

OUP Copyright

CHAPTER 23

The mind–body problem and mental health, a philosophical update

Chapter contents

- Session 1 **The mind–body problem in clinical neuropsychiatry** 642
- Session 2 **Functionalist accounts of the mind** 644
- Session 3 **Davidson's Anomalous Monism** 649
- Session 4 **Arguments against mind–body identity theories** 654
- Session 5 **Is there any reason to believe in supervenience?** 660
- Reading guide** 664
- References** 665

OUP Copyright

Introduction

The previous chapter provided an historical introduction to the mind–body problem. In this chapter we will return to the issue of the relation of mind and body and examine some of the rival contemporary approaches within the philosophy of mind. The aim of the chapter will be to shed light on the range of different ways that mind and body might be related and the kind of arguments, both based on a priori reasoning and on very general empirical evidence, that can be brought to bear on the issue. It will do this by examining a small sample of views and indicating the range of other choices that might be defended.

This chapter will also provide further introduction to the debates in later chapters. This is because the mind–body problem is one of the central issues in the philosophy of mind and views taken on this question will have ramifications into other debates. (By the same token, views here are influenced by views taken in those other debates.)

We will look at two influential views of the mind: functionalism, which is perhaps the most widely held view of philosophers who do *not* specialize in the philosophy of mind and the late Donald Davidson's 'Anomalous Monism'. We will also look at two general ideas about the relation of mind and body, which crop in different theories or models. The first is a codification of the view that mental states just are physical states. Mental states are identical to physical states: an *identity* thesis. The second is little more complex. Determining the physical properties of a person determines their mental properties but not vice versa. This is called *supervenience*. We will find that while there are a range of different models all face some difficulties. In other words, our intuitions about the relation of mind and body do not form a neat coherent set and accommodating some of them leaves others unsatisfied.

The plan of the chapter

This chapter will begin in *Session 1* by looking at some of the assumptions about the relation of mind and body found in neurology and neurologically informed psychiatry.

Session 2 will then review the advantages for holding a *functionalist* account of mind. This is the orthodox position (especially among philosophers who do not themselves specialize in the mind). It promises to underpin a science of the mind in which mental states are identified as second order physical states of the brain, picked out by their causal or functional roles. But it faces criticisms, the most important of which has to do with the role of rationality in the mind.

Session 3 will examine a more modest position: Donald Davidson's Anomalous Monism. This is modest in the sense that it holds out no promise of reducing mental properties to any other sorts of properties. (Functionalism, recall, identifies mental properties with functional properties, and these are just second order properties of the brain.) Nevertheless, Davidson argues that each individual mental state or event just is a physical state or event, differently described. That is why he calls his position a form of 'monism'.

However, Davidson's position comes under criticism partly because of his advocacy of a further claim about the relationship

of mind and body: that the mind *supervenies* on the body. This is a constraint expressed by the claim that no two people who were physically identical could differ in their mental properties. There are arguments that the holding of both a token identity claim and supervenience is incompatible.

Session 4 examines two arguments against identity theories of mind and body. The first argues that if an identity theory were true it would have to be necessarily (rather than merely contingently) true. However, we have reason to believe that mental states might not have been physical and thus there can be no identity. The second argument focuses not on the nature of the relation but on what is related. It argues that even an identity claim as weak as a token-identity thesis (of which more below) is implausible.

Session 5 examines the arguments for the truth of supervenience: the claim that the physical properties of people or the world as a whole determine the mental properties. Even this modest claim turns out to be more problematic than it might at first appear.

Session 1 The mind–body problem in clinical neuropsychiatry

In the philosophy of mind, the relation of mind and body, the mental and the physical, is a contentious metaphysical issue. As we began to see in Chapter 22, some philosophers have argued that types of mental state are really types of physical state while others have argued that this cannot be so. In psychiatry, however, the connection between the mental and physical is not merely of academic interest. We begin with a reading that outlines the practical possibilities and current limitations in brain imaging techniques.

EXERCISE 1

(30 minutes)

Read the extract from:

Posner, M.I. (1993). Seeing the mind. *Science*, 262: 673–674 (Extract p 673).

Link with Reading 23.1

In this exercise, we want you to start to think about the significance of the recent growth in the technology for *brain* imaging for making discoveries about the *mind*. Given the title of the reading, to what extent do the experimental techniques described reveal the nature of the mind? Is there an obvious *empirical* answer to this question? Think about the significance of the finding that local areas of brain activity often seem to be highly correlated with kinds of mental activity (such as speaking to oneself). Does this show that one is literally seeing the mind in a brain scan? If not, why not? Think also whether the relation between mind and brain could have been different. Could God have so designed the universe (so to speak) that there was a different relation?

Mental and physical correlations

Although there is some disagreement about the assumptions that underpin the interpretation of imaging techniques and expressing some scepticism about their validity, Posner presents a rosy picture. Take the following passage:

It is a popularly held belief in psychology that the cognitive functions of the brain are widely distributed among different brain areas... Nevertheless, imaging studies reveal a startling degree of region-specific activity... When thought is analysed in terms of component mental operations, a beautiful localisation emerges. In word reading studies, words activate specific posterior visual areas that are not affected by consonant strings, and specific frontal and temporo-parietal areas are active when subjects are required to indicate the use of a noun... or its classification into a category. (p. 673)

On the assumption that such localization of brain activity correlated to specific mental activities is a genuine phenomenon of brain imaging what does this show about the relation of mind and brain?

It is interesting to note that that despite calling his short paper 'Seeing the mind', Posner does not go on to discuss at any length *how* the mind is seen. The implication, suggested in the passage above, is that brain activity can be correlated with mental activity (this is simply an empirical finding, according to Posner) and furthermore, imaging that brain activity just is imaging the mental activity. This second assumption deserves careful scrutiny. Is it clear that mental and physical correlations imply or justify in a weaker sense mental and physical identities?

Something like this latter assumption also underpins the very idea of cognitive neuropsychiatry. In the case studies described by Halligan and Marshall (1996), a clear, although implicit, connection between the mind and the body or brain is made simply through the juxtaposition of descriptions of patients' states of mind and behaviour and accounts of structural features of their brains. This implicit connection is a reflection of the key methodological assumption behind 'neuropsychiatry': structural accounts of the brain can shed light on psychiatric symptoms and vice versa.

But on the other hand, just what the nature of the connection is between neurology and psychiatry remains obscure. As the Introduction to *Method in Madness* notes

We have no doubt that a neuropsychiatry is possible, despite the fact that many (or perhaps most) psychiatric conditions remain as yet 'functional'... [But] the disciplines of neurology and psychiatry seem to be investigating 'similar disorders viewed with different perspectives'. It will be a major challenge... to reconcile these perspectives... How much of a gap will always separate the two disciplines is debatable. Likewise, the extent to which the gap can be bridged by in vivo functional brain imaging is likely to remain highly controversial for some time to come. Halligan and Marshall (1996, p. 7)

Thus a discipline such as neuropsychiatry provides both the most concrete empirical evidence for the connection between mind and brain while at the same time demonstrating the still inconclusive nature of that evidence.

The element of cognitivism in the identity

It is also worth noting here that the connection speculated upon in the reading (Posner, 1993) is between *cognitive* psychology or psychiatry and neurophysiology. Posner says 'When thought is analysed in terms of component mental operations...' suggesting the same sort of cognitive analysis of thought processes.

So a first stage in the proposed connection will be the analysis of mental phenomena or states in cognitivist or information processing terms. This is a substantial assumption. Can the mental be described satisfactorily in these terms? An abstract form of cognitivism is a functionalist account of the mind. As we will see, however, in this chapter and Chapter 24, there are profound philosophical objections that functionalism would have to overcome for this assumption to prove correct.

One response to this marrying of bizarre psychological symptoms with underlying brain structures is to say—as we have already seen—that psychological and neurological descriptions are *different perspectives* on the *same phenomena*. There are not two sorts of phenomena—the mental and the physical—but one sort, which can be described in two different ways. This is not to say that the kinds of descriptions so far available can make this relation transparent. The connections between structural and functional descriptions of the brain and between those and functional or cognitivist descriptions of the mind cannot yet be filled out. But the underlying goal of a cognitive neuropsychiatry is precisely to articulate those connections. Furthermore, the kind of materialism expressed above may seem to be the best underpinning for such a project.

Such a materialist might seem to suggest that in, say, functional magnetic resonance imaging scans we see not only the working of the brain but also the working of the mind. It may seem that such a view is almost obligatory for any serious neurologically informed psychiatry because to doubt it would be to subscribe to some form of dualism of mind and matter. But as attention to the philosophical and metaphysical debate shows that is not a true opposition. Options in the mind–body debate do not comprise a binary opposition of a romantic dualism versus a serious minded scientific monism or materialism but comprise a range of different options that differ in the tightness of the connection of mental and physical properties. This chapter will steer a course through some of the recent debate in order to shed light on the range of available positions.

The next session will turn to functionalist accounts of mind. As well as being hinted at in the reading linked with Exercise 1 (Posner, 1993), these are also still the most popular accounts in the philosophy of mind.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then, thinking back over the chapter as a whole, write brief notes about the following:

1. How does empirical work contribute to our understanding of the relation of mind and brain?
2. What argument is given in the reading for the connection between mind and brain?
3. What suggestion is made about the connection between a cognitivist approach to the mind and the brain?

Session 2 Functionalist accounts of the mind

Introducing functionalism

This session introduces the current orthodoxy in the philosophy of mind: functionalism. It also considers some of the criticisms it faces. One of these criticisms is that functionalism cannot accommodate the essential connection between the mental and rationality. Such a connection will play an important role throughout this module while also introducing significant difficulties for the philosophy of psychiatry including, for example, an account for delusions.

The inspiration for functionalism is the computer. The basic idea is that mental states are software states, i.e. patterns of information-processing, whereas brain states are causal states of a physical substrate.

The attractiveness of this model is obvious. Unlike its main historical rival, behaviourism, it allows for a person's internal states to differ, although the outward behaviour may remain the same. Relatedly, it accounts for the fact that not all desires lead to action. Acting on one desire may be incompatible with satisfying another, perhaps stronger, desire. It also provides for internal states to drive or cause behaviour in some way. Unfortunately, though, as we will see, all these *prima facie* advantages have proved extremely difficult to pin down in a systematically consistent way.

Types and tokens

Functionalism can thus be seen as a way to avoid some of the objections to behaviourism. It also escapes an objection to crude forms of physicalism. Before summarizing its advantages over those positions it is worth revisiting the distinction between types and tokens. Talk of a *type* of mental state is a way of grouping together different individual or *token* mental states (at different times or in different people), which have a property in common.

Thus everyone who thinks that the prime minister is doing a good job has something in common. They all share the same *type* of belief. Everyone who has an ache in their left knee shares the same *type* of pain. But each person's pain is their own. One cannot, for example, give someone else one's pain in the knee (although one can cause them to have a pain of their own of the same sort). Each person has a different *token* state of the same general type. The same goes for beliefs.

A closely related distinction is between *qualitative* and *numerical* identity. Two ties can be qualitatively identical if they look exactly alike, are made of the same material and so on. But they cannot be numerically identical unless 'they' are the very same tie. Now it may seem that numerical identity is an idea of limited use as it seems to require both that there are two things (to stand in the *relation* of identity) and that really there is only one (as the relation is one of *identity*). But in fact there are often two different names or descriptions of the same thing whose coincidence comes as a discovery. In many cases, qualitative identity corresponds to sameness of type while numerical identity is sameness of token. (But this neat assimilation is complicated in the case of the putative identity of mental and physical *properties*: such as the claim that pain just is the firing of C-fibres. This is a claim about the numerical identity of properties: i.e. *types* of mental and physical state.)

With the distinction of type and token in place we can now summarize a key advantage that functionalism has over both behaviourism and physicalism.

Behaviourism as a type identity theory

Type identity versions of behaviourism identify types of mental state with types of behavioural disposition. Even Ryle's modest version of behaviourism described in Chapter 22 seems to require such a connection. However, as a mental state only results in action in the context of other mental states, a one to one correlation between mental states and dispositions to act is implausible. The desire for a drink of water may not lead to drinking if it is combined with the belief that the only available sources of water contain poison. But such a failure to drink does not imply that the desire for water did not exist after all, as a crude behaviourism would imply. Functionalism avoids this problem because it claims that mental states mediate between perceptions and actions as a network rather than individually. It also differs from behaviourism in that mental states are not merely disguised descriptions of behaviour but are internal states playing a genuinely causal role in the production of behaviour.

Type identity physicalism

Type identity physicalism identifies types of mental state with types of physical state. But this identification is implausibly *chauvinist* as it seems possible that creatures with very different physiology might possess the same mental states. Functionalism escapes this charge because it claims that what matters for the identity of a type of belief is a second order property rather than

the first order properties of physical states. Nevertheless, it is still consistent with the more modest physicalist claim that each token (of a type of) mental state is a physical state.

Functionalism

Functionalism is the view that what make a mental state the state it is—whether a pain in the knee or a belief about frogs—are the causal relations holding between it, other mental states, perceptions, and behaviour. Thus each mental state is constituted by its functional relations to the rest of one's mental economy and the inputs and outputs to that economy.

What is the connection between functionalism and physicalism?

By concentrating on these causal or functional interrelations, functionalism *need not* take a view on what underpins these causal or functional relations. But in fact, some functionalists do commit themselves to materialism or physicalism. Ned Block (1980) makes a useful distinction between 'functional state identity theory' and 'functional specification' in his 'What is functionalism'. The former identifies pain, for example, with a functional state and inquires no further as to its underlying physical make up or ontology. Block ascribes this view to Putnam and Fodor. On this account pain just is the role or property. Such a view is consistent with a dualism of mental and physical substances. In other words, it provides no support for physicalism.

By contrast, functional specification uses functionalism to specify a role that is filled by some underlying and probably physical state with which mental states can then be identified. This is the view of Lewis and Armstrong. Pain, for example, is said by them to be the first order physical state or thing which has the second order role or property identified by functionalism. (This contrasts with saying simply that pain just is that second order role or property.) So construed, functionalism specifies a description that is satisfied by some physical state. Hence the philosopher David Lewis, for example, argues for an identity or for bridging laws between the mental and physical on the basis of functionalism. He claims, in other words, that, rather than conflicting with it, functionalism offers support for a kind of type-identity physicalism.

Lewis' argument goes like this:

1. Mental state M = occupant of causal role R (by definition of M)
2. Neural state N = occupant of causal role R (by physiological theory)
3. Therefore, mental state M = neural state N (by transitivity of identity)

Lewis claims that we can deduce identities rather than simply postulating them for ease or simplicity. (Unlike Saul Kripke—see later in this chapter—however, he does not go so far as to ascribe necessity to them.)

What justifies Lewis' extra claim that functionalism supports type-type physicalism? Consider the status of a putative identity of a type of mental state and a type of physical state in the face of possible discoveries about other creatures with very different physiologies. Lewis argues that such identities may be species specific. Thus some type of physical state will occupy a functional or causal role for a given population and will thus just be pain, say, for that population.

Mad pain

Interestingly, this suggests that if an instance of that type of physical state is realized in a person under abnormal circumstances and playing the 'wrong' causal or functional role, it will still be pain. Lewis calls this the possibility of 'mad pain' and imagines a person for whom characteristic pain behaviour was the solving of complex mathematical problems. By contrast, a behaviourist would have to say that such a person was not in pain.

Martian pain

And, in Lewis's picture Martian pain is also possible. This comprises a different type of physical state but one that plays the same causal role for Martians as our pain states do for us. Thus a variety of physicalism based on functional specification does not rule out the idea that creatures with very different physiologies might still share types of mental state with humans. It is not so chauvinistic.

It is not a fully physicalistic identity theory because what all pains have in common is not, however, any physical property. What makes Martian pain the same kind of state as our pain is not answered in physical terms. Hence Lewis's account is not in fact a victory for type-identity physicalism. In fact functionalism gives grounds for believing type-identity physicalism is false—we need not postulate sameness of physical state given that we can ascribe mentality on the basis of functional states. Given the possibility of the multiple realization of mental states, advocating type-identity physicalism is a needless epistemological risk. Thus functionalism looks at first sight an attractive account of the metaphysics of mind.

Can functionalism capture the experiential quality of mental states?

Two sorts of criticisms of functionalism

Functionalism has, however, faced a number of substantial criticisms. In 'Troubles with functionalism' Ned Block fires a broadside. His strategy has two main strands.

1. To argue that functionalism is fundamentally flawed as a model of mind as it privileges the quantitative aspects of mental states and processes (i.e. their measurable aspect) while ignoring or denying the qualitative aspects of mental states (e.g. belief-states, emotions).
2. To demonstrate that functionalism, in its various forms, is either too liberal (inclusive) or too chauvinistic (exclusive)—i.e. the functionalist project must include some systems without minds or exclude some with minds.

One criticism is that functionalism necessarily took no account of the *qualitative* aspects of some mental states. As well as mediating between certain sorts of causal inputs and behavioural outputs, pain, for example, has a characteristic feel. Similarly, it is often claimed that direct colour perception not only provides information about the ripeness of tomatoes but also involves a qualitative experience or *qualia*. At first sight it seems that functionalism lacks the resources to articulate this aspect of our mental lives.

Block considers connecting the population of China together with two-way radios to mimic the functional organization of a human mind. They can then be connected to a robot body. In the thought experiment a single mind is staffed by a number of other minds and thus Block calls it a homunculi-headed system. But, he suggests, although this fits a functionalist criterion for a mind, it is not one.

What makes the homunculi-headed system . . . just described a *prima facie* counterexample to machine functionalism is that there is a *prima facie* doubt whether it has any mental states at all—especially whether it has what philosophers have variously called ‘qualitative states’, ‘raw feels’, or ‘immediate phenomenal qualities’. (You ask: What is it that philosophers have called qualitative states? I answer, only half in jest: As Louis Armstrong said when asked what jazz is, ‘If you got to ask, you aint never gonna get to know’). In Nagel’s terms, there is a *prima facie* doubt whether there is anything which it is like to be the homunculi-headed system. (p. 278)

In fact it is far from clear that this is a decisive objection. All turns on the kind of functional account of consciousness which may become available. (A defence of functionalism to this sort of objection is made in Van Gulick, 1994.)

A second objection to functionalism raised by Block is that it faced a fatal dilemma in specifying its inputs and outputs. The charge is that functionalism is guilty of either chauvinism or liberalism. It is chauvinistic if the inputs and outputs are so described that genuinely mental systems are wrongly ruled out. This is a danger if inputs and outputs are defined by analogy with the human case. This might be done in two different ways. One might define the outputs of a functional characterization of the mind as the movements of arms or legs. But this would imply that multiple amputees could not have minds. Or one might retreat to the outputs of the brain and describe the firings of the terminal neurones that normally cause bodily movement. But this would still be chauvinistic in that it would require that only those entities with human brains could have minds.

A radical alternative is not to describe the outputs to a functional system in physical terms at all. One might instead map out the system of functional interactions and then define the outputs in the most abstract and general terms: simply by numbering them, for example. But the danger now is one of liberalism. Such a formally described system could be instantiated by ‘obviously’ mindless systems such as economies. (One could equate output 1 with the inflation rate, and output 2 with the balance of trade, for example.)

Economic systems have inputs and outputs, e.g. influx and outflux of credits and debits. And economic systems also have a rich variety of internal states, e.g. having a rate of increase of GNP equal to double the Prime Rate. It does not seem impossible that a wealthy sheik could gain control of the economy of a small country, e.g. Bolivia, and manipulate its financial system to make it functionally equivalent to a person, e.g. himself. If this seems implausible, remember that the economic states, inputs, and outputs designated by the sheik to correspond to his mental state, inputs, and outputs, need not be ‘natural’ economic magnitudes . . . The mapping from psychological magnitudes to economic magnitudes could be as bizarre as the sheik requires. (p. 225)

This is a powerful dilemma. But it takes the form of a challenge to functionalism rather than a proof that a functionalist account of mind is impossible. The challenge of liberalism relies on the intuition that an abstract functional system equivalent to a human mind could be instantiated on a mindless system. A functionalist may reply that any such intuition relies on an oversimplified picture of what such a functional system will be like. A worked-out functional account of the mind will possess a complexity that precludes its instantiation by an economic system. If the objection is to prove decisive something more has to be added to it.

Syntax and semantics and Searle’s Chinese room

One such addition stems from the fact that a functional system can be described in merely syntactic terms rather than semantic terms. (Syntax is the set of rules that govern the correct use and combination of symbols regardless of the meaning: the grammar, as it were. Semantics adds to those connections between words or sentences and parts of the world.) Thus describing the inputs and outputs to a functional system in purely formal terms does nothing to ensure that functional states possess intentionality or aboutness.

This is the heart of the philosopher John Searle’s (1984) famous ‘Chinese room’ thought experiment. Searle likened artificial intelligence to a man in a closed room with a kind of Chinese phrase book but containing only questions and answers in Chinese. Phrases posted into the room could be paired with the same phrase in the book and a corresponding answer posted out. However, in the thought experiment, the man himself does not understand Chinese even though the box as a whole appears to. Searle suggests that the same applies to any artificial intelligence system brutally following a computer program. He called artificial intelligence simulated rather than real intelligence. (This thought experiment and Block’s China thought experiment ask related questions. Searle asks: where is the intelligence, the meaning? Block asks: where are the qualia, the experiences, the point of view?)

Of course a functionalist may reply to this by saying that if mental states are functional states of the brain then they also *could* be described in merely syntactic terms but that does not show that they *cannot* also be described in semantic terms. Still, the challenge to functionalism is to show what grounds a semantic interpretation of its states.

Causal theories of reference

The standard contemporary answer is to add to a broadly functionalist account of the mind, a causal theory of reference. We will return to this idea in Chapter 24. But the basic idea is that internal functional states gain their semantic status by standing in causal relations to those things in the world that they are about. The causal theory of reference is problematic. But it provides a defence for this challenge to functionalism. The reason why economic systems lack intentionality is that they do not stand in the right additional causal relations to things in the world.

Block's dilemma and the reductionist hopes of functionalism

In fact, what makes functionalism vulnerable to Block's dilemma is that it attempts to characterize the mental in non-mental or reductionist terms. Reductionist forms of *behaviourism* attempt to define mental states in terms of behavioural dispositions. This presupposes that each state can be paired with a behavioural disposition. *Functionalism*, however, need not presuppose this because it recognizes that mental states only lead to action in the context of other mental states. One may be thirsty but this will not dispose one to drink fluid from a beaker if one also believes that it is poisoned. Thus mental states can only be defined in terms of perceptual inputs and behavioural outputs holistically. 'Ramsey sentences' provide a formal mechanism for defining the mental in non-mental terms but only as a whole.

The underlying motivation for this project is to shed light on the mental by defining it in terms of something supposedly less mysterious. For this reason, both inputs and outputs are described in non-mental terms. One does not describe an output as, for example, 'reaching one's hand out for the coffee cup' because the verb 'reaching' is purposive and implies possession of a mind. Instead this action has to be described in non-mental, non-intentional terms; perhaps as a movement of a five-pouched manifold in an easterly direction at one metre per second. But the problem in realizing this reductionist project is that it may not be possible to define the mental in physical terms even at this holistic level. So far, however, we have only had Block's powerful intuition that that is so and not an argument that it must be the case.

Davidson's attack on reductionism

The American philosopher of language and mind Donald Davidson (1984) has put forward just such an argument. He claims that there is a principled reason why one cannot define the mental in non-mental terms. The mental is irreducible. His argument centres on the idea that rationality is a constitutive principle that necessarily governs the ascription of mental states. In making sense of each others' beliefs and actions we have to assume that we are largely rational. Thus, Davidson claims that the mental domain is governed by a constitutive principle of rationality while the physical is not. As a result, he claims, there can be no law-like relations between the two realms. But before moving on to Davidson's proposed solution to the mind-body

problem developed as a response to the claimed central role of rationality, it is worth noting a different response.

Can functionalism capture the rational character of mental states?

Non-reductionist functionalism?

Block claims, effectively, that functionalism cannot spell out inputs and outputs to a mental system both in non-mental terms and also in a way that avoids both chauvinism and liberalism. Thus one response might be to preserve something of functionalism by dropping the requirement of reductionism and allowing the inputs and outputs to be described in intentional or mind-involving terms. This would no longer shed light on the mental from the 'outside'. But it would at least provide an account of the metaphysics of mind. Such a position would stand to functionalism proper as non-reductive behaviourism of Gilbert Ryle stands to the behaviourism advocated by the psychologist B. F. Skinner (and some but not all of the writing of the philosopher W.V.O. Quine). It would, however, retain a key element of functionalism in construing mental states as internal functional states standing in nomic causal relations. As the following discussion shows, however, there is still something fundamentally wrong with this position.

McDowell's Davidsonian argument

The reading (linked with Exercise 2) is by the contemporary British philosopher John McDowell (1985). Examples of his work formed readings in Part 3. A running theme in those readings was the idea that the 'space of reasons' is different from the 'realm of law'. This had consequences for, among other things, the connection between observation and theory (Chapter 12) and the relation of the human and the natural sciences (Chapter 15). In this reading some of its consequences are explored in the context of the philosophy of mind.

EXERCISE 2

(30 minutes)

Read the extracts from:

McDowell, J. (1985). *Functionalism and Anomalous Monism*. In *Actions and Events* (ed. E. Lepore and B. McLaughlin). Oxford: Blackwell, pp. 387–398. Reprinted in McDowell, J. (1998) *Mind Value and Reality* Cambridge, M.A.: Harvard University Press, pp. 325–340 from which we have taken the extracts (Extracts pp. 328–331, 336–337)

Reading 23.2

This article is a criticism of a defence of functionalism by B. Loar in *Mind and Meaning*. But the key points can be understood independently of that book.

- ◆ What role does rationality play in McDowell's argument?
- ◆ What is the key difference between explanations in 'folk psychology' and natural science?

The key role of rationality

McDowell's (1985) attack on functionalism develops an argument from Donald Davidson (on whom more shortly) based on the role of rationality for the mind. Davidson focuses mainly on the reduction of the mental to the *physical*. McDowell is concerned with the reduction of the mental to the *functional*. But both agree that what prevents reduction is the constitutive role that rationality plays for the mind. McDowell argues in support of this by criticizing one particular functionalist account proposed by Brian Loar. He picks on Loar because Loar explicitly attempts to reduce at least some elements of rationality to a functionalist story. If Loar's project were successful then even though there is a close connection between meaning, mental content, and rationality this connection would not prevent a functionalist account. McDowell, however, argues that functionalism cannot after all account for the central role of rationality.

One of McDowell's arguments is quite straight forward. If functionalism is to capture the constitutive role that rationality plays for the mental it must capture the deductive links between an agent's beliefs. Loar claims that the deductive links between beliefs can be reduced to a formal structure or pattern. But as McDowell argues, one cannot in general reproduce these links without using an intentional language and describing the beliefs in terms of their contents. We know this because there is no general mechanical test for logical validity.

There is a mechanical test for validity of the simplest logical system: propositional logic. This is the logic that governs whole sentences and the logical connectives AND, OR, NOT, IF... THEN, etc. Valid arguments in propositional logic can be checked using truth tables in a way that can be mechanized. But there is no general mechanical method of determining whether arguments in predicate logic are valid. And there are arguments, based on Godel's theorem to show that this is not merely a matter of our ignorance. It is a necessary truth that there are no general mechanical tests for validity.

Furthermore, accounting for deductive logic is the easiest task for functionalism because it is the narrowest conception of rationality. But in claiming that mental states must necessarily stand in rational relations, Davidson includes the claim that the rationality of beliefs should also include induction and perception-based beliefs. There seems to be no prospect of providing a formal schema that abstracts away from the *content* of beliefs and prescribes when it is justified to form inductions or when one can form beliefs on the basis of direct perceptions. This is not to say that no such practical advice could be set down; however, it would turn on the *particular subject matter* of the inductions or perceptions.

Explanation by ideals and by subsumption under universals

McDowell (1985) argues that mental states stand in a different sort of structure to the nomological or law-like structure deployed by a causal functionalism. This is why the one structure cannot be mapped on to the other. McDowell goes on to provide

a second argument against any such mapping, which has to do with the different kinds of explanation suitable for the different structures. It thus concerns the role of psychological explanations. McDowell argues that rationality also functions as a constitutive concept in explanation. Here the idea is that everyday folk psychological ways of making sense of each other are a species of *ideal* explanation. By contrast, explanation by subsumption under laws has a different 'logic'.

Explanation of speech and action works by fitting behaviour into a normative pattern of how rational subjects *ought* to act, given their beliefs and desires. One explains beliefs by citing what they ought to believe given other beliefs and perceptions. This use of 'ought' reflects the fact that psychological explanation turns on providing *good* reasons, reasons that one should have, which rationality ideally requires. Even mistaken beliefs are explained by providing a context of good reasons. This context is an a priori rational structure imposed on all folk psychological explanation.

This style of explanation by ideal contrasts with scientific or nomic explanation. The latter typically subsumes particular cases to be explained under a posteriori laws or generalizations as the covering law model attempts to formalize (Chapter 14). Unlike explanation by reference to ideal functioning, explanation by subsumption under generalizations is undermined by exceptions to the generalizations (unless they are merely statistical). But this is the way functionalism has to construe psychological explanation because of the law-like relations that functional states stand in. So functionalism cannot account for the ideal nature of psychological explanation.

Another way of making this claim is to point out that a functional specification of mental states is an empirical theory that charts their law-like structure. From this, however, it is impossible to milk a notion of having *good* reasons for beliefs or actions. Nor can it provide the resources for explaining how the relations that govern mental states mark the limits of intelligibility of other people. The relations between mental states that we employ in folk psychology do not comprise a merely contingent structure. If that were the case we could conceive of other relations holding. But in fact any widespread and significant deviation from the rational structure would be incomprehensible. (Quite how much divergence from rational belief formation and action is clearly an important matter for psychiatry.)

Functionalism is a form of Cartesianism!

McDowell (1985) provides an instructive diagnosis of what is wrong with a functionalist approach to the mind. He suggests, to begin with, that a broadly Cartesian account of mind results from thinking of reality as an objective realm while at the same time thinking of the mind as both real and subjective. Cartesianism attempts a quick ontological solution to this tension by construing the mind as real but immaterial. Functionalism is in no sense a dualist theory of mind but, like Cartesianism, it also tries to make the mental objective. It does this by suggesting that mental states are second order causal states of the body.

By contrast, recognizing that rationality plays a constitutive role makes the mental subjective for two reasons. First, it marks off mental states as both states *of* a subject and also *about* the world. And it makes the limits of intelligibility and mind forever open to further criticism and change. In short, McDowell's account of Davidson makes propositional attitudes as irreducible as qualia and as dependent on subject's points of view. (As mentioned above: functionalism has been criticized for being unable to account for the qualitative aspect of some mental state; however, most philosophers assumed that it was at least adequate to account for the intentional properties of mental states.) These dense comments will become clearer by examining Davidson's account of the mind: Anomalous Monism. Ironically it will turn out that McDowell's criticisms of functionalism also apply to Davidson's own account.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What, in a nutshell, is a functionalist view of the mind. On what analogy is it based?
2. Does functionalism support or undermine a broadly physical or materialist view of mind? What is the relation of type and token versions of identity claims?
3. What is its relation to behaviourism?
4. What criticisms does it face? (We name two.)

Session 3 Davidson's Anomalous Monism

Functionalism is one account or model of how the mind might be related to the brain or body and thus serves as a potentially appropriate model for psychiatric research into the neurological underpinnings of mental illness. But, as we have seen, it faces considerable a priori criticism. The next reading (Davidson, 1980) introduces a different model: Donald Davidson's (1917–) more recent account of the metaphysics of mind.

Davidson was a recent American philosopher much influenced by W.V.O. Quine. His work can be divided very roughly into the philosophy of mind and the philosophy of language. He never wrote a book and his work comprises a series of articles in journals published since the 1960s many of which have been gathered into two collections *Essays on Actions and Events* and *Inquiries into Truth and Interpretation*. This reading comes from the collection that focuses on the philosophy of mind. (A clear introduction to Davidson's philosophy is in the Reading guide at the end of this chapter.)

What should be striking about Davidson's account is that it is a move to a more modest account. Whereas type identity physicalism or functionalism promise to explain mental properties as either physical properties or functional properties (and the latter are usually thought of as second order *physical* properties), Davidson assumes that mental properties cannot be reduced in this way. Types of mental state are not identical with any type of physical or functional state. But each *token* mental state is a (token) physical state.

Davidson thus presents, through a priori reasoning, an alternative to both a simple-minded assumption that mental properties are just physical properties seen from a different perspective and also to the more sophisticated view that mental properties are functional properties. Thus the issue for this session is whether Davidson's abstract conception of the relation of mind and body can provide a coherent model for clinical and research work in mental health.

EXERCISE 3

(30 minutes)

Read the extracts from:

Davidson, D. (1980). Mental events. In *Essays on Actions and Events*. Oxford: Oxford University Press, pp. 207–227 (Reprinted in N. Block (ed.), *Readings in Philosophy of Psychology*. Oxford: Oxford University Press, and in D. Rosenthal (ed.), (1991) *The Nature of Mind*. Oxford: Oxford University Press, pp. 247–256 from which we have taken the extracts (Extracts: pp. 247–248, 248–249, 250))

Link with Reading 23.3

- ◆ What are the three premisses or principles Davidson aims to reconcile?
- ◆ Would Davidson's account allow for neuropsychiatry?

The challenge

Davidson's account of mental states, Anomalous Monism, aims to reconcile three principles or assumptions:

1. that mental states stand in causal relations;
2. an underlying nomological account of causation;
3. the anomalism of mental; i.e. there are neither laws linking mental and physical or mental and mental properties.

At first sight, these assumptions appear to be in tension. The first premiss is that mental states stand in causal relations. The second is that whenever there is causality, there are laws. But the third says that the mental realm is not governed by laws. Before summarizing Davidson's proposed reconciliation it is worth asking why he holds these principles. Given their at least apparent mutual tension, one response would be simply to reject one of the premisses.

Davidson's three premisses

Davidson's reasons for claiming that mental states stand in causal relations were first set out in his article 'Actions, reasons and

causes' and are the subject of Chapter 26. Briefly his central argument is as follows. When we act for a reason, the reason *rationalizes* our action. It sheds light on why we acted by showing how the action is justified in light of our beliefs and desires. But it is also possible to hold a reason for acting, and act, but not act because of that reason but instead for some other reason. The question is, what is the difference between having a reason and acting (where 'and' is construed purely conjunctively) and acting because of the reason. As in both cases the reason would rationalize the action the difference between them must be something else. Davidson suggests that the only plausible candidate is that the reason why one acts is also the cause of the action.

The second ingredient of Anomalous Monism is a nomological account of causation. As we saw in Chapter 15 it is a widely held consequence of Hume's investigation that singular instances of causation require the existence of general laws. Davidson stresses that he is not assuming that saying that one thing causes another *means* that a general law-like relation exists. Nevertheless such a connection must exist.

The third ingredient is the claim that the mental is anomalous, i.e. it is not governed by laws that link mental properties. This involves a denial both of psychological laws (including those linking beliefs with other beliefs or with actions described in mental terms) and of psychophysical laws (linking types of mental states and types of underlying physical states—e.g. type-identity physicalism). One element of Davidson's motivation is his earlier work on testing and consequent criticisms of a Decision Theory approach (see his paper 'Psychology as philosophy' also in *Essays on Actions and Events*). But the justification for claiming that the mental is anomalous turns on the role of rationality in the ascription of mental states to which we will return in Chapter 25.

Davidson's token-identity theory

The solution that Davidson proposes to the problem of reconciling the three principles is to postulate a token-identity theory. Every mental event is identical to a physical event. Recall that in this context, the notion of identity is that of numerical rather than qualitative identity. (Two ties in a shop may be qualitatively identical while not being the very same tie. On the other hand I may wear the very same tie every Friday.) Here the claim is that each mental event just is the very same thing as a physical event. One way of thinking of this is to think of one sort of event having two different descriptions.

What makes it merely a token-identity theory is the claim that there is no alignment of types of mental event and physical event. If two people have the same type of mental state—perhaps they all share a deep respect for the Queen—then they will have two different tokens of that mental state type because one will be one person's respect and other will be the other's. In virtue of the token-identity theory these mental state tokens will be the very same things as physical state tokens. But because the mental is anomalous, there is no reason to believe that the physical state

tokens will have any physical properties in common. Types of mental state do not align with types of physical state.

Notice what a strong claim this is. It implies that there will be no physical laws linking types of psychological symptoms—for example, the phenomenon of thought insertion—with types of physical state, such as a type of brain lesion. Think of an analogy. It is like denying for a priori reasons that there might be a law-like relation between colours and wavelengths of light.

Anomalous Monism is *non-reductive* materialism

The reconciliation of the three premisses for Anomalous Monism can now be quickly summarized. Mental states stand in causal relations. Therefore they are governed by general laws. There are no general laws that govern mental properties or mental types. But because each mental event is identical to a physical event it can be governed by a general law that connects its physical properties to other physical properties. Thus reasons can be causes even though the mental cannot be *reduced* to the physical. It cannot be reduced because mental properties cannot be

Box 23.1 Davidson on events

Davidson himself is careful to talk of a token-identity theory of mental and physical *events*. This concentration on events is partly the result of an investigation in the philosophy of language. It follows from the fact there is some difficulty in codifying deductions that involve adverbs. Adverbs form no part of first order predicate logic, the logic devised by Frege and which is best understood. From the statement that Jones buttered the toast carefully we can all infer that Jones buttered the toast. The difficulty is explaining why this is a valid inference. Davidson realized that this can be explained as a case of conjunction elimination in first order logic if these statements are interpreted as making claims about events. Thus: there was an event and the event was a buttering of toast and the event was by Jones and the event was careful. From this statement, the last conjunct can be eliminated in a deduction of the same form as if A and B are both true then A is true. Davidson concludes from this analysis that the world order comprises a succession of events. This is rather a startling claim. Davidson is drawing a conclusion about *ontology* or *what there is in the world* from an examination of which inferences can be codified in first order logic.

There is a second reason for talking of a token-identity theory of mental and physical events. This is that Davidson suggests that causation should be analysed as a dyadic relation between events. This account of causation differs, for example, from the Cambridge philosopher Hugh Mellor's claim that causation is a fact that relates facts. The relevance of causation to the token-identity theory will be made clear shortly, but we will speak more loosely and informally of a token-identity theory of *events* or *states* or *facts*.

equated with any physical properties. (In fact Davidson allows that mental and physical types could be aligned but this would be a matter of accident rather than law.) Thus Anomalous Monism might be called non-reductionist or non-reductive monism or materialism.

Supervenience

The role of supervenience in Davidson's philosophy

There is one other feature of Anomalous Monism that is merely hinted at in 'Mental events' but will be important shortly. This is the claim that the mental *supervenes* on the physical. During the last 30 years, Davidson has taken different views as to whether supervenience is an ingredient of Anomalous Monism or not. But it is defined in the following way: *determining or fixing the physical properties of a person (or possibly the person and their environment) determines or fixes their mental properties but the converse implication does not hold.*

One part of the reason to hold supervenience to be true is in response to the intuition that mental properties could be realized in more than one way. One of the things that makes this a plausible intuition is an analogy with computers. In their case, the same type of software state can be realized on *different* kinds of hardware and thus by different types of physical state. Thus determining the type of mental or software state does *not* determine the type of physical or hardware state. This explains the second half of the supervenience claim: that fixing the mental does *not* fix the physical. But why should one believe the first half? Why believe that determining the physical determines the mental? This will be the subject of Session 5 in this chapter. But a preliminary case can be made as follows.

Supervenience in ethics

Supervenience was first deployed in moral and aesthetic cases. There the analogue of the negative argument is the thought that moral values can be realized by different physical set-ups. The colour of the murderer's tie, for example, is of no consequence to the evil of the act. But on the other hand, imagine that all the physical properties that go to make up an act were reproduced: the movement of the hands, the trajectory of the knife, etc. Could the moral value of the act be different? Most of us would say no. Moral properties may not be reducible to physical properties but nevertheless they are not completely free floating. The same kind of intuition applies to the mental.

Here the analogous thought experiment is whether two people who were (qualitatively) identical with respect to their physical properties could differ in their mental states? By claiming that the mental is constrained by the physical through the relation of supervenience, Davidson attempts to demystify mental properties. As well as individual or token mental events being identical with physical events, mental properties are fixed by physical properties. As we will see shortly, there is a further reason for invoking supervenience. But the combination of token-identity thesis plus supervenience is not without its critics.

Mental properties and the charge of epiphenomenalism

Davidson aims, via Anomalous Monism, to reconcile the non-reducibility of the mental to the physical with a causal role for the mental. But one recent focus of criticism has been that he fails to make the mental causally relevant in the right sort of way. The problem might provisionally be put like this. Mental states rationalize and justify other beliefs or actions 'in virtue of' their mental properties. They are causally active 'in virtue of' their physical properties. But their mental *properties* are causally irrelevant. Suppose that the mental properties of an event are not essential to them being the event they are (in the way that the colour of a person's hair is not an essential property of them). Then an event would have just the same causal consequences whether or not it possessed its mental properties. They do not play a role in the causal order.

Davidson's twofold response

Davidson has made two complementary responses to this kind of criticism:

1. He emphasizes that causal relations are extensional. If one event causes another then they stand in a causal relation that holds of them no matter how they are described. This case contrasts with an intensional (with an '*s*') relation such as acting intentionally (with a '*t*'). Described one way an act may be intentional (turning a light on) while in another it may not (alerting a prowler). Because it is description-sensitive it is intensional. But causal relations are extensional. Thus one should not say that an event causes another 'in virtue of' its properties or description (as we did above for temporary clarity). It either causes it or not, however described. Thus mental events do not cause other events 'in virtue of' their physical properties at all.
2. He argues that providing the mental supervenes on the physical, if two events differ in mental properties (for example one has mental properties and the other does not) then they must differ in at least one physical property. Thus they will have different causal powers.

But are these adequate?

The relative strengths of the objection and Davidson's responses are by no means clear-cut (it is the main subject of the collection of essays edited by Heil and Mele in *Mental Causation*, 1993). But it is worth noting that the appeal to supervenience to ground causal efficacy may not be as convincing as it at first appears here. Think again of the intuition that moral properties supervene on physical properties. Thus if two events differ in moral properties then they must differ physically as well. A similar intuition might be thought to apply to aesthetic properties. If two pictures differ in their beauty then they must also differ in some physical property. (Although one might wonder whether a spectacularly good forgery was as good as a genuine Picasso. Financial worth certainly does *not* supervene on (non-relational) physical

properties alone.) Granting for the sake of argument that these two other examples are cases of supervenience, does this imply that moral and aesthetic properties are causally relevant? Most people would think not. The supervenience relation is a meta-physical relation. It does not itself transmit causal relevance from physical properties to the properties that supervene on them.

If supervenience is not a satisfactory response to the claim that Davidson's account makes mental properties epiphenomenal, one possibility is to accept that mental properties are unlike physical properties and are not part of the causal order of the world. In this chapter we have simply assumed, so far, that the mental does play a causal role. This is a claim that Davidson makes explicit but which is also implicit in functionalism and type-identity forms of physicalism. We will return to arguments for and against causal accounts of the mind in Chapter 26.

Does supervenience cohere with Anomalous Monism anyway?

As well as the question of whether supervenience ensures the causal relevance of mental properties, Davidson's use of it has also come under separate criticism. One issue is whether it can be coherently combined with a merely token-identity theory. Another is whether supervenience is itself a coherent response to an underlying intuition about the relation of mind and body, or whether it 'inflates' into full-blown psychophysical laws. In the rest of this session we will begin to examine these in the context of Davidson's account.

Does supervenience imply psychophysical laws thus contradicting the claim of anomalousness?

A key tenet of Anomalous Monism is that there are no psychophysical laws. But, as Simon Evnine (1991, chapter 4) argues in his *Donald Davidson*, supervenience implies the existence of something like psychophysical laws. 'No mental change without physical change' implies 'same physical state implies same mental state'. Thus if supervenience holds, and if two people are in the same type of physical state, then they must be in the same type of mental state.

This, however, comprises 'half' of a psychophysical law in that it allows inferences to be drawn from the physical characteristics of a person to conclusions about their mental characteristics. This in turn seems to allow the ascription of mental states in ways that do not necessarily respect the central role of rationality in that ascription. One way of bringing out the worry here is to focus on the fact that it seems to be merely a contingent matter whether or not one could use a collection of these laws to ascribe wildly incompatible sets of beliefs to someone. This runs counter to the Davidson's view that it is an essential or necessary feature of mental states that they stand in rational relations.

(Consider the following thought. Suppose one person's neural state 123 is identical with the belief that it is raining and neural state 234 is identical with the belief that it is not. Now in general, according to Davidson, one cannot believe contradictory

beliefs. This follows, he argues, from the claim that it would never be rational to ascribe such beliefs to another and such ascriptions are the only way of finding out about others' beliefs (see Chapter 25). But there seems no reason to hold that this necessary principle about the combination of mental states applies to physical states. Even if, as a matter of fact, neural states 123 and 234 were never combined in the person in the actual world, there seems to be no a priori reason why they could not be in some possible world. And now we can imagine that this might happen quite generally, contrary to Davidson's insistence about the overall rationality of the mental.)

William Child, however, proposes a response to this in *Causality Interpretation and the Mind* (1994, chapter 2, sections 1–3). The mental side of these half psychophysical laws still have to be arrived at in the normal way we ascribe mental states to one another. But these everyday methods of mutual interpretation respect the central role of rationality. Thus the link we establish between physical states and mental states always depends on first having interpreted the mental state using principles of rationality. Given that such interpretation depends on behaviour, it would always produce the same results for physically identical behaviour. Thus using the half psychophysical laws of supervenience is both subject to rational interpretation (to establish them) and will never deliver results that run counter to rational considerations.

Does supervenience inflate into psychophysical laws?

But there is another argument which Evnine (1991) aims against supervenience. This is that supervenience is an inadequate response to a particular metaphysical intuition. His argument is based upon progressive modifications to a thought experiment related to one we have already discussed. Imagine that there are two people with (perhaps only slightly) different *mental* properties. Supervenience guarantees that there must be at least one physical difference between these two. But would there be anything wrong with the story if we were told that this difference comprised merely the length of their eyelashes?

Evnine suggests that it would be inconsistent to insist that the mental supervenes on the physical but to be satisfied in this case that the physical difference accompanying mental difference might be something as irrelevant as eyelash length. But this objection would also apply if the physical difference concerned a part of the brain that has nothing to do with the mental properties concerned. One should only be satisfied by a difference in whatever parts of the brain are causally responsible for mental properties. In short, Evnine suggests that the metaphysical intuition that motivates the adoption of supervenience cannot be satisfied by anything short of full psychophysical laws, which in turn undermine Davidson's merely token-identity theory.

Given the kind of findings discussed in the first reading of this chapter (Posner, 1993), Evnine's argument is clearly relevant to any neurologically informed approach to psychiatry. What such an approach appears to show is that quite specific parts of the

brain are involved in particular kinds of mental activity. Because of this, the claim that the mind supervenes on the brain appears to be a reasonable methodological assumption partly justified by subsequent empirical findings. But one might hold that, although parts of the brain were responsible for mental activity (either by causing it or by just being that very activity seen from a different perspective), there were no strict laws that linked mental types to underlying physical types. One motive for the latter doubt is the extent to which brains can differ in their physical properties. Evidently's argument appears to undercut such a modest middle ground. Supervenience inflates into either a type-identity theory or at least causal laws linking types of mental and physical state.

What is the moral of these objections?

Evidently's argument has a prima facie plausibility. On the assumption that one already believes in a token identity version of physicalism, supervenience adds a further constraint to the effect that the mental properties of (physical) events are themselves fixed by their physical properties. Without supervenience two physically identical persons could differ in mental properties. A token-identity theory by itself does not rule that out because they are numerically distinct persons (they are not the same person or they are at least temporally distinct stages of the same person) their mental states will not be numerically the same (the very same token of the same type of state) and thus there is no reason without supervenience to believe that they will be of the same type. But once supervenience is added it seems implausible to stop short of the claim that the mental properties of an event must supervene on a specific set of physical properties of that event (although so far unknown to us). Thus it inflates into a psychophysical law.

We will return to this claim a little later in the chapter. But one possible line of defence of supervenience is the thought that it is that constraint in combination with the token-identity thesis which is problematic. So one possibility is to keep supervenience and discard the token-identity thesis. We will look at arguments for that in Session 5.

As Evidently makes clear, supervenience need not be construed as an integral part of Anomalous Monism, which comprises at heart a token-identity theory and a denial of reductionism. It is an 'add-on' and thus one possible response to the arguments directed against it would be simply to dispense with it. (This is what Davidson has sometimes suggested despite also invoking it to explain the causal relevance of the mental.) Thus a different alternative is to abandon the supervenience claim while retaining the token identity claim. We will examine an argument against identity theses in the next session.

The many responses to dualism

What can already be seen from the discussion so far is how many different opposing positions there can be to Cartesian dualism and how much they can differ in their form of materialism. So to return to the issue raised in the discussion on clinical neuropsychiatry in

the first session, it is not a straight-forward choice of hard line reduction of mind to brain or unscientific dualism. Philosophical reflection reveals both a number of different options and some critical arguments in addition to clinical findings, by which to judge them.

Consider Box 23.2 regarding some of the positions taken in the mind-body debate. So far we have encountered, in this and earlier modules, the first four positions.

In moving on to consider alternatives to Davidson's position it is worth bearing in mind what the other options are for explaining the relation of mind and body. If, for example, it turns out that even a token-identity theory makes too strong a connection between the mental and the physical then one will be forced to adopt a position lower down the order such as 5a or 5b, 6, or even 7.

Note also that the list in Box 23.2 contains only some of the available positions. Eliminativism, the view that psychological

Box 23.2 Some of the positions taken in the mind-body debate

1. *Type-identity physicalism.* Type of mental state/event are identified with types of physical state/event—so mental properties.
2. *Token-identity physicalism + type identity behaviourism.* Each individual mental event is a physical event + types of mental state are types of dispositions to behave.
3. *Token-identity physicalism + type identity functionalism.* Each individual mental event is a physical event + types of mental state are types of internal causal state individuated by their function (inputs and outputs).
4. *Token-identity physicalism + supervenience.* Each individual mental event is a physical event + fixing the physical state fixes the mental state but not vice versa. All world events are thus fixed by their physical properties only. A weak property dualism.
- 5a. *Supervenience without identity (weak supervenience).* The full mental history supervenes on the full physical history but no elements can be correlated/identified.
- 5b. *Token-identity physicalism without supervenience.* Each individual mental event is a physical event but fixing the physical history (of bodies? of whole worlds?) does not fix the mental history.
6. *Full property dualism (without supervenience).* Mental events are instantiations of mental properties which are as fundamental as physical properties—physical properties do not fix the history of worldly happenings.
7. *Substance dualism.*

mentalist descriptions should be discarded in favour of physical or rather neurophysiological descriptions, is an obvious omission. (Perhaps it belongs at the very top as position 0.)

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What general philosophical claims does Davidson aim to reconcile?
2. How does his account of mind attempt this and what is it called?
3. What is supervenience? Is it compatible with Davidson's model of mind?
4. How many possible solutions are there to the mind–body problem?

Session 4 Arguments against mind–body identity theories

This session will examine two arguments against mind–body identity theories. The first is a general argument developed by Kripke from his work on the semantics of modal claims. In other words, it concerns the meaning of claims about necessity and possibility. Nevertheless it employs simple intuitions (which may or may not be correct!) about what is and what is not necessary and possible.

The second is an argument best set out by Child (1994) against Davidson's token-identity theory. It is, in effect, an application of one of the arguments from McDowell (1985) in the reading earlier in this chapter (in Session 2).

The development of some form of identity theory has been a key tool in contemporary materialist or physicalist accounts of the mind. Recall, the point of talking of an *identity* thesis is that this is a clear way of saying (with the first neuropsychiatric reading) that mental language and physical language are different ways of talking about *the very same things*, or events, or states. If it turns out that there are arguments against even such a weak construal of this approach as a token-identity then this is a significant blow to materialism.

In fact the two arguments considered here differ in their focus. Kripke argues against construing the *relation* between mental and physical states as one of strict identity. But, it will turn out, that this need not dramatically affect clinical findings. Child (1994), by contrast, attacks the very idea of aligning mental and physical states or events whatever the nature of the relation that is then supposed to hold between them. This does have profound consequences for what might be found clinically.

The background to Kripke's argument

Kripke's argument against mind–body identity

Saul Kripke is most famous for writing two books. One is an interpretation of Wittgenstein's later philosophy in which Kripke advocates a form of scepticism about meaning (see Chapter 25). The other, *Naming and Necessity*, is a prose interpretation of Kripke's own formal work on modal logic in which he explores the semantics (or rules governing the meaning) of names and descriptions. The two works are thus in some tension. The argument against mind–body identity theories is contained in the latter.

The background to Kripke's argument

In order to understand Kripke's argument against identity theories, one must first understand a little of the background that concerns the connection between names and necessity. The key claim that grounds the argument is that an identity statement that links natural kind terms—such as types of mental state and types of physical state—holds of necessity and the source of that necessity is genuinely metaphysical rather than the result of a linguistic convention.

In other words, statements of the form: 'necessarily $X = Y$ ' or 'necessarily X has feature Y ' do not just report rules for the use of the symbols ' X ' and ' Y '. Kripke rejects the idea that the source of the necessity in these lies in a connection forged in language and depends on the way the thing is described. This is called a *de dicto* account of necessity. In its place he favours a *de re* account, which is implicitly a form of essentialism. The origin of these claims about necessity in turn lies in his account of how names work especially within counter-factual conditional claims. (Remember, a counter-factual conditional claim is a claim about

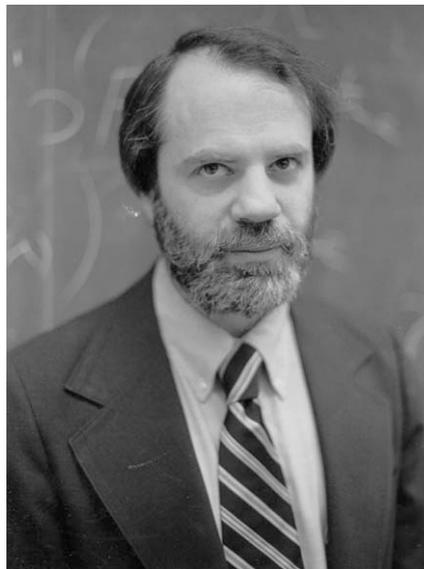


Fig 23.1 Saul Kripke

what would have been the case if some fact that is not or was not true had in fact been true. They were discussed in Chapter 15 on causation.)

Kripke's brief history of descriptive theories of names

Kripke gives the following potted history of the philosophy of names. (As an accurate intellectual history it leaves much to be desired but that is neither Kripke's nor our purpose here.) John Stuart Mill (whose inductive methods were discussed in Chapter 16) claimed that names have *denotation* but not *connotation* or *reference* but not *sense*. (This contrast will become clearer shortly.) But if so, the philosophers and logicians Frege and Russell asked at the turn of this century, *how* do we know what the name picks out? Or, to put the point less epistemologically, what is the connection between a name and a thing?

Frege and Russell instead proposed that what we generally take to be names stand for and mean some definite description such as the 'the leader of the Israelites'. (In fact there are important differences between Russell and Frege that Kripke ignores. Russell, for example, distinguished 'names' such as Napoleon for which this is true from 'logically proper names' like 'this', 'that', and 'I' but this need not concern us.)

But how should one pick out just one defining attribute? A natural modification of the idea is that names stand for a definite description is to employ instead a cluster of descriptions. The contemporary philosopher John Searle argued for such a position in the 1960s. Moses, for example, is that individual who possesses the majority of a cluster of descriptions for which his name stands: the leader of the Israelites, who was found in bulrushes, whose brother was Aaron, etc.

Theories of names of this sort aim to explain:

1. how a name can point to an individual person or thing including those who are no longer present or alive;
2. the content of identity statements such as that 'Hesperus is Phosphorus' (the Greek names for what turns out to be the same planet: Venus). Such statements can be informative on this account because they connect two different sets of attributes.

Kripke's attack on descriptive theories of names

Kripke rejects outright the idea that names latch on to things via descriptions. His main argument is that that theory cannot account for how we can describe counter-factual possibilities. But there are a number of other criticisms worth note in passing:

1. We might not know any of the cluster descriptions associated with a name. Alternatively we might know of two people who share the few things we know. Perhaps all we know about Feynman is that he was a famous physicist. So were many others.
2. We might not be able to specify the attribute independently. If all we know about Einstein is that he devised 'Einstein's theory' and all we know of that is that it was Einstein's theory, we cannot latch on to Einstein.

3. Or properties may be satisfied by the wrong person. Columbus was not, after all, the first to discover America. But when we talk of 'Columbus' we do not mean the Vikings or whoever really first discovered it.

But the main objection Kripke raises to descriptive theories of names is that Moses *might not* have led the Israelites or done any of the cluster of things that we use to pick him out. Nevertheless if we talk about the counter-factual case in which Moses led a quiet life, we are still talking about *him* and contemplating possible circumstances for *him* (even if under those circumstances Aaron would have risen to the occasion and done those things actually done by Moses).

The moral

The conclusion Kripke draws is that definite descriptions or property clusters are a bad analysis of names because they behave differently in counter-factual circumstances. In another possible world the individual who satisfies the description 'the winner of the 1997 UK general election' might not have been Tony Blair. In such worlds Tony Blair might have narrowly lost the election, or might have taken no part in it having chosen to continue his earlier career in the law in which case various descriptions true of him in this world would not apply. In other circumstances he might not even have been called 'Tony Blair' if his parents had preferred the name 'Algenon'. But we could still talk about what Tony Blair would have done, had he been called 'Algenon' and been a successful lawyer. 'Tony Blair' always picks out the very same person in different possible worlds. But descriptions such as 'the winner of the 1997 UK general election' need not.

As Kripke puts it, names are rigid designators and definite descriptions or property clusters are non-rigid designators.

- ♦ *Rigid designators* (e.g. names): pick out the same individual in every possible world in which they pick out anything.
- ♦ *Non-rigid designators* (e.g. definite descriptions): pick out different individuals in different possible worlds.

Necessary identity statements

Given this definition then if two rigid designators pick out the same individual in one world, they will pick out that same individual in all worlds in which either applies. This follows from the definition of rigid designators that Kripke proposes. But it is supposed to answer a prior intuition we have. Thus if an identity statement such as 'Hesperus is Phosphorus' is true in any one possible world, there will be no possible world in which it is not true. There will be possible worlds in which neither name applies but none in which only one does. Thus if it holds it holds necessarily. By contrast 'Moses was the leader of Israelites' is not a necessary truth because it contains a non-rigid designator: 'the leader of the Israelites'.

Natural kind terms

One intuitive theory, similar to the cluster theory of names, defines natural kinds in terms of a cluster of properties. Gold,

for example, might be a yellow, malleable metal. But, Kripke argues, we may *discover* that gold is not usually yellow. Or it might not have been usually yellow had other things been different. Or, similarly, we might find that some yellow-striped tiger-like things are really reptiles while others are not. In this case we should not conclude that some tigers are reptiles but that some of what we took to be tigers are not and are reptiles. Thus, Kripke argues that natural kind terms are not defined by clusters of properties but are also rigid designators (like names), which designate things or stuff with the same internal structure, whatever it is.

Natural kinds and identities

If natural kind terms are rigid designators then the same considerations which apply to identity statements linking names will also apply to identities between natural kinds. Thus if ‘water is H₂O’ is true in any possible world it will be necessarily true.

Some cautionary notes on possible worlds and other matters

1. Possible worlds are stipulated not observed. There is therefore no need to consider the epistemological question of how we could recognize individuals under other circumstances widely different from those in the actual world. Suppose we speculate on what would have happened if Tony Blair had become a plump sports commentator with no interest in politics who changed his name to ‘Desmond’. One might think that this would raise a problem of how we might find him in this possible world. How would we recognize that that person was Tony Blair under different circumstances? But that is a mistake: we can simply specify that we are still talking about *him!*
2. Some *necessary* conditions govern stipulations about individuals in other possible worlds. Bill Clinton could not have been an armchair. (There might have been an armchair we called ‘Clinton’ for amusement.) Similarly, this! table (on which I am writing) could not have been made of ice. (A qualitatively identical one might have been ice but then it would have been this! table.)
3. Sometimes the reference of a name is fixed via a definite description, e.g. ‘Hesperus is the evening star’ but this description does not enter the meaning of the name—it doesn’t have a meaning. Had Hesperus been hit by comet and pushed to a different orbit it would not have been an evening star but would still have been Hesperus (although presumably not called ‘Hesperus’ by the people in that possible world).
4. Recall the three distinctions between sorts of truth: necessary versus contingent, a priori versus a posteriori, analytic versus synthetic. These distinctions are, respectively, metaphysical, epistemological, and semantic. Since Hume there has been a school of thought that these distinctions classify truths into the

same two sets. Any truth that is necessary will also be a priori and analytic. (Kant’s claim that there are synthetic a priori truths is a notable exception.) There is an intuitive link between necessity and a priori and vice versa (if a truth is necessary, one does not need to look which world one is in) and analytic truths will be true in virtue of meaning in all possible worlds. Kripke claims instead that there are both a priori contingencies (the Paris bar is one metre) and a posteriori necessities (Hesperus is Phosphorus).

5. Kripke’s positive account of naming depends on there being suitable causal relations between names and the objects named, mediated through the social use of the name. Causation provides the answer to the question the cluster theory was meant to answer. This is a notoriously problematic claim. How can the correct functioning of the causal claim be specified non-question beggingly? But it lies outside the scope of this book.

The argument against mind–body identity theories

We can now examine Kripke’s argument about mind–body identity theories. Kripke argues that identity claims for natural kinds as well as for names express a posteriori necessities because natural kind terms are also rigid designators. Thus ‘water = H₂O’ is found by experiment but expresses a necessary linkage because both ‘water’ and ‘H₂O’ are rigid designators. But, he argues, this shows that the claim that pain is a certain type of neural state (e.g. state N) is not a true identity statement because we have independent reason to believe it is at best merely contingent. (Had it been true it would also have had to be necessarily true.)

EXERCISE 4

(30 minutes)

Read the extract from:

Kripke, S. (1980). *Naming and Necessity*. Reprinted in Rosenthal, D. (ed.) (1991). *The Nature of Mind*. Oxford: Oxford University Press, pp. 236–246. (Extract, pp. 242–243)

Link with Reading 23.4

- ◆ Try to identify the kernel of Kripke’s argument.
- ◆ What is the role of the Cartesian intuition that one’s mental states need not be physical states?

Kripke’s argument against mind–body identity theories can be summarized like this:

- i Assume that pain = neural state N.
- ii If an identity holds between rigid designators then it is necessary.
- iii Both ‘pain’ and ‘neural state N’ are rigid designators.
- iv Therefore pain = neural state N is a necessary truth.

- v But, by the Cartesian intuition, pain = neural state N is (at best) contingent. (Note that all that is required is that the Cartesian claim is possible not that it is true.)
- vi Therefore it is both necessary and not necessary. This is an absurdity so one premiss must be rejected.
- vii Therefore premiss (i)—the identity statement—is false.

Because this form of argument is *reductio ad absurdum* it is a matter of choice which premiss to reject. All that can be directly inferred is that at least one premiss must be false. But Kripke provides some argument that premisses (iii) and (v) are true and this leaves only (i) to be false.

Premiss (iii)

'Neural state N' names a natural kind (like electron) and thus picks out the same kind of state in every possible world. Thus it is a rigid designator. 'Pain' is also a rigid designator. Kripke says that if something is a pain it is essentially so. It would be absurd to suppose that a pain could have been some phenomenon other than it is. This contrasts with the case of 'the 42nd president of the USA', who might have been a different man.

The Cartesian intuition

This intuition is that either my pain state could have existed without some particular brain state or that my brain state could have existed without the pain—i.e. without consciousness. Either seems possible. How could this intuition be rejected for a principled reason? What would justify the counter claim that in no possible world could this pain have existed without its particular corresponding brain state?

Could the argument undermine the claim that water = H₂O?

Kripke's argument does, however, face an obvious counter objection to which he sets out a response. The objection is that there is a similar intuition to the Cartesian intuition available in the case of water and H₂O. We may claim to have a similar intuition here that would undermine the identity for water and H₂O, which, Kripke supposes, is a true identity. The argument would be of the same structure. If there is an identity between water and H₂O it must be a necessary identity. But it is at best contingent. Therefore there is no identity.

Kripke claims that one can defuse this second argument by explaining away the intuition on which it relies. He suggests that the intuition of contingency is the result of confusing the false thought that water might not have been H₂O with the true thought that a substance that brought about qualitatively identical experiences in us when we looked at, touched, and tasted it might not have been H₂O. But that substance would not have been water so there is no real Cartesian intuition here. Thus the identity is not undermined.

One might therefore think that the Cartesian intuition could also be explained away in a similar manner. But according to Kripke it cannot because if some state has the 'look and feel' of

pain it simply is pain (whereas to look and feel like water is not necessarily to be water). If one has the experience as of pain, one has pain. Thus the intuition cannot be explained away and the identity of pain with any neural state cannot hold.

Responses to Kripke's argument

Two counter-arguments

Is Kripke's argument an elegant refutation of any identity theory of mind and body? If so, what is the connection—if any—between them? Kripke himself provides no positive account saying merely: 'I regard the mind–body problem as wide open and deeply confusing.'

McGinn's response to Kripke

One response to Kripke's argument is to argue that the Cartesian intuition on which it is based can in fact be explained away. In 'Anomalous Monism and Kripke's Cartesian intuitions' (1977), Colin McGinn argues that this is possible providing that one adopts a token rather than a type-identity theory of mind and body. He suggests that Kripke is right that if an identity theory is true then it is necessarily true. But this is consistent with a merely token identity. Thus, if a token mental state is identical with (i.e. it just is) a token physical state, it is so necessarily. Thus the Cartesian intuition that a particular pain might not have been a physical token is misleading because it is strictly false. But it can be explained away as the result of a confusion with the true claim that a state qualitatively (but *not* numerically) identical to a particular mental state might not have been this physical state nor a physical state like it with respect to its physical properties. Thus a token identity can explain away the Cartesian intuition in a similar way to Kripke's account of the case of water.

It is worth taking this again. If a merely token-identity theory is true, then mental state types gather together token states that are qualitatively the same as far as their mental properties go, but which are very different from a physical perspective. Thus the intuition that a particular mental state token could have had a physically different underpinning is strictly false but there is a true intuition with which it might be confused. That is that a mental token of the same mental type might have had a different physical basis.

Kripke, however, makes the point that there are two related broadly Cartesian intuitions. One is that one's mental states might not be physical. The other is that one's physical states might not be mental. McGinn can accommodate—by explaining away—the former, but not the latter.

Feldman's 'third-person response' to Kripke

Fred Feldman attempts to undermine Kripke's interpretation of the Cartesian claim that it is possible that a given mental state is not identical with any physical state in a different way (Feldman, 1980). He claims that it is possible that a token mental event might occur without any awareness of it by a subject. But this possibility depends on the idea that the qualitative, characteristically

mental, aspect is not an essential property of the event. Thus construed a token-identity theory is a kind of contingent identity theory: all mental events are physical events but it is a contingent feature of these events that they are mental in the sense of having mental properties. A particular or token event that has mental properties in our world might have existed without having its mental properties in another possible world. A pain state or event would still be that very pain state or event but shorn of the contingent property of being felt by someone.

The idea that events have essential and non-essential properties fits with Kripke's general account of necessity as far described. It is a necessary property of Bill Clinton that he is a human. Bill Clinton could not have been a mouse (although there might have been a mouse called 'Bill Clinton'). But it is not a necessary property of him that he is or was the president, nor that he often wears a suit, nor even that he is his current height or weight. These latter properties could have been different. They are non-essential. Feldman claims that possession of mental properties is a non-essential element of the events which in our world have them. It is in this specific claim that he disagrees with Kripke.

Given this counter-intuitive account of mental events, Feldman can explain the other Cartesian intuition away and thus defuse Kripke's argument against identity theories. The intuition that the physical event could have occurred without the mental event is mistaken (as both are the very same event, if one happens, the 'other' *must* happen) but it is true that it could have occurred without its mental *properties* such as being felt. As Feldman advocates a token-identity theory he can employ McGinn's (1977) argument to defuse the other intuition that this pain, for example, might not have been a physical state.

What is the moral of Kripke's argument?

Kripke and Feldman trade opposing intuitions. Kripke claims that its painful feel is an essential property of every pain. Feldman denies this. One consideration in support of Feldman's position is this. Kripke's argument requires that both sides of identity statements are rigid designators. The physical description could well be. Types of neural state might be like types of atomic state in being natural kinds. One feature of them being natural kinds is that there is some independence between the underlying kinds and the way, relative to a given body of scientific knowledge, they are picked out. Particular experiential properties might turn out to be poor guides to the underlying kinds. (Kripke makes this point by saying that gold might turn out not always to be yellow and malleable.) But 'pain' does not seem to label a kind in the same way.

Two further considerations support this. First, there does not seem to be the same independence between the kind and how we pick it out. Kripke emphasizes this by saying that to feel like pain is to be pain. But if so this makes pain behave quite differently to other natural kinds. (To feel like gold is not to be gold.) Secondly, pain, unlike gold, does not seem to have a constant underlying *physical* structure. This stems from the thought that it could be

multiply realized. Why in that case should one think of it as a natural kind at all? What other underlying structure might it have?

On the other hand, to say that mental states possess mental properties only contingently presupposes the following sort of picture of mental states. They are free-standing internal states of the body, which also, contingently, possess the qualitative or intentional properties that they do. The question of whether this is a cogent conception of mental states is one to which we will return throughout this Part. It is a key element of an identity claim. But as will be made clearer in later chapters, the cost of this assumption is that it makes the answer to the question of how some mental states possess *intentionality* mysterious.

Note also that neither McGinn's argument nor that of Feldman can help with the intuition that a pain might have been non-physical. Token-identity physicalism is physicalist! So if physicalism amounts to the claim that the class of mental events is identical to a subclass of physical events, then Kripke's argument should apply to this class-wide identity claim.

What is the alternative to the *identity* of mind and brain?

If not identity then what? 'Composition' or 'realization'

It thus seems that if the modal intuitions on which Kripke bases his argument are correct then there cannot be an *identity* between mental states, events, or properties and physical states, events or properties. But if there cannot strictly be an identity theory, what sort of relation might there be between mind and body?

One possibility is to retreat to a less metaphysically loaded notion. Think of the relation between a car and its components. Suppose one were to claim that a car is simply identical to its components (perhaps in a working structural configuration). The reason for thinking this is that there is in some sense nothing more to the car than the components. Remove all the components from a driveway and one has ipso facto removed the car. But according to Leibniz's Law, if two things are numerically identical (if they are really just the same thing) then they cannot differ in any properties. But the car I get back from a garage mechanic after a service may not have exactly the same properties (perhaps it now performs better) and it may not have the same components (because the brakes have been replaced). So if the 'first' car is identical to its components and the 'second' is to its components and these are not the same then the two cars are not the same.

As a result of such examples philosophers sometimes talk of the car being *composed* of its components while not being identical to it. Similarly, a ring may be composed of a certain amount of gold but is not identical to it. The ring can be destroyed in a furnace without destroying the gold for example. A building is also composed rather than identical to its bricks and concrete.

It seems, in other words, that Kripke's argument does not the count against *some* version of the claim that psychological and neurophysiological descriptions are different descriptions of the

same thing. This might be true in the way that there could be descriptions of the historic buildings of England both in architectural and in stone-masonry terms. Retreating to constitution does not preclude some such relation of descriptions.

But there is some cost to the retreat to constitution. By contrast with the claim that water is H₂O (a strict identity claim) the claim that a ring is composed of its gold is less explanatory. The former identity explains the properties of water through the properties of H₂O. But while properties such as density and conductivity are transmitted from the gold to the ring, these do not fix others, such as its shape. And those properties are not transmitted from the ring to its substance—gold—in general. So if the mind is merely *constituted* by brain stuff, the brain stuff alone does not explain the nature of mind.

Child's non-modal argument against a token-identity theory

Kripke's argument turns on the following modal claim: since the identity between mind and body would be at best contingent, but since any identity statement linking such rigid designators must be necessary, there can be no such identity. As we have seen, one possible response to this is to retreat from the metaphysically loaded idea of identity and retreat to a weaker relation such as constitution. As this relation appears to work in the case of many objects and their parts, such a retreat appears to pose no threat to the direction of neuropsychiatric research. Child (1994, chapter 2 sections 4–6), however, takes a different tack in his *Causality, Interpretation and the Mind*. He argues against explicitly a token-identity theory not through consideration of the nature of the relation between the mental and physical but instead of what is linked by this relation. If Child is right, retreating to claims of constitution is irrelevant to the key problem that is *correlating* in whatever way mental and physical events or states.

EXERCISE 5

(15 minutes)

Look back at McDowell's (1985) argument against functionalism. Functionalism is a form of type-type identity theory. It relates types of mental state to types of functional state (states identified by their causal properties). Now think about Davidson's token identity physicalism. Could McDowell's arguments also apply to this weaker position? Think for yourself before going on.

Child (1994) presents an argument against the modest token-identity theory. Recall the reading linked with Exercise 2 (McDowell, 1985, reprinted in McDowell 1998, pp. 387–398). A central argument he raised against functionalism could be summarized roughly like this:

- i Functionalism is a form of type-identity theory in which mental states are identified with functional states that are determined by codifiable functional relations, which resemble computer programs.

- ii Mental states (intentional states at least) stand in a rational structure. Rationality plays a constitutive role for what mental states are.
- iii But rationality cannot be codified.
- iv Because functionalism is committed to the identity of mental and functional states it is committed to the codification of rationality.
- v So functionalism is false.
- vi And we might add: so is any identity theory that links mental states with states which are so codifiable such as physical states by physical theory.

Child applies the same kind of argument against even a token-identity theory. If a particular mental state (or event) is identical with a physical state (event) then the rational structure of mental states will have to be mirrored by an isomorphous structure of physical events (as each mental event will be a physical event and both will be governed by mental or physical relations). However, that impossibly presupposes the codifiability of rationality. Thus even token identity theories are false.

One way of putting this point is that it is just inconceivably unlikely that the structure of causal relations and rational relations will keep in step over time if they are not bound together by psychophysical laws.

The moral?

The moral of this argument is that the token-identity theory is an unstable middle ground between two more stable positions. One is a reductionist type identity theory that does offer some—albeit unsatisfactory—explanation of what keeps the mental and physical in step. The other position denies that any identification can be made between mental and physical states. Davidson's position has the initial attraction that it promises to reconcile antireductionism about mental properties with an account of how mental states can be causal in just the same way as physical events. However, it is not a stable middle ground because the identification of mental and physical elements requires a harmony that it no longer has the resources to explain.

If Child's (1994) argument is right—and it presupposes the central constitutive role for rationality in the mental domain for which we have so far encountered only limited argument—then no form of identity theory seems plausible. Child applies the argument to token- as well as type-identity theories. If that is the case, is anything left of a materialist account of the metaphysics of mind? One answer to that is supervenience, which is the subject of the next session. However, that claim is just that sameness of physical properties implies sameness of mental properties but not vice versa. Such a relation could obtain between two discrete realms of substances. An epiphenomenalist dualist of substance might hold that physical changes in the brain causal mental changes in one's thinking stuff such that sameness of physical

state implies sameness of mental stuff. But this would be a dualist rather than a materialist position.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. What is Kripke's argument against the identity of mental and physical states? How is it based on his view of names and on possible worlds?
2. Is it necessary to believe that possible worlds are real?
3. What alternative view of the mind does he put forward instead?
4. Can the argument be refuted?

Session 5 Is there any reason to believe in supervenience?

If there are good reasons to doubt the soundness of an identity theory of mind and body, what is the relation between them? In addition to an identity (and more recently composition or realization), the other relation so far discussed has been that of supervenience. This session will consider the idea of supervenience in more detail.

EXERCISE 6 (10 minutes)

Supervenience was introduced above in the context of Davidson's Anomalous Monism as making the following claim: determining or fixing the physical properties of a person (or possibly the person and their environment) determines or fixes their mental properties but the converse implication does not hold.

- ◆ Could this claim be true in the absence of any sort of identity claim?
- ◆ What would supervene on what?

Haugeland's version of supervenience

It is easiest to think of supervenience in the context of an identity theory of some sort between the mental and the physical. But this is not necessary. All that is required for supervenience is something like this: the whole history of physical happenings determines or fixes the whole history of mental happenings without there being any further correlations. This latter thought amounts to this. There need be no event at the physical level that corresponds to a mental event. The two kinds of description of

the world might not be commensurable. Again, the easiest way of expressing the supervenience relation does trade on something like an identity claim. One might say: there are two different descriptions—mental and physical—of the same events. But this initial description can then be modified by saying that there need not be identifiable elements—events—common to both descriptions.

The US philosopher of mind John Haugeland set out this view in an article called 'Weak supervenience' (1982). In it he attempts to articulate and motivate a weaker and less contentious version of physicalism than Davidson's token-identity thesis (a thesis that we have already seen reason to criticize). The paper does this by first attempting to undermine Davidson's argument for his position (which, to reiterate, is an argument based on the reconciliation of three assumptions). Then it presents an illustration of an case of supervenience without identity (the case of loop and arrow languages described below). Finally, it attempts to suggest that there will be a great difficulty in finding identities between the mental and the underlying physical given the considerations already put in place.

1. Haugeland's argument against Davidson turns on the charge that Davidson equivocates on the meaning of 'event' between the claim that mental events play a causal role and the claim that whenever events are related as cause and effect then there must be a strict law linking appropriate types. Haugeland argues that the first claim requires a robust construal of events as distinct disruptions, sudden changes or happenings. By contrast the second requires a mathematical construal of related gradual variations in properties. Haugeland's suggestion is that mental and physical 'events' are unlikely to be picked out in the same way. A distinct mental event might not correspond to any particular distinct physical event.
2. Haugeland provides an argument that supervenience need not require the identity of elements within the two sets of descriptions (or 'realms') by describing a single such case. A language of loops and a language of arrow vectors can both describe patterns in a two-dimensional world (you might think of pixels on a TV screen) but the loop language is a richer higher level language. But it supervenes on the arrow language because loops depend on pixels. However, the two languages do not share an ontology. There are no elements that feature as individual things in both descriptions. Supervenience is a relation of true descriptions (or truths) between complete histories rather than requiring an identity of a common grid of, say, mental and physical events or states.
3. Finally, Haugeland considers both the relation of waves on the motions of water molecules and chess playing behaviour on lines of programming code. In both cases he argues that the lower level description fixes or determines the higher level description while also arguing that there is no candidate at the lower level with which to identify 'robust' events of the higher level.

Haugeland, however, presents no positive argument for supervenience

Haugeland presents intuitive and illustrative support for the idea that one can have supervenience between two sets of properties without being able to correlate individual features, items, events, or states in both sets. It is an argument against the assumption that supervenience requires a token-identity thesis. So even if one gives up that claim about the identity of mind and body as too strong this does not preclude holding that the mental supervenes on the physical.

Haugeland's paper, however, does not provide a positive argument for mental supervenience. It simply removes a potential obstacle to holding that view. (The obstacle being the assumption that supervenience requires a token-identity theory, which is itself implausible.) If one were not already a supporter of supervenience then Haugeland's paper would not provide further argument in its favour. So why should one believe that mental facts are determined by physical facts? What independent justification is there for supervenience or, related by physicalism?

An argument for supervenience

Why should one believe in physicalism (the claim that the world contains just what a true complete physics would say it contains)? In response to an attack on physicalism initiated by the English metaphysicians and philosophers of mind Hugh Mellor and Tim Crane, there has been a flurry of discussion mainly conducted in the journal *Analysis*. We will consider a short statement and defence of physicalism and supervenience by David Papineau and a critical response by Tim Crane.

An argument for a token-identity thesis of sorts...

In an article called 'Why supervenience' (1990) David Papineau offers a defence of supervenience. In a sense the paper attempts to defend a combination that we have already rejected: a combination of a token-identity theory and supervenience. For the moment, however, that does not matter to the question of whether Papineau provides any positive support for supervenience.

Papineau provides the following argument first for a token-identity theory:

1. Mental occurrences have physical effects.
2. Physical effects also have physical causes.
3. Physical effects are not overdetermined—there are not two sufficient causes.
4. The physical and the mental cause are the same.

Note that Papineau takes a relaxed attitude to what the claim of sameness here amounts to. It does not matter whether it is a strict identity claim or a weaker claim such as that the mental is *realized* or *constituted* by the physical. These weaker relations still provide a way to avoid the overdetermination of physical effects by separate mental and physical causes. It is, however, a further question of whether the yet weaker claim that Haugeland puts

forward fits Papineau's argument, although that is better postponed until later.

This brief argument so far is based on an underlying claim about the *completeness of physics*: all physical events are determined entirely by prior physical events according to physical laws. This is a substantial assumption that will also merit further consideration. But first let us look at Papineau's further claim that supervenience can also be derived from this principle.

... and for supervenience

Papineau's argument for supervenience goes like this:

1. Consider a mental event or state, which is the cause of some physical effect.
2. Given the completeness of physics (CP), the physical effect must be determined by a physical cause in virtue of physical features.
3. Now consider the more complex case of a mental event or state that is the cause of a *mental* effect and that this has some further physical effect.
4. Therefore the original cause has a physical effect via an intermediate cause.
5. By CP the middle (mental) cause must cause its physical effect in virtue of its own physical features.
6. By CP these physical features must be caused by the physical features of the first cause.
7. Given that the final effect is caused by the intermediate cause (we say that the final event is caused via the intermediate cause rather than just saying it is caused via its physical features), the latter as a whole (not just its physical features) is determined by the physical features of the first. So there are no independent mental causal powers.
8. The mental is not complete in the same way so no similar result can show that physical effects are determined by mental features. Some mental effects are the result of non-mental causes.
9. Thus whenever a mental cause has an effect that effect is determined by physical features of the mental cause. By contrast, only some effects of physical causes are fixed by mental features of physical causes. Thus there are no mental differences without physical differences, whereas there are physical differences without mental differences.

This argument aims to show that whenever a mental event causes either a physical event or a mental event, its physical features would be sufficient to underpin that causal relation. Thus if two people were in different mental states there would have to be some physical difference between the two of them because being in different mental states would have different (physical) causal consequences and these differences in causal powers stem from physical differences.

In fact it should be clear that this argument does not require that mental and physical tokens are strictly identical or even

merely *correlated*. The weaker claim that the mental *as a whole* is realized by the physical *as a whole* would suffice. No mention of identity is required to set out the underlying claim, which is that mental difference has to be accompanied by physical difference because otherwise there are events that are effects of non-physical causes. So on the assumption that the mental can play a causal role and that physics is complete, there can be an argument for supervenience, whether or not this is added to an identity claim. (And we have seen arguments that it should not be.)

A counter-argument

Crane's response

Tim Crane, however, responded to Papineau—in 'Why indeed: Papineau on supervenience' (1991a)—that the second assumption simply begs the question in favour of supervenience. His argument can be summarized as follows:

1. CP is just as contentious as supervenience. If one doubts the latter one will claim that some effects are determined by mental as well as physical causes. Such effects are *not* determined by physics alone and thus CP is not true.
2. 'Physics' cannot be defined in such a way that CP is both substantial and plausible. If it were simply (and *insubstantially*) defined as whatever would be included in an explanatorily complete science that need not rule out psychology as an essential part of a future complete science.
3. If, by contrast, physics is *substantially* defined by reference to the *micro*-sciences one could claim that by determining the micro-features of a situation one determines the macro-features. The motivation for this claim is the thought that at heart physics deals with the fundamental particles that make up matter and whose behaviour ultimately determines what happens in the universe. But if one simply assumes that micro-level phenomena determine what happens at the macroscopic level, and thus that the physical determines the non-physical this comes to assuming that there will be no macroscopic difference without microscopic difference. And this assumption is just supervenience again. One will in other words have simply assumed supervenience in the argument that was supposed to give it independent support.

Thus one needs an argument to establish the completeness of *physics* (hence *physicalism*) to establish supervenience in this way. What would one be? The alternative is that psychological factors may form part of a final theory of the world—the mental (mental properties) may be fundamental in a final description of the world.

A stalemate?

This dialogue between Papineau and Crane suggests a stalemate. If one can help oneself to the principle of the completeness of the physical then supervenience follows. But there seems little prospect of a truly independent argument for supervenience because that principle is too close to it. Now this need not matter.

All arguments start with premisses or assumptions of one sort or another. So one response would simply be to say that supervenience is an assumption founded on the likely course of physical theory (this is what Papineau has subsequently said).

In fact Crane has gone on to argue that things are worse than that for supervenience. Whether or not he is successful, he highlights a very real question. In the absence of psychophysical laws, *how* is the physical supposed to determine the mental? Examining the issues raised by this last optional reading will help shed further light on the first of Papineau's assumption: the mental is causal.

In 'All God has to do' (1991b), Crane puts forward a dilemma for a defender of supervenience, which turns on a theological thought experiment. Having created the physical features properties and laws of the world, what more does God have to do to create the mental features? Does he have to create anything further? Specifically does he have to create further psychophysical laws?

[D]oes God, in creating the physical facts, have to create laws linking those facts to the mental facts? If he does, then he has to do more than simply create the physical facts: he has to create laws in which mental properties figure. But if he doesn't, then the claim that he is creating genuine mental facts loses its bite: for if there are no mental laws, then arguable there are no genuine mental properties either. So the desired combination of physicalism and realism is unstable: the first horn of the dilemma threatens physicalism, while the second threatens realism. (p. 237)

The first horn

Crane's dilemma has the following underpinning. Consider, for a start, the option that God *does* have to create psychophysical laws in order, in addition to having fixed the physical, to fix the psychological. If this is so, then in the act of so creating psychophysical laws, God effectively creates *in addition to the prior act of fixing the physical facts* the mental facts. This follows from the fact that such laws will characterize the mental facts by characterizing (at least partially) the mental properties that figure in them. If so, then the physical facts alone do not fix the mental facts and thus the mental floats free of the physical in contrast to the assumption of physicalist supervenience.

The second horn

But if, on the other hand, God does *not* have to fix anything other than the physical facts, then, at least according to Crane, that undermines the idea that there are any mental facts to be fixed. At this stage in this chapter, this is the horn which requires greater consideration. (The motivation for supervenience is in part the thought that there are no psychophysical laws but that the mental has, nevertheless, to be domesticated as part of the physical universe.)

Taking the first horn first. Crane argues that if God does have to add psychophysical laws to the physical facts then these laws will (partially) constitute the mental facts. So God will, in this second act of creation, create further mental facts.

The second horn presents more of a problem for the position that reflection on the readings of this chapter have tentatively supported: something like Haugeland's weak supervenience or Child's physicalism without identity thesis. Crane argues that if there are no psychophysical laws then there are no mental facts. Part of his argument here turns on rejecting Davidsonian arguments against psychophysical laws—which we have echoes of in McDowell's and Child's arguments against identity theories. But we will focus here on the claim that without psychophysical laws there would be no mental properties.

An analogy with aesthetic and ethical supervenience

Crane considers the analogies that defenders of mental supervenience draw on from the cases of aesthetics and ethics. (Recall that this is how mental supervenience was introduced above.) Crane makes two criticisms of the idea that supervenience in these two cases provide analogical support for psychophysical supervenience:

1. In those other cases the supervenience thesis is a necessary truth. Crane plausibly glosses this by saying that it is part of the *concept* of aesthetics that there is such a relation to the physical. Anyone who doubted aesthetic supervenience would have a different understanding of the aesthetic. By contrast, according to Crane, mental supervenience is a contingent matter. And if so, what underpins its contingent truth are the (psychophysical) laws of nature.
2. Mental properties are supposed to be part of the causal order. Thus, first, they need to instantiate laws and secondly this undermines the analogy with aesthetic and ethical properties, which are not directly part of the causal order.

Neither of these points are decisive but they do constrain appropriate responses. Further light will be shed on the nature of the mental throughout this book. But one preliminary response would be, on the one hand, to accept both these arguments. That is one would accept that the claim that the mental supervenes on the physical is a necessary truth. But on the other hand one could also deny that mental properties are causal at least in the way that physical properties are. The first claim can be made more plausible by considering the supervenience base of the mental to be physical properties that describe behaviour rather than simply brain states. It requires construing with the behaviourists that there is a necessary connection between mental states and behaviour. We will return to this issue in Chapter 27. The second requires an abandonment of one of the key assumptions of Davidson's Anomalous Monism. We will return to it in Chapter 26.

It is also worth considering the kind of picture Crane himself advocates. He suggests that mental properties are not reducible to physical properties but are related to them by laws. This is like saying that the gravitational facts about the universe are not reducible to the facts about electromagnetism but they are related by (in this case physical) laws. Such a position shares some features with type-type identity theories (higher up the earlier table) in that it accepts

law-like relations between mental and physical properties or types. But unlike those positions, it does not accept that these law-like connections amount to ways of reducing mental properties to other properties (any more than electromagnetism can be reduced to gravity at least at the moment). So in another sense it is close to the bottom of the table in that mental properties are genuinely ontological distinct properties. There may not be mental substances but there are, according to Crane, distinct mental properties.

So what is the relation of mind and body?

EXERCISE 7

(15 minutes)

Think back to the original requirements for an understanding of the relation of mind and body. Do brain scans show the brain in action or the mind in action? What is the connection between the two? Is the answer obvious? What would be the consequences of adopting each of the different theories of mind and body summarized in Box 23.2 in Session 3 for interpreting brain scans?

The multiplicity of options

What the arguments in this chapter have shown is that there is very much more to the issue of thinking through what a sensible scientifically informed account of the relation of mind and brain might be. It is not, for example, a case of serious scientifically minded thinkers rejecting the absurdity of a substance dualist position and thus all agreeing on a common materialist position. There are many different forms that materialism can take.

But one of the general morals that the discussion in this chapter makes more plausible, if by no means obligatory, is that with respect to the general question of the relation of mind and brain, modesty is called for. The general trajectory of philosophical thinking in recent years has been away from the neat picture of type-type identity versions of physicalism. On such an approach, types of mental state simply are types of physical state. This promises a reduction of psychological properties to neurophysiological properties. But there is reason to believe that such an attractive picture is not true. As we have seen, it is much more likely that there are no such general connections between mental and physical descriptions.

A modest proposal

Now one response open to a philosophically informed neuropsychologist or psychiatrist is to accept such a general conclusion and make more restrictive claims about what brain imaging reveals. It does not reveal general connections between all forms of mentality and types of brain state. What it reveals are correlations applicable to humans (and their relatives) only. Thus it is still possible to discover fairly general connections between mental activity and brain activity, but these correlations are restricted to the human form of mind.

But the cost of this retreat is that it puts a key question beyond the reach of such studies. They will never show what it is that *all*

forms mind have in common. They will not, that is, reveal what is fundamental to having a mind, or intentionality, or consciousness, or whatever. To restrict one's interpretation to human mindedness allows for empirical discoveries about the relevance of particular brain structures to our having minds but sheds more limited light on the general question.

A computational route to generality?

Now one way round that point is to follow Lewis and combine a more modest physicalism with functionalism. That has the advantage that a physical story can be told about the working of the human mind and brain and a more general account can be told of what it is for any species to have a mind in functional terms. We already saw in the first reading something like this approach in the thought that mental activities had first to be broken down into functional components, which were themselves to be identified with brain activity. This is how neuro-*computational* psychology works.

But as the second session argued, there are a priori arguments against functionalist accounts of mind. Mental states stand in a different sort of structure than the law-like structure described by functionalism.

Weaker physicalism?

What would be the consequences of the weaker philosophical positions for the interpretation of brain imaging techniques? Davidson's Anomalous Monism involves the denial that there can be any law-like connections between mental and physical types (and we saw that it could be extended to rule out functional types as well), but retains an identity claim at the more modest level of individual tokens. In fact, despite his general argument that there are no correlations between the mental and the physical as a matter of strict law, he has in recent years suggested that this is consistent with there being local correlations at the level of lore. This might seem, therefore, an attractive philosophical interpretation of the results of imaging experiments.

But Davidson's position does assert that mental states are identical with brain states and we have seen that there are arguments any such idea of identity, even merely token identity. These arguments suggest that when one looks at the brain one never sees the mind, because mental states are different sorts of state to free-standing brain states. A brain state can be the state it is independently of what happens outside the skull, but mental states are relational and world involving and thus cannot be found simply within the skull. And while brain states stand in nomic relations of cause and effect, mental states stand in rational relations.

This leaves a yet weaker claim. Although not identical with brain states, mental states supervene on them. Such a claim is probably a methodological assumption of anyone working in the neurosciences. (They may also believe some stronger claims about the relation as well.) Supervenience ensures that where there are mental differences between people (or in one person over time) then there must be some physical difference, but it

does not dictate just what kind of difference is involved. We also saw that this assumption is not unproblematic. It is far from clear that an independent argument can be given in its support. But it does provide a background against which more specific empirical findings can be viewed. Such findings might well unpack the kind of causal preconditions there are in localized areas of the brain for being able to think in certain ways. But, as we have seen, this does not imply that when we image the brain we image the mind. Things are more complicated than that.

One of the features of the rest of Part V is that it will shed light on how the mental is constrained by the physical. It will be a thread throughout the discussion of meanings in subsequent chapters.

Reflection on the session and self-test questions

Write down your own reflections on the materials in this session drawing out any points that are particularly significant for you. Then write brief notes about the following:

1. Supervenience was introduced in Session 4 in the context of an identity theory of mind and brain. Can it apply without even a token identity theory?
2. Is there an independent argument that the mind supervenes on physical states? If so, on what other claims might it be based?
3. What are the two horns of the dilemma that Crane argues faces arguments for supervenience?

Reading guide

- ♦ The mind–body problem is central to the philosophy of mind and covered in introductions such as: Braddon-Mitchell and Jackson (1996) *Philosophy of Mind and Cognition* (chapters 1–3), and Burwood *et al.* (1999) *Philosophy of Mind* (chapters 1 and 2).
- ♦ For a thorough, but difficult, introduction to the recent philosophical debate about the relation of mind and body together with an original contribution to that debate see: Macdonald (1989) *Mind-Body Identity Theories*.
- ♦ A very useful collection of modern readings is edited by Rosenthal (1991) *The Nature of Mind*.

Davidson's Anomalous Monism

- ♦ Davidson's philosophy is discussed in Evinne's (1991) excellent introduction *Donald Davidson*.
- ♦ A good collection of critical essays is edited by Lepore and McLaughlin (1985) *Actions and Events*.

Identity theories

- ◆ Kripke's argument about identity is discussed in Boyd's (1980) 'Materialism without reductionism' (in N. Block (ed.) *Readings in Philosophy of Psychology*).
- ◆ A more difficult discussion of Kripke occurs in Macdonald's (1989) *Mind-Body Identity Theories* (pp. 143–155).
- ◆ Debates about identity constitution realization and so forth form a difficult area of metaphysics. A good introduction to such issues is Loux (1998) *Metaphysics*.

Supervenience

- ◆ The locus classicus for discussion of supervenience is Kim's (1993) *Supervenience and Mind*.
- ◆ There are a number of essays on the connection between it and Davidson's Anomalous Monism in Heil and Mele (ed.) (1993) *Mental Causation*.

Applied work

- ◆ For an account of how neuroimaging and analysis of brain function impacts on the mind-body problem see: Christen and Churchland (1992) *Neurophilosophy and Alzheimer's Disease*, and Northoff (1999) 'Psychomotor phenomena, functional brain organization, and the mind-body relationship: Do we need a 'philosophy of the brain'? (*Philosophy, Psychiatry, and Psychology*; with commentaries by Stein, Graham, and Spence, and a reply by Northoff).
- ◆ For a defence of the claim that how mind and body interact in depression should be treated as a medical rather than metaphysical mind-body problem see:
- ◆ Glannon (2002) 'Depression as a mind-body problem' (*Philosophy, Psychiatry, & Psychology*; with commentaries by Martin and Fuchs, and a response by Glannon). Among calls to leave the mind-body split behind, see Bracken and Thomas (2002), and Garnar and Hardcastle (2004).

References

- Block, N. (1980). What is functionalism, and Troubles with Functionalism. In *Readings in Philosophy of Psychology* (ed N. Block) London: Methuen, pp. 171–184, 268–306.
- Block, N. (ed.) (1980). *Readings in Philosophy of Psychology*. London: Methuen.
- Boyd, R. (1980). Materialism without reductionism. In *Readings in Philosophy of Psychology* (ed. N. Block). London: Methuen, pp. 67–106.
- Bracken, P.J. and Thomas, P. (2002). Time to move beyond the mind-body split. *British Medical Journal*, 325: 1433–1434.

- Braddon-Mitchell, D. and Jackson, F. (1996). *Philosophy of Mind and Cognition*. Oxford: Blackwell.
- Burwood, S., Gilbert, P., and Lennon, K. (1999). *Philosophy of Mind*. London: UCL.
- Child, B. (1994). *Interpretation, Causality and the Mind*. Oxford: Oxford University Press.
- Christen, Y. and Churchland, P. (1992) *Neurophilosophy and Alzheimer's Disease*. Heidelberg: Springer-Verlag.
- Crane, T. (1991a). Why indeed: Papineau on supervenience. *Analysis*, 51, 32–37.
- Crane, T. (1991b). All God has to do. *Analysis*, 51, 235–244.
- Davidson, D. (1980). *Essays on Actions and Events*. Oxford: Oxford University Press.
- Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.
- Evnine, S. (1991). *Donald Davidson*. Oxford: Polity Press.
- Feldman, F. (1980). Identity, necessity, and events. In *Readings in the Philosophy of Psychology* (ed. N. Block), Vol. 1. Cambridge, MA: Harvard University Press, pp. 148–155.
- Fuchs, T. (2002). Mind, meaning, and the brain. (Commentary on Glannon, 2002a) *Philosophy, Psychiatry, & Psychology*, 9(3): 261–264.
- Garnar, A. and Hardcastle, V.G. (2004). Neurobiological models: an unnecessary divide—neural models in psychiatry. In *The Philosophy of Psychiatry: a companion* (ed. J. Radden). New York: Oxford University Press, pp. 364–380.
- Glannon, W. (2002a). Depression as a mind-body problem. (Commentaries by Martin, M.W. (2002, 255–260), Fuchs, T. (2002, 261–264), and response by Glannon, W. (2002, 265–270)) *Philosophy, Psychiatry, & Psychology*, 9(3): 243–254.
- Glannon, W. (2002b). The psychology and physiology of depression. (Response to the Commentaries on Glannon, 2002a) *Philosophy, Psychiatry, & Psychology*, 9(3): 265–270.
- Graham, G. (1999). Mind, brain, world. (Commentary on Northoff, 1999a) *Philosophy, Psychiatry, & Psychology*, 6(3): 223–226.
- Halligan, P.W. and Marshall, J.C. (1996). *Method in Madness: case studies in cognitive neuropsychiatry*. Hove, UK: Psychology Press.
- Haugeland, J. (1982). Weak supervenience. *American Philosophical Quarterly*, 19: 93–104.
- Heil, J. and Mele, A. (ed.) (1993). *Mental Causation*. Oxford: Oxford University Press.
- Kim, K. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Kripke, S. (1980). *Naming and Necessity*. Oxford: Blackwell.



- Lepore, E. and McLaughlin, B. (ed.) (1985). *Actions and Events*. Oxford: Blackwell.
- Loar, B. (1981). *Mind and Meaning*. Cambridge: Cambridge University Press.
- Loux, M.J. (1998). *Metaphysics*. London: Routledge.
- Macdonald, C. (1989). *Mind-Body Identity Theories*. London: Routledge.
- Martin, M.W. (2002). On the evolution of depression. (Commentary on Glannon, 2002a) *Philosophy, Psychiatry, & Psychology*, 9(3): 255–260.
- McDowell, J. (1985). Functionalism and anomalous monism. In *Actions and Events* (eds E. Lepore and B. McLaughlin). Oxford: Blackwell, pp. 387–398. Reprinted in McDowell, J. (1998). *Mind Value and Reality*. Cambridge, MA: Harvard University Press, pp. 325–340.
- McGinn, C. (1977). Anomalous monism and Kripke's Cartesian intuitions. *Analysis*, 37: 78–80.
- Northoff, G. (1999a). Psychomotor phenomena, functional brain organization, and the mind-body relationship: Do we need a 'philosophy of the brain'? (Commentaries by Stein, D.J. (1999, 217–222), Graham, G. (1999, 223–226), Spence, S.A. (1999, 227–230), and reply by Northoff, G. (1999, 231–235) *Philosophy, Psychiatry, & Psychology*, 6(3): 199–216.
- Northoff, G. (1999b). Neuropsychiatry, epistemology, and ontology of the brain: a response to the Commentaries on Northoff (1999a). *Philosophy, Psychiatry, & Psychology*, 6(3): 231–235.
- Papineau, D. (1990). Why supervenience. *Analysis*, 50: 66–70.
- Posner, M.I. (1993). Seeing the mind. *Science*, 262: 673–674.
- Rosenthal, D.M. (ed.) (1991). *The Nature of Mind*. Oxford: Oxford University Press.
- Searle, J. (1984). *Minds, Brains, and Science*. Cambridge, MA: Harvard University Press.
- Spence, S.A. (1999). Does a philosophy of the brain tell us anything new about psychomotor disorders? (Commentary on Northoff, 1999a) *Philosophy, Psychiatry, & Psychology*, 6(3): 227–230.
- Stein, D.J. (1999). Philosophy and cognitive neuropsychiatry. (Commentary on Northoff, 1999a) *Philosophy, Psychiatry, & Psychology*, 6(3): 217–222.
- Van Gulick, R. (1994). Deficit studies and the function of phenomenal consciousness. In *Philosophical Psychopathology* (ed. G. Graham and G.L. Stephens). Cambridge, MA: MIT Press, pp. 25–50.