks that physical objects do not exist over and above ideas, ideas either in the infinite mind of God or ideas in the finite minds of persons such as you and I. Hegel thinks that the physical world is ultimately to be understood as an expression of Spirit, a pantheistic cosmic mind for which human consciousnesses are agents or (rather as Plotinus thought) points of view.

Berkeley and Hegel are very different thinkers. Berkeley is an empiricist, that is, he thinks all knowledge depends ultimately upon experience. Hegel is a rationalist in that he thinks it is possible to obtain by intellectual means knowledge of the essential properties of reality as a whole. The differences between the two should not be overemphasized, however. Berkeley is a theist, and there can be no conclusive empirical proof of the existence of God; and, while Hegel's 'logical' writings exhibit his rationalism, an empirical or experiential side of his philosophy is found in his 'phenomenology'. Idealism is logically independent of both empiricism and rationalism.

I have excluded from consideration the critical philosophy of Immanuel Kant as he expounds it in his book The Critique of Pure Reason, even though that philosophy was called by him 'transcendental idealism'. Kant has been read as an idealist by almost every commentator since Hegel, but I think this view of his work is almost certainly false. Kant himself tried to dispel the misreading by including the chapter 'The refutation of idealism' in the second edition of The Critique of Pure Reason but the legend has persisted none the less. (I cannot argue the point in this book, but I think Kant was not an idealist but a materialist philosopher.)

The German idealists who have systems partially similar to Hegel's are J. G. Fichte and F. W. J. Schelling. It is in their hands that transcendental idealism becomes a kind of idealism. I shall say something briefly about their rather convoluted philosophical systems.

Fichte's book The Science of Knowledge is a synthesis of metaphysics and epistemology. Metaphysics is the study of what exists as it really or essentially is in itself. Epistemology is the
study of philosophical problems about knowledge. By generating an idealist metaphysics, Fichte also hoped to define the limits of knowledge. What knows and what is most fully real is the individual, purely spiritual Ego or subjective consciousness. The Ego is irreducibly active and through its mental acts it is conscious of itself. The Ego's self-consciousness depends upon a self—not self distinction or upon an Ego-world distinction which is itself created by the Ego's 'positing' of the empirical world. The world of spatio-temporal enduring items which we confront in everyday life is, so to speak, constructed out of the experiences of the Ego, or, more precisely, out of appearances which the Ego is acquainted with. Because what passes for the objective world is an intellectual construct by oneself as Ego, Fichte's idealism is sometimes known as 'subjective idealism'. Hegel calls it that.

Fichte's partial contemporary, Schelling, thinks of his own 'absolute idealism' as a holistic unification or overcoming of a set of dualisms or separations which are in the last resort artificial or unreal. In reality there is no genuine bifurcation between object and perception, concept and image, person and external world. All these are created by conscious reflection, and are most appropriately viewed as aspects of a spiritual and unitary whole. Schelling considers persons as 'separated' from themselves by acts of reflection, and it is one ambition of his philosophy to reconcile persons with themselves. Reflection, because it produces psychologically damaging and metaphysically illusory dualisms, is regarded by Schelling as a kind of mental illness and he seeks to replace the oppositions of ordinary thinking by a 'philosophy of identity' which exhibits the mutual dependence between opposites. Philosophy itself, in so far as it is dualistic, is an evil, albeit a necessary one in the sense of necessary prerequisite for the formulation of his own philosophy. When we live and act without reflection, we are not conscious of dualisms between, say, objects and our mental representations of them. In this, and especially in the thesis that the overcoming of dualisms is in a universal spirit, Schelling anticipates many themes in Hegel's system.

Although Berkeley is the most celebrated British idealist, towards the end of the nineteenth century British philosophy was dominated by a group of thinkers deeply influenced by German idealism, especially by Hegel. This group comprised F. H. Bradley, Bernard Bosanquet, Thomas Hill Green and John McTaggart Ellis McTaggart. These thinkers are little read nowadays, mainly because of the severity of criticism levelled at them, first by G. E. Moore and Bertrand Russell and later by the logical positivists. However, their writings display logical rigour and considerable metaphysical imagination. They are, in my view, just as worthy of study as their anti-metaphysical opponents.

In his Appearance and Reality (1893) Bradley argues that it is not possible to specify what something is, independently of specifying that thing's relations to other things. In particular, rather like Fichte, Bradley thinks it is not possible to make sense of the idea of self except in contrast with a not-self. The 'Absolute' is Bradley's name for the sum of all relations in their unity. It is a spiritual whole, which is more than the sum of those parts which are appearances to conscious selves, and it is more real than the physical world, which Bradley regards as an ideal construction or useful fiction postulated by the natural sciences.

Like Bradley, Bernard Bosanquet thinks that dualisms or oppositions are united in reality as a whole, which he too calls the Absolute. The most fundamental philosophical dualism to be overcome is that between individual and universal, or things and sorts of things. The unity of universal and particular is expressed in the political, religious and artistic progress of conscious beings. The Absolute, or reality as a whole, is best understood on the model of self-consciousness. We finite minds are parts or aspects of a single infinite mind which is identical with the Absolute, and the physical universe does not exist independently of its experience by finite minds. Bosanquet advocates this quasi-Hegelian metaphysics in his book The Principle of Individuality and Value (1912).

Thomas Hill Green acknowledges that mental events may have physical causes but argues that everything physical is nothing but a constituent of the world of experience, part of the content of experience. He concludes from this that what makes the whole of experience possible cannot be anything physical. Both experience and its contents depend, in fact, on certain spiritual principles which Green characterizes in his book Prolegomena to Ethics (1883).

In his two-volume work The Nature of Existence (1927) —
perhaps the master-work of this phase of British idealism – McTaggart argues that conscious spiritual selves are the fundamental constituents of the universe. Each of us is essentially one of these, and the empirical world logically depends upon its perception by us for its existence. According to McTaggart, there is no such substance as matter and, as spiritual selves, we are immortal. Interestingly, McTaggart rejects the claim that God exists, so combining a doctrine of personal immortality with atheism. This perhaps illustrates my point that theism and belief in one's survival of death are logically independent (see page 2, above).

Two contemporary British idealists with systems very different from one another are John Foster and Timothy Sprigge. Foster advocates an idealist phenomenology in his The Case for Idealism (1982) and Sprigge a pan-psychic neo-Hegelianism in his The Vindication of Absolute Idealism (1983). Foster argues that the physical world is a logical construction out of sense contents and that ultimate contingent reality is at least not physical and plausibly mental. Sprigge argues that pan-psychism, the doctrine that everything that exists has at least one mental property, is the only metaphysics that adequately recognizes the existence of conscious subjects. Indeed, for Sprigge, consciousesses are the ultimate constituents of the universe. If they did not exist, nor would anything else.

Why should anyone believe idealism? Two kinds of argument for idealism, one empirical and one metaphysical, can be discerned in this variety of thinkers. The empirical argument is that physical things are known to exist only through our perception of them. Further, physical things are not known with certainty to exist over and above our perception of them. Finally, it might even be incoherent – contradictory – to suppose that physical things exist independently of our perceptions at all. The metaphysical argument goes as follows. Science and much empirical thinking give us only a partial account of reality. This is because science and empirical thinking are essentially objective or third-person in their approach. They treat their subject matter as ‘other’. Neither science nor empirical thinking can provide an account of subjectivity, in particular of consciousness, which is a subjective or first-person phenomenon. It follows too that purely objective modes of thought cannot explain the relationship between conscious subjects and the objects they experience. Science treats everything as physical. It cannot explain consciousness and it cannot explain the location of the conscious subject in the universe. Idealism explains just what science cannot explain.

**BERKELEY**

Berkeley's idealism is best understood against the background of his empiricism. Empiricism is the view that all knowledge is acquired through the use of the five senses. Sometimes it is held, in the weaker form: all knowledge is acquired through experience; but nearly all empiricists agree that no experience would be possible if sense experience were not possible. Berkeley shares his empiricism with the seventeenth-century English philosophers, Thomas Hobbes and John Locke, and the eighteenth-century Scottish philosopher, David Hume. Empiricism by no means necessarily leads to idealism: Hobbes, for example, was a materialist and Locke a dualist. But Berkeley thinks that strictly correct empiricist thinking should lead to idealism. He takes the view not only that knowledge is all acquired through experience but that only experience can be certainly known to exist. One's experiences are all one is ever directly acquainted with, so belief in anything else cannot be based on direct experience. Thus we can understand Berkeley's idealism as an extreme form of empiricism. He conceives of himself as pushing empiricism to its logical conclusions.

Experience gives us absolutely no evidence for belief in two features of physical objects which we normally take to be essential to them. Experience does not teach us they exist unperceived; that they exist before and after not just during our perception of them. Experience does not teach us that they are material, that is, ultimately composed of a substance called ‘matter’. Berkeley also thinks there is no sound philosophical justification for these two fundamental assumptions, and he argues they are both false.

**Matter in Question**

To understand why Berkeley denies the existence of material
4 MATERIALISM

Materialism is the theory that if something exists then it is physical. For example, the theses that only physical objects exist and only physical events exist are both materialist theses. Sometimes the doctrine is held in a weaker form which allows that physical objects or events may possess non-physical properties — for example, mental or abstract properties — but even then it is maintained that any such putatively non-physical properties logically depend upon the existence of physical things. No non-controversial definition of ‘physical’ has ever been given, but it is thought that if something is physical then it is spatio-temporal. In addition, physical objects are those objects which possess essentially the properties of shape, size and solidity and are capable of motion. More controversially, materialists usually maintain that physical things are composed of a substance called ‘matter’. Indeed, the term ‘materialism’ is sometimes used to denote the theory that matter exists, as well as the theory that everything that exists is physical. ‘Matter’ has proved even more recalcitrant to definition than ‘physical’.

Several versions of materialism have been advocated in the history of philosophy. In what follows I shall concentrate on a particularly sophisticated and influential, twentieth-century version: the so-called mind–brain identity theory. This is the view that any mental event is literally identical with some event or state in the brain. It has been developed by two contemporary Australian philosophers, J. J. C. Smart and David Armstrong, but finds an early and succinct expression in a 1956 paper by the English philosopher and psychologist U. T. Place. The paper is called ‘Is consciousness a brain process?’. After examining the version of materialism defended by Place there, I shall turn to an important adaptation of the mind–brain identity theory by the North American philosopher Donald Davidson. Finally, I shall examine the philosophy of mind of a materialist who finds shortcomings with the mind–brain identity theory and who advocates instead a materialistic or physicalist non-identity theory: Professor Ted Honderich of the University of London.

Before discussing these three contemporary views, however, I shall distinguish some varieties of materialism through their exponents in the history of philosophy.

Democritus, the Greek atomist philosopher who wrote in the fourth century BC, maintained that everything that exists is composed of physical objects which are so minute as to be imperceptible. These he called ‘atoms’. The word ‘atom’, in its etymology, means ‘indivisible’ and it is indeed Democritus’s view that the atoms are indivisible and impenetrable. They are physical because they have the properties (size and shape) essential to any physical object. Not only do atoms exist, according to Democritus, but only atoms exist. Everything that exists is either an atom or a collection of atoms, so it is clear that Democritus is a thoroughgoing materialist. The atoms are located in what Democritus calls ‘the void’ — pure nothingness — absence of being. All the atoms are in motion and there exist an infinite number of them, but if we ask what the atoms move in or what they move through, the only true answer on the Democritian view is ‘the void’ or ‘nothing at all’.

The atomism of Democritus was used as a set of premisses for a humanistic ethical philosophy by Epicurus, also a Greek, who lived later in the same century. Epicurus thought that all the atoms move downwards but not every atom moves parallel to all the others, so collisions sometimes occur. The material universe which we inhabit is the result of one such original collision. The fact that collisions occur between atoms introduces a certain indeterminacy into the universe which makes it impossible to predict human actions with certainty, and Epicurus thinks that this leaves room in the universe for human freedom. It follows that Epicurus combines materialism and libertarianism — the view that humans have free will — in his philosophy. It is clear too that, although Epicurus thought that the gods have no influence over human affairs, and although the human soul is not immortal
because it is just a collection of atoms that will disperse, he is nevertheless not an atheist. The gods do exist but they too are composed of material atoms, albeit atoms of an especially minute or rarefied kind. It follows that Epicurus not only combines libertarianism with materialism but also theism with materialism.

Theism and materialism are also combined by the seventeenth-century English philosopher and political theorist, Thomas Hobbes. According to Hobbes, everything that exists has the physical dimensions of size, that is, everything that exists has length, depth and breadth. Also, everything that exists is composed of matter. He not only maintains that all our thoughts and sensations are caused by physical objects, he also maintains that those very thoughts and sensations are themselves physical. Hobbes does not deny that God and the soul exist, he merely denies that they are immaterial. God and the soul are physical entities, but the particles of matter which compose them are so minute as to be undetectable by the human senses. For this reason it is impossible for us to perceive God or the soul. Hobbes notes that the soul allegedly travels to heaven or to hell, where it may burn, and that souls in the form of ghosts allegedly haunt people on earth. His thought about these phenomena is that their intelligibility depends upon their being conceived in physical terms. To talk of travelling, or burning, or being present in a churchyard is necessarily to talk in physical terms. To think of something non-physical doing these things is patently absurd, in Hobbes's view. This is a powerful materialist thought: when we think we have succeeded in thinking of something non-physical, we have in fact only succeeded in imagining something physical, for example an invisible or perhaps transparent physical object.

Hobbes is both a materialist and an empiricist. He thinks that everything is physical, and everything known depends upon observation through the five senses. These two positions, the first ontological and the second epistemological, are logically independent of each other. Nevertheless, not only Hobbes but also the French Enlightenment thinker, de la Mettrie, subscribed to both philosophies. De la Mettrie tried to present a purely mechanistic account of human thought and action in his L'Homme Machine (1748). His project was adequately to describe a person as a highly complicated physical object with moving parts and hence to dispense with the Cartesian notion of the person as essentially an immaterial soul. Thought and sensation, according to de la Mettrie, are nothing over and above the complicated motion of matter. The most complete expression of the materialist strain in the French Enlightenment is Baron d'Holbach's System of Nature (1770). D'Holbach argues that the universe is one enormous deterministic system of physical objects, and that nothing else exists.

The most famous materialist of the Western intellectual tradition was not a materialist at all. I am thinking of the nineteenth-century revolutionary thinker, Karl Marx. True, the topic of Marx's doctoral dissertation was the materialism of Democritus, mentioned above. True, Marx did try to reverse the order of priorities in the philosophy of his idealist predecessor, Hegel, by 'turning Hegel on his head'. It is important to note, however, that Marx, when called a materialist, is called a special sort of materialist. He is sometimes called a 'dialectical materialist' and sometimes a 'historical materialist'. What both of these views entail is that it is the material, especially the economic, facts about a society which determine that society's other features. What laws, religion and other patterns of thinking obtain in a society is closely dependent upon the way in which that society is organized economically: the way in which it can reproduce itself. Now, the view that the physical determines the mental is not the same as materialism. Materialism is the view that the mental is the physical. Materialism is the theory that all the mental facts are nothing over and above all the physical facts, and there is nothing in Marx's writings to suggest that Marx was a materialist in that sense. The view that the mental is determined by the physical is consistent with most of the ontologies of the mind described in this book. To show that determinism implies materialism requires special argument, in fact argument of the sort deployed by Hölderich and to be examined later in this chapter.

Ironically, Marx's determinism may be inconsistent with materialism. Arguably, if A determines B then B is 'something over and above' A. If Marx did adopt an ontology, then this was the 'naturalism' he describes in the Economic and Philosophical Manuscripts of 1844. His naturalism, he claims, will dissolve the contradiction between idealism and materialism. Marx, I think, was
not much interested in the mind–body problem. His concern was the overthrow of capitalist society and its replacement by communism. His materialism is a doctrine of material determinism and not an ontological materialism. It is fine to call Marx a materialist so long as we use the term in this special sense.

Could materialism be true? Could it be true that you and I are just highly complicated physical objects? One common misunderstanding needs to be dispelled straight away. When people hear about materialism for the first time, they often say: no, I am not just a complicated physical object because I think, I perceive, I have emotions, and so on. Now, it is not a very good objection to materialism to mention the fact that people think. No materialist, or hardly any materialist, would deny this. Materialists accept - just as you or I or anyone else would - that we all have thoughts, perceptions, emotions, mental images. The materialist is not denying that we think. The materialist is saying that our thoughts are physical. That fully fledged mental life each of us has is a series of physical events - a set of electro-chemical processes in the brain, according to the most modern version of materialism. Matter can think. You and I and everyone else are thinking matter.

We may turn now to the materialism of Place, Davidson and Honderich.

PLACE

When Place suggests that consciousness is a brain process, he does not mean that our thoughts and experiences are merely caused by events in the brain; he means that those mental events are exactly the same events as events in the brain. A causal correlation between mental and physical is not of itself sufficient for the truth of materialism, because a dualist might plausibly maintain that mental events are caused by physical events yet cling to the view that those mental events nevertheless occur in a non-material mind. So Place wants a stronger theory than causal correlation.

A Scientific Hypothesis

He says that he advocates that consciousness is a brain process as a scientific hypothesis and goes on to argue that this claim cannot be dismissed on logical grounds alone. What does Place mean by calling 'consciousness is a brain process' a scientific hypothesis? A hypothesis is a statement which is designed to be a solution to a problem and which is open to test for its truth or falsity. If a statement has the status of a hypothesis, then we do not yet know whether it is true or false, but it is true or false. We are putting it forward for confirmation or refutation. By calling the hypothesis 'scientific', Place means that it is the procedures of the natural sciences - especially perhaps neuro-psychology - which will demonstrate its truth or falsity. So we need to be clear that Place does not take himself to have proved the mind–brain identity theory. He conceives of himself as removing logical obstacles to its plausibility so that it may be handed over for use by scientists in explaining what consciousness is.

Place's first clarification of his thesis is a denial that sentences about consciousness mean the same as sentences about processes in the brain. He takes it that his thesis would be manifestly false if it was that our terms for sensations and mental images were semantically equivalent to our terms for describing the central nervous system. This is clearly right to the degree that what a person intends to convey by saying, for example, that they are in pain is not that they are in a certain neurological state but that they are undergoing a certain sensation. The two sorts of meaning seem to be quite distinct, whether or not the mind–brain identity theory is true. So even if pain is exactly the same thing as C-fibre stimulation, 'pain' does not mean 'C-fibre stimulation'. It is worth pointing out here Place's relationship to logical behaviourism. Place thinks a logical behaviourist programme for translating our dispositional mental concepts - like 'knowing', 'believing' and 'intending' - into sentences about our actual and possible behaviour is fundamentally sound. Nothing essential to their meaning would be lost. But he thinks our occurrent mental concepts like 'is having an after image' or 'is in pain' contain an irreducible experiential dimension which resists the analytical translation attempt. He allows that in principle there may come a time when the behaviourist thesis may be extended to cover concepts of occurrent states also, but Place is pessimistic about the current prospects for this. So we may read Place as in a sense
which are personal, some moral and some political. They are: (1) What may we each realistically hope for in the remainder of our life? (2) How should we feel about others who affect us for good or ill? (3) Is it possible for us to know the truth? (4) Should we hold ourselves and others morally responsible for actions? (5) What is the moral worth of people? (6) Should particular actions be judged right or wrong? (7) What should politics be like? Each of these is an enormous and complex philosophical question, and to explore them would unfortunately take us beyond the scope of this book. I should say, however, that Honderich distinguishes two families of response to these questions: one based on the idea of an originator — a self as an uncaused cause of actions — and one which rejects such an originator. It is the second of these which is compatible with his own theory of determinism. For example, with regard to question (7), political conservatism must be ruled out in so far as it presupposes an originator — an individual self who is not caused to be what he or she is. The theory of determinism is no threat to the hopes we have about the future that are worth having. We may have 'life hopes' even if determined where 'to have a life hope ... is either to hope that one will achieve a certain thing, or else that some state of affairs will come to obtain' (A Theory of Determinism, p. 508). The idea of punishment as simple retribution should be given up because people are caused to do what they do, they do not act as uncaused causes. Honderich ends his book by distinguishing two fundamental attitudes to life: dismay and affirmation. He urges affirmation.

5

FUNCTIONALISM

Functionalism is the theory that being in a mental state is being in a functional state. A functional state is a state which may be individuated or picked out in virtue of its causal relations; so a mental state is one with a particular kind of cause, say a sensory input, and a particular kind of effect, say a behavioural output. Mental states are also causally related to one another, and the totality of the causal relations which a given mental state enters into is known as that state's 'causal role' or, sometimes, 'functional role'. Being a particular sort of mental state is having a particular sort of functional role. Being a particular mental state, a particular one, is having a particular functional role, just that functional role.

Functionalism is, in a sense, an attempt to bypass the mind-body problem. It provides a philosophical framework within which to devise a scientific psychology without any need to address the ontology of the person. It is notable, however, that many functionalists are also materialists. Many people who think that the answer to the question, what is the mind?, is the answer to the question, what is the mind for?, also think that mental states are physical states, and that a person is just a highly complicated physical object. It is necessary, therefore, first to say something about the relation between functionalism and materialism.

Functionalism and materialism are not the same theory. Functionalism is the theory that a mental state is essentially an effect of some perceptual input and a cause of some behavioural output. Mental states are also among the causes and effects of one another. Materialism is the theory that every mental state is identical with some physical state, say a state of the brain or the rest of
the central nervous system. Now, clearly, somebody who is functionalist may be tempted by a materialist view of the person, because brain states seem good candidates for the effects of sensory inputs and the causes of behavioural outputs. Clearly also, brain states stand in causal relations to one another. It is important to note, however, that functionalism per se does not logically entail materialism. From the fact that there exist mental states, states with a particular kind of causal role, it does not logically follow that those mental states are physical states. That they are physical states, for example brain states, is a scientific hypothesis not a necessary consequence of functionalism.

Functionalism does entail materialism with the addition of just one extra premiss. This is: all causes and effects are physical causes and effects. Clearly, if only physical items - physical states, events, objects and so on - may enter into causal relations, then, if functionalism is true, materialism is true. Functionalism entails that every mental state is both a cause and an effect. The new premiss is that only physical states may be causes or effects or both. It follows that, if functionalism is true, then every mental state is a physical state: materialism.

The premiss that, if something is a cause or an effect, it is physical needs argument; but two considerations may be adduced here in its favour. We saw in the chapter on dualism that there is a metaphysical problem as to how qualitatively distinct substances may interact causally. If mental causes and effects are physical causes and effects, that problem need not be addressed. Also, a non-physical cause of a physical effect or a non-physical effect of a physical cause could not be incorporated into the system of scientific laws as we know them. Such relations would be outrageously at variance with any natural scientific view of the universe. They would be 'something new under the sun'. Despite this, I see no sound and so decisive argument in favour of the thesis that causal relations hold only between physical entities.

Functionalists therefore divide between those who are materialists and those who feel committed to no particular ontology of the mind. All functionalists hold that a particular kind of mental state may be realized in a number of different ways; for example, on a functionalist view, being in pain is being in a functional

state. Being in pain is being in a state which results from certain sensory inputs and causes pain-behaviour. This account holds true for any beings capable of being in pain: humans, cats, dogs, Martians. However, the way in which the state of being in pain is realized may differ radically according to the make-up of these beings. Pain may be realized by the firing of C-fibres in a human central nervous system, but the way they fire in dogs may be physiologically different. Perhaps Martians do not have C-fibres - or, indeed, central nervous systems. A mental state of one kind may have several different sorts of realization, and in principle this realization need not be physical. Suppose you and I are essentially souls or essentially consciousnesses. In that case, being in pain could be realized as an intrinsically spiritual event or an intrinsically mental event, so long as it resulted from a specifiable kind of perceptual input and behavioural output and stood in the requisite causal relations to the rest of our mental states. This is what I meant when I said that functionalism in a sense bypasses the mind-body problem. Being in a mental state is being in a functional state. It logically follows that any kind of being capable of being in that functional state is capable of being in that mental state. It does not matter what that being is made of, so long as he, she or it is capable of realizing that state.

If the thesis that all causes and effects are physical is false, then functionalism is logically consistent with any of the theories of the mind in this book, with the exception of logical behaviourism. Being a functionalist is quite compatible with being a mind-body dualist, an idealist, a materialist, a neutral monist or a phenomenologist. (Notice that all mental states might be physical states, even if not all causes and effects are physical, so functionalism is still compatible with materialism about the person even if the materialist premiss about causes and effects is false.)

The reason why functionalism is logically inconsistent with logical behaviourism is that, as we saw in Chapter 2, according to the logical behaviourist, being in a mental state is being in a behavioural state. Having a mind is either behaving or having a propensity to behave in certain ways. This is quite incompatible with functionalism. Functionalists define the mental not as behaviour but as the cause of behaviour. Logically, if A is the cause
of B and B is the effect of A, then A and B are not identical. Only distinct entities may stand in causal relations.

Functionalism has come to be a theory of the mind by two fairly independent routes. We saw in Chapter 4 that one of the motivations of the mind–brain identity theorists was to overcome one of the deficiencies of logical behaviourism. They thought a logical behaviourist analysis could plausibly be given for one important sub-class of our mental states: the dispositional ones, that is, beliefs, intentions, motives and desires. However, they thought there was not much intuitive plausibility in maintaining that our occurring mental states are only behavioural. Thoughts and emotions, for example, were treated not as behavioural states but as brain states. Now, one mind–brain identity theorist in particular, the Australian philosopher David Armstrong, argued forcefully that the mind is not behaviour but the inner cause of behaviour: the inner physical cause of behaviour. Armstrong is, of course, a materialist, but he argues for a functionalist materialism in his important 1970 paper ‘The nature of mind’. Armstrong agrees with the behaviourists that the concept of mind is logically tied to the concept of behaviour but not by way of identity. The mind is defined as that which brings about behaviour. In Armstrong’s paper, the central moves of functionalism are made: the individuation of mental states through their causal relations and the identification of mental states with functional states.

The other avenue in recent intellectual history issuing in functionalism is thought about artificial intelligence. In particular, the English mathematician and computer scientist Alan Turing wrote an influential paper in the philosophy journal Mind in 1950. This was called ‘Computing machinery and intelligence’. Turing is there concerned to defend the logical possibility of machine intelligence, and part of his strategy in doing so is the devising of a test that has subsequently become known as the ‘Turing Test’. It provides us with a criterion for deciding whether computers are intelligent. The idea is that, first, a man and a woman in one room should provide answers to questions posed by an interrogator in a separate room. The questions are typewritten or sent via a teleprinter to the interrogator, who is to decide in each case whether the question has been answered by the man or the woman. Then, secondly, the woman is replaced by a computer. In this second case the interrogator has to decide whether each question has been answered by the man or the computer. To the extent that the interrogator succeeds in judging correctly more often in the computer case than in the woman case, we should say the computer is intelligent.

Now, whatever the merits or demerits of the Turing Test as a criterion of artificial intelligence, the connection with functionalism is as follows. The states of a computer which is running a program are defined by input and output relations and by their relations to each other. The software description of the operations of a computer is a functionalist description where, if a certain input is received, a certain state results, or where, if the machine is in one state but no input is received, then it moves to a new state, and so on. Importantly, a system may be described functionally in this way without any ontological commitment to what would realize the running of such a program. Software descriptions are functionalist descriptions, and this is the second large impetus to functionalism as a theory of the mind.

In what follows I shall consider the work of two philosophers, each of whom came to functionalism by a different route. They are the contemporary North American philosophers, Hilary Putnam and David Lewis.

PUTNAM

Brain States and Pain States

Putnam begins ‘The nature of mental states’ by arguing that the mind–brain identity theory cannot be dismissed on a priori grounds. He addresses the particular issue as to whether pain is a brain state. (Putnam uses the term ‘brain’ broadly to denote the central nervous system as a whole. He is not here concerned with the partially empirical question: which part of the central nervous system is pain to be identified with? but with the philosophical question, whether it may be significantly identified with any part of it.) His conclusion is that, from the fact that ‘pain’ and ‘brain state’ have different meanings in English, it does not follow that pain is not a brain state. From the fact that the concepts are
The double aspect theory is the theory that mental and physical are two properties of some underlying reality which is intrinsically neither mental nor physical. It follows that it is incompatible with dualism, idealism and materialism because it includes the denial of mental substance and the denial of physical substance. It has it in common with idealism and materialism, however, to be a monism rather than a dualism because it includes the thesis that only one substance, or only one kind of entity, fundamentally exists.

It would be misleading in the extreme to suggest that there is just one kind of theory which could correctly be called 'double aspect theory'. Nevertheless, the three philosophers I have chosen to examine under this heading do share something essential in common. The English twentieth-century philosophers, Bertrand Russell and Peter Strawson, as well as the seventeenth-century Dutch philosopher, Benedictus de Spinoza, all reject dualism, idealism and materialism. Further, each of them in his different way thinks that the mental/physical distinction is not ontologically fundamental, and that our making the mental/physical distinction depends upon either the existence of, or our prior acquaintance with, some substance or entity which is in itself either (a) neither mental nor physical or (b) both mental and physical.

Spinoza thinks there exists only one substance and that this is identical with the totality of what exists. It has two and only two attributes: consciousness and size. In this way Spinoza devises a dualism of properties or characteristics with which to replace the Cartesian dualism of substances. Spinoza is a monist about substance. Russell maintains that our talk about mental and physical events depends logically upon our talk about the contents of sense perceptions he calls 'sense data'. Sense data, according to Russell, are intrinsically neither mental nor physical but count as such only according to the kinds of knowledge that are possible of them. Not only does Russell reject Cartesian dualism, he argues against the existence of any kind of substance: mental or physical. Strawson argues that, unless we were already possessed of the concept of the whole person, we could not have the concept of a consciousness as something distinct from a body. Our making the mental/physical distinction in our thought and language depends logically upon our prior acquaintance with the whole person, the concept of which cuts across any clean mental/physical distinction.

One version of the double aspect theory is sometimes called 'neutral monism'. Russell, for example, calls his theory this. As the term 'monism' implies, this is partly the view that only one kind of entity exists. These entities are purportedly 'neutral' between mental and physical descriptions of them. Russell's sense data, for example, fall into this category. I shall now discuss some varieties of double aspect theories presented in the history of Western philosophy, including 'neutral monism'.

David Hume, the eighteenth-century Scottish philosopher, does not use the expressions 'double aspect theory' or 'neutral monism' to refer to his philosophy of mind. However, he thought that we are directly acquainted through sense perception with a class of entities which are intrinsically neither mental nor physical, and our having knowledge of these is a condition of our having the concepts of mind and matter. Indeed, on Hume's empiricist epistemology, it is impossible to have an idea of anything without having been aware of certain impressions. If we ask what mind and matter are on Hume's account, then a mind is nothing over and above a collection of perceptions. Matter is a mere fiction, postulated to explain the identity of physical objects; and to speak of those physical objects themselves is to speak of the possibility of certain impressions. Minds and physical objects exist for Hume, but they are not Cartesian entities. Hume rejects the view that they are substances and refuses to separate out the questions of what they are and how we can know that they exist. Any talk about them is to be explained in terms of the possibility
of ideas and impressions. Minds and physical objects do possess an identity over time but this is not the identity of substances. The identities of minds and physical objects over time is constituted by relations between the contents of perceptions which give those minds and physical objects their unity and coherence. These relations are thought by the imagination, a faculty which makes empirical knowledge possible.

'Neutral monism' is usually used in connection with the philosophy of William James, the North American philosopher and psychologist who lived and wrote at the turn of the twentieth century. It is true that James's 'pure experience' cannot be straightforwardly classified as either mental or physical, and it is also true that pure experience is what makes possible the classification of things as mental or physical. However, James's ontology of the mind is less clear than the label 'neutral monism' might lead us to suppose. For example, in his celebrated *The Principles of Psychology* he argues that the self has a fourfold structure: the material self, the social self, the spiritual self and the pure ego. The material self is essentially the body but may include physical objects we call our own, our clothes and home for example. The social self is the recognition a person obtains from others. Indeed, a person has as many social selves as there are persons who recognize him or her, according to James. Most interesting for our ontological purposes is James's account of the spiritual self. He says that various views of this are possible. Intuitively or pre-philosophically, it is perhaps best regarded as the awareness or consciousness that a person has of the contents of their perceptions. We obtain the concept of the spiritual self by thinking of ourselves as thinkers. It is a product of reflection. James conjectures that the spiritual self may be nothing over and above the stream of our consciousness, or perhaps some essential or conspicuous component of that stream. There seems to exist a distinction between the active and the passive ingredients of perception. Although I passively receive a sensory input, I seem to be actively aware of such input. Perhaps the spiritual self is this mental activity.

James's speculations about the spiritual self are derived from introspection; the paying attention to one's own mental processes by a kind of direct inner awareness or mental 'watching'. Now, interestingly, James maintains that, if we introspect carefully enough, we will notice that in introspection we never manage to catch an act of spontaneous awareness. All we are really aware of is some bodily sensation, especially in the head and throat. It follows that what normally passes for spiritual activity is in fact physical. What people take to be their innermost self is in fact some bodily movements between the head and the throat. The pure ego is a pure fiction. Philosophers have postulated this because of the shortcomings of previous empiricist theories of mind.

We should not conclude from this account that James is a materialist philosopher. He is not a materialist, a dualist or an idealist. For him, sensations are intrinsically neither mental nor physical, and there exists no mental substance and no physical substance. It is part of his 'radical empiricism' that there is no empirical justification for the postulation of such substances.

Neutral monist or double aspect theories of the mind have one singular advantage and one singular drawback. The advantage is that they avoid the shortcomings of materialism, dualism and idealism. The materialist tends to underestimate the reality of the mental and the idealist tends to underestimate the reality of the physical. Double aspect theorists try to do justice to both, without facing the metaphysical difficulty the mind–body dualist faces of accounting for causal interaction between two qualitatively dissimilar substances. The drawback is that it is typically unclear what the 'neutral' entities postulated by the neutral monist are. It is equally unclear what mental and physical properties are properties of on the double aspect theory. These unclarities are at least as serious as the unclarities about what mind and matter are on all these theories. Perhaps Hume, however, would reply that there is nothing we could know more directly than our own impressions and, if we cannot know what they are, then we are unlikely to know what minds and physical objects are. The neutral monist reverses our normal order of priorities. We normally think we are directly acquainted with our own minds and external physical objects; and there then arise the questions: whether both of these are real; if so, what the relations are between them; whether one may be reduced to the other, and so on. The neutral monist typically postulates the content of our experience – our
impressions of sound, shape and colour for example – as that with which we are most directly acquainted. Minds and physical objects are then described as intellectual or logical constructs out of those contents of experience. The contents themselves are neither mental nor physical.

Whether or not any of these strategies work as a solution to the mind–body problem, some intuitive plausibility may be adduced for neutral monism. It might be that most of the things we come across in everyday life are not clearly either mental or physical, and it might be that some are both mental and physical. Perhaps the classification of things into mental and physical is neither mutually exclusive nor collectively exhaustive. In that case, the mind–body problem could be a product of the philosophical assumption that everything is either mental or physical but not both. Now we may examine the arguments of three philosophers who call this fundamental assumption into question: Spinoza, Russell and Strawson.

SPINOZA

One Substance

Spinoza's central metaphysical idea is that only one substance exists and that this substance may be thought of in two important ways. It is conscious and it has size. He means that the one world which exists – the one we are acquainted with, are a part of, and may think about – has two essential properties or attributes: thought and extension. What exists has these two attributes objectively, that is, whether we believe it or not; and reality may be named in two ways, depending upon whether we are considering it under the attribute of thought or the attribute of extension. If we think of the world as extended, then we should call it 'Nature'. If we think of it as conscious, then we should call it 'God'. 'God' and 'Nature' are two alternative terms denoting one and the same single substance which possesses both mental and physical characteristics.

To appreciate Spinoza's metaphysical theory, we need to make sense of some of the vocabulary he uses. His concept of substance is rather similar to the one Descartes inherited from Aristotle via the medieval scholastics:

I understand 'substance' [substantia] to be that which is in itself and is conceived through itself; I mean that the conception of which does not depend on the conception of another thing from which it must be formed. (Ethics, p. 1)

If something is in itself, then it exists but does not depend upon the existence of anything else for its own existence. If it can be conceived through itself, then, as Spinoza explains, it is possible to obtain the concept of it without thereby needing to think the concept of anything else upon which it might be thought to depend. So a substance depends upon nothing except itself for its own existence and nature. No other thing can cause a substance to be or to be what it is.

Spinoza thinks the only plausible candidate for a substance in this sense is reality as a whole – the totality of what exists. We can suggest, perhaps, why the whole of what is might meet this specification. First, it exists – what is, is. Secondly, what is, is all that there is, so there cannot possibly be anything upon which the totality of existence depends for its own existence. It follows that all that there is depends only upon itself for its own existence, or 'is in itself' as Spinoza puts it.

It follows that the only substance, if it has a cause, is the cause of itself. Given that it exists, and that it is all that there is, there cannot be any cause of its being, nor any cause of its being what it is, except itself. Spinoza explains something's being the cause of itself as follows: 'I understand that to be "cause of itself" [causa sui] whose essence involves existence, and whose nature cannot be conceived except existing' (Ethics, p. 1). Something's essence is what that thing is, so Spinoza is saying that, if something is the cause of itself, then in order to specify adequately what that thing is it has to be mentioned that that thing exists. Its existing is part of what it is. It is reasonable to suppose that the individual existing things we come across depend upon one another for their existence. For example, one event is caused by an earlier event and so on, and this seems equally true of mental and physical events. The Spinozistic thought is that this sort of regress of explanations cannot go on ad infinitum; explanations have to stop at some point. His solution is that the world system as a whole – the one substance – includes its own explanation. Its existence is part of its essence. It is its own reason for being.
conceivable only because it is not true.

So far as disembodied existence is concerned, this is not absolutely or logically impossible on Strawson’s account — although the final two pages of the chapter on persons in Individuals have, I suspect, more than a touch of irony. His point is that it is not difficult to conceive of oneself as disembodied — and I think Strawson is right about that. You have to imagine experiences, just like the ones you are having now, continuing — except there is no experience of your body: your body, for example, does not feature in your visual field. Also, you have no power to initiate changes in the world around you.

However, such a disembodied existence is thinkable only because we are persons: entities to which both mental and physical predicates apply. The conceptual apparatus used to imagine disembodied existence is drawn from, is parasitic upon, its ordinary use as described by Strawson. It is just about intelligible to say that a person who was embodied may become disembodied; but, again, this kind of talk depends upon our being persons in Strawson’s sense. But, according to Strawson, a disembodied being would soon lose any sense of self, being unable to engage with the physical world; and so the sense of ‘person’ in that context would soon be lost; it was, after all, dependent upon the fact that we are persons.

Phenomenology is the attempt to produce presuppositionless descriptions of the contents of experience, without any prior commitment to the objective reality of those contents. This procedure has two goals. It is hoped to exhibit the perennial features of human thought and perception — their ‘essences’ — and it is hoped to ‘ground’ all other kinds of inquiry. Phenomenologists frequently maintain that phenomenology is prior to any other kind of inquiry. It demonstrates the possibility of other inquiries, including philosophy, by showing how all knowledge is made possible by experience. Thus phenomenology may be correctly viewed as an extreme form of empiricism, because empiricism is the doctrine that all knowledge is derived from experience; but it may also be viewed as a kind of Cartesianism, because it tries to place the whole of our knowledge upon secure, indubitable foundations; and it may be viewed as a kind of Kantianism, because it is partly an attempt to show how knowledge is possible. Because of this feature it is sometimes called a kind of ‘transcendental’ philosophy. These three ways of thinking of phenomenology are complementary, not mutually inconsistent, and I think each provides a strong strain in the movement.

It will be apparent that phenomenology is not straightforwardly an ontology of the mind in the sense in which, say, materialism and idealism are. Nor is phenomenology in any interesting sense a ‘theory’; rather, it is a practice. It is the practice of observing and characterizing the contents of experience just as they appear to consciousness, with a view to capturing their essential features. Thus, although phenomenologists usually aspire to a rigour of expression in philosophical language, it is fair to say that doing phenomenology requires an almost aesthetic or artistic ability to contemplate the qualities of one’s experience.
Phenomenology is primarily a German- and French-speaking movement in modern philosophy. Its principal exponents have been the German philosopher and psychologist, Franz Brentano, whose most important work Psychology From an Empirical Standpoint was published in 1874, the German philosopher, Edmund Husserl, whose phenomenological writings are voluminous and include Logical Investigations (1900–1), Ideas (1913) and Cartesian Meditations (1929), and the most profound thinker of the three Germans (and, some would say, of the twentieth century) the philosopher Martin Heidegger. Heidegger was Husserl’s most brilliant student, but his celebrated 1927 work Being and Time exhibits such originality as to constitute the break with Husserl’s phenomenology that Heidegger intended it to be. Heidegger replaced that phenomenology with what he calls ‘fundamental ontology’: a philosophical inquiry into the meaning of being, a study of what it is to be.

The salient French exponents of phenomenology are Jean-Paul Sartre and Maurice Merleau-Ponty. Sartre is as well known as a novelist, playwright and left-wing political polemicist as he is as a philosopher. Merleau-Ponty too was on the Left in politics and shared with Sartre a version of existentialism: a radical and practical brand of philosophizing which gave priority to questions of human existence such as anxiety, relations with others, death and political commitment, over questions of epistemology and metaphysics. In its obsession with action over cognition, existentialism is a partial reaction against phenomenology which incorporates many of the fundamental ontological insights of Being and Time. The two French works which are classics of the phenomenological movement are Sartre’s Being and Nothingness (1943) and Merleau-Ponty’s Phenomenology of Perception (1945). These works are syntheses of phenomenological and existentialist thinking.

The two phenomenologists I have selected for study in this chapter are Brentano and Husserl. I shall concern myself with their work only in so far as it bears on the mind–body problem, but before turning to that I shall say something about the central concepts of Husserl’s phenomenology. When people think of phenomenology, it is paradigmatically Husserl’s brand of it they have in mind.

The most important distinction for understanding Husserl’s phenomenology is that between the so-called natural attitude and transcendental subjectivity. The world of the natural attitude is the everyday, common-sense world which we inhabit when we are not practising phenomenology. It is full of physical objects and other people, and I am one of those people. The world of transcendental subjectivity is that same world, but viewed phenomenologically. It is the world as I directly experience it when I have suspended all belief in the objective reality or causal relations of the objects of my awareness. Husserl calls this suspension of belief the ‘transcendental reduction’ or sometimes the ‘epoché’ (‘epoché’ is the Greek for ‘suspension of belief’). It is important to note that Husserl is not committing himself to a Berkeley-style idealism by the imposition of the epoché. Husserl does not deny that the external world of physical objects and other minds in causal interaction exists; he merely suspends this belief to do phenomenology. We can understand this by making a distinction between disbelieving something and not believing something. For phenomenological purposes, Husserl does not believe in the external world, but he does not disbelieve in the external world either. He leaves the question open.

The result of the epoché is the phenomenological presentation of the ego and the life of consciousness. Rather like Descartes, Husserl endorses the view that I could in principle continue to exist even if there were no external world. Clearly, however, ‘I can no longer denote a thinking, living human being with a body. I too have been transformed by the phenomenological reduction. I exist as a bare, subjective precondition for experience. Phenomenologically I am a transcendental ego.

From the phenomenological standpoint, certain fundamental structures of consciousness are exposed. Perhaps the most important of these is the distinction between a mental act and its content, or the ‘noesis’ and the ‘noema’. The noema is what is perceived, what is remembered, and so on; but the noesis is the actual act of perceiving or the actual act of remembering. This distinction is not revealed within the natural attitude. It is available only to phenomenological reflection. Once the phenomenological reduction is carried out, the essential structures of consciousness may be discerned a priori by such reflection.
In what follows, considerable emphasis is given to the doctrine of intentionality. The intentionality of the mental is its alleged property of being directed towards some object or some content. Thus, all perception is perception of something or other, all hating is hating something, all fearing is fearing something, and similarly for all the other possible kinds of mental state. What is thought, feared or perceived is the intentional object of the mental act. Such intentional objects do not have to exist mind-independently in order for mental acts to exhibit this alleged feature. The idea of intentionality is not original to the phenomenologists; it is to be found in the writings of the medieval scholastics. However, it is of central importance for the way in which Brentano thinks the distinction is drawn between mental phenomena and physical phenomena. It is to Brentano that we should now turn.

**BRENTANO**

The interest of Brentano’s work for the mind–body problem lies largely in his attempt to find a clear line of demarcation between the mental and the physical. His sharpest and most sustained treatment of this problem is in Chapter 1, Book 2, of his work *Psychology From an Empirical Standpoint*, called “The distinction between mental and physical phenomena”. Brentano examines several definitions of ‘mental’ which, in differing degrees, he finds deficient, before deciding upon a solution he finds wholly satisfactory. We may follow the course of Brentano’s arguments in this chapter and so trace the trains of reasoning which lead him to his own solution to the problem.

Brentano begins by noting that we possess an intuitive or pre-philosophical distinction between mental and physical phenomena but that this is not drawn in a precise way. We are in some sense aware of both the mental and the physical: ‘All the data of our consciousness are divided into two great classes—the class of the physical and the class of mental phenomena’ (*Psychology From an Empirical Standpoint*, p. 77). The problem is that the meanings of the two words, ‘mental’ and ‘physical’, are not clear, so we do not possess a strict criterion for distinguishing mental and physical. He describes his project as follows: ‘Our aim is to clarify the meaning of the two terms “physical phenomenon” and “mental phenomenon”’ (*Psychology From an Empirical Standpoint*, p. 78). In what follows we should try to decide whether he succeeds.

**Mental and Physical Phenomena**

Brentano’s first attempt at marking the mental/physical distinction is the amassing of examples of mental and physical phenomena. It might be thought that this is a question-begging procedure because, surely, Brentano must already be in possession of a mental/physical distinction to decide which examples belong to which class. In fact the procedure is less question-begging than it sounds. He has already conceded that we mark the distinction commonsensically, and Brentano may be thought of as making more precise a distinction we already have. In any case, there is a difference between, on the one hand, possessing the ability to distinguish As from Bs and, on the other hand, being able to state explicitly what the difference is between an A and a B. Brentano’s view is that we can, by and large, distinguish mental from physical phenomena, but we are as yet unable to formulate precisely what this distinction consists in.

Brentano provides us with the following examples of mental phenomena: ‘hearing a sound’, ‘seeing a coloured object’, ‘feeling warmth or cold’, ‘the thinking of a general concept’ (*Psychology From an Empirical Standpoint*, p. 79) and he claims that all the ‘states of imagination’ are mental phenomena (ibid.). On top of this:

Every judgement, every recollection, every expectation, every inference, every conviction or opinion is a mental phenomenon. Also to be included under this term is every emotion: joy, sorrow, fear, hope, courage, despair, act of will, intention, astonishment, admiration, contempt etc. (*Psychology From an Empirical Standpoint*, p. 79)

These examples are contrasted by Brentano with the following physical phenomena:

A colour, a figure, a landscape which I see, a chord which I hear, warmth, cold, odour which I sense; as well as similar images which appear in the imagination. (*Psychology From an Empirical Standpoint*, p. 80)