

Brain and Mind

MARTIN DAVIES

History of the Mind-Brain Relation

The thesis that the brain, rather than the heart, is the seat of the mind was already widely accepted by the ancient Greeks; but it was not universally accepted – Aristotle was an exception. Many issues in psychiatry resonate with the ancient debates over the roles of the heart and the brain. But a brief review of modern thinking about the mind-brain relation can begin two millennia later with René Descartes, who held that minds are real things of a fundamentally different kind from material bodies.^[1]

Dualism: Descartes

Descartes's world-view included bodies or material things, whose essence is to be extended in space, and minds, which are immaterial things whose essence is thinking. According to **Cartesian dualism**, the mind is not literally housed within the body, because spatial properties belong to matter and not to mind. But, when he talked about the way we experience the states of our own body, Descartes sometimes spoke of the mind being 'mixed up with' the body.

Early theories of the brain as the seat of the mind assigned an important role to the ventricles. On Descartes's view, mechanical operations involving the release of animal spirits in the ventricles were adequate to explain animal behaviour but intelligent human action required something more. He postulated that the immaterial mind could modulate processes in the material brain by way of a causal interaction operating through the pineal gland.

The motion of bodies and the completeness of physics

Dualist interactionism is challenged by theories about the motion of bodies. According to Descartes's own theory, quantity of motion (defined as mass times speed) is conserved. Because motion is not a directional notion, this conservation law allowed that the immaterial mind could bend the trajectory of a physical particle in the pineal gland. But Gottfried Leibniz's superior theory, with conservation laws for momentum (a directional notion) and kinetic energy, had the consequence that only impacts with other bodies could cause changes in the direction or speed of physical particles. This left no

room for immaterial causes of material changes and, while Leibniz was a dualist, he was not an interactionist dualist but believed, instead, in a **pre-established harmony** between the material and immaterial worlds.

By departing from the idea that impact was the only force on bodies, and allowing action at a distance, Isaac Newton reopened the possibility of distinctively mental forces affecting the trajectory of bodies. These forces were not even ruled out by the law of conservation of energy, which was widely accepted by the middle of the nineteenth century, but advances in biochemistry and neurophysiology during the first half of the twentieth century made appeal to vital and mental forces seem increasingly unmotivated. Since around 1950, the dominant theories of the mind-brain relation have been compatible with a broadly physicalist world-view and with the completeness of physics: physical effects have wholly physical causes.^[2]

Behaviourism: Ryle

From the 1920s to the 1950s, particularly in the United States, behaviourism was a dominant approach within psychology. This was not just methodological behaviourism, which is a restriction on the kinds of evidence that can be used, but a radical reconception of psychology as the science of behaviour rather than the science of the mind. In philosophy, **analytical behaviourism** was a doctrine about the meaning of our mental discourse. The idea was to analyse or translate our mental talk into talk about patterns of behaviour.

Gilbert Ryle promoted behaviourism as a response to what he called ‘Descartes’s myth’ of ‘the ghost in the machine’.^[3] A dualist would regard talk about being in love, or wanting to visit Paris, as talk about an immaterial mind whose states lie hidden behind observable bodily behaviour. Ryle proposed to analyse this mental talk as being about the observable behaviour itself. He did not, however, aim to replace all mental terms by terms appropriate to the science of material bodies moving through space. He analysed believing that the ice is thin as, in part, being ‘prone to skate warily’ and it was enough, for his purposes, that skating warily is an observable and recognisable kind of behaviour, even if it is not readily defined in terms of the trajectories of body parts.

Because action is explained in terms of what the agent believes *and* what the agent wants, analytical behaviourism faces a major objection of principle. There is no pattern of behaviour associated with a belief, by itself. Someone who believes that the ice is thin but has an unusual desire to be immersed in ice-cold water may not skate warily. So there is

no prospect of analysing any belief in terms of behaviour. We might elevate this point into a general requirement on the description of any creature as having beliefs. Attributions of beliefs are not warranted if they merely summarise the creature's dispositions to exhibit patterns of behaviour. A belief is a mental state that can figure in the explanation of indefinitely many different actions in pursuit of different goals.

The Identity Theory: Place and Smart

Ryle's behaviourism involved a clear rejection of Descartes's duality of material and immaterial substances, but **central state materialism** (also known as the **identity theory**) encapsulated a more thoroughgoing commitment to the physicalist world-view. If the physical effects of our experiences, thoughts, and volitions have wholly physical causes then there is no causal work left for distinct mental items to do. To avoid epiphenomenalism, mental states, processes and events were to be identified with physical states, processes and events, and mental properties with physical properties. U.T. Place advanced a precursor of the identity theory, restricted to the case of conscious experiences^[4], and this was generalised by J.J.C. Smart, who identified beliefs and desires, intentions and hopes, as well as sensations and experiences, with brain states or processes.^[5]

The identity theory defends the idea of **mental causation** by identifying each mental state with a physical state that is a locus of causal powers. But, taken literally, the identity theory is bound to seem chauvinistic. No being with a physical constitution radically different from ours could be described as feeling anything, or thinking anything, or wanting anything.

Functionalism: Putnam and Lewis

The functionalist response to the identity theory is that what a system does is more important than what it is made from. Physically different computing machines can run the same software and one version of the functionalist theory of the mind-brain relation is that the mind is the software of the brain.^[6]

In an early version of **functionalism**, Hilary Putnam proposed that mental states are functional states like the states of an abstractly defined Turing machine rather than physical states like the states of a human brain.^[7] This **machine functionalism** had the advantage of not tying mental states to a particular physical substrate but also a

disadvantage. Since a Turing machine is in only one state at a time, machine states are not analogous to mental states like being in love or wanting to visit Paris.

The dominant version of contemporary functionalism, attributable to David Lewis, is **analytical functionalism**.^[8] The leading idea is that commonsense specifications of the interconnected causal roles of mental states can be taken as interlocking analyses of mental state terms. For any physical being with a mind, there will be physical states playing each of the mental state causal roles but different physical states may play the same causal role in physically different minded beings – in humans beings and Martians, for example.

Functionalism thus avoids the apparent chauvinism of the identity theory by allowing that a human being may be in the same mental state as a being with a very different physical constitution. But functionalism faces the opposite problem of apparently being too liberal. It seems to be possible to make up examples in which physical states play the causal roles that are supposed to define mental states, yet where, intuitively, there is no intelligence and no mental life.^[9]

Challenges to Functionalism

The dominant contemporary theories of the mind-brain relation are compatible with a broadly physicalist world-view and analytical functionalism, in particular, is consistent with a version of physicalism, **a priori physicalism**, that is both ontologically and conceptually **reductionist**.^[10]

Ontologically, analytical functionalism is like the identity theory in its commitment to types of physical state that *realise* mental states. Functionalism does not quite say that being in pain is to be identified with having C-fibres firing (the standard example for the identity theory); but it does say that the causal role of the mental state of being in pain is played, in human beings, by the physical state of having C-fibres firing.

Despite this ontological similarity to the identity theory, analytical functionalism is *conceptually* more like analytical behaviourism in being a thesis about the meanings of our mental terms. According to behaviourism, it is a matter of meaning, or **conceptual analysis**, that being in a mental state is being disposed to produce particular patterns of behaviour. According to functionalism, it is equally a matter of meaning that being in a mental state is being in a state that plays a particular causal role. Consequently, analytical functionalism is *conceptually reductionist*. The mental facts, as conceived by the functionalist, are *entailed a priori* by the physical facts.

As we shall now see, both the ontological and the conceptual commitments of analytical functionalism face challenges.

Rylean behaviourism revisited

A theorist of the mind-brain relation who was sympathetic to Rylean behaviourism might challenge the *ontological* commitments that are shared by functionalism and the identity theory. The neo-behaviourist might accept the idea that if a system has a disposition to exhibit a particular pattern of behaviour then there must be a basis for this disposition in the system's inner constitution. But he might argue that identifying individual mental states with physical states, or insisting that mental states are individually realised by physical states, goes beyond what is required by this idea.

Dispositions do not float free of inner constitution and the behavioural dispositions of human beings are, presumably, underpinned by states and processes of the brain. But it is not obviously required that there must be a single brain state that underpins precisely the dispositions that are associated with the attribution of a single mental state. This neo-behaviourism may draw support from remarks made by Ludwig Wittgenstein.^[11]

Neo-behaviourism will be open to objection so long as it retains the unattainable commitment to an analysis of belief attributions in behavioural terms. But there is an alternative view that abandons those analytical ambitions. The **interpretationist** says that mentalistic interpretation is answerable to a creature's behaviour in various actual and hypothetical circumstances, but that this answerability is a matter of 'making sense' of the creature and cannot be codified mental state by mental state.^[12] Rather, the interpreter casts a net of psychological description – 'X is in pain; X is in love; X wants to visit Paris; X believes that the ice is thin; . . .' – over a writhing mass of behaviour. Tracts of human behaviour normally support this interpretive project and, presumably, the behaviour is susceptible of causal explanation. But we should not assume that the physical causes of behaviour must have an articulation that matches the structure of the interpreter's description.^[13]

Interpretationism is compatible with a broadly physicalist world-view but it involves some departure from apparently plausible claims about mental reality and mental causation. The interpretationist is not committed to the claim that there are individual mental states – for example, individual beliefs such as my belief that there is a bottle of white wine in the refrigerator, or that I have an appointment at 9 am – that are bearers of causal powers.

Consciousness and the explanatory gap

The *conceptual* commitments of analytical functionalism are challenged by our conception of **conscious mental states**.

According to functionalism, all mental states are realised by physical, specifically neural, states and the phenomenal properties of conscious mental states are physical properties of those neural states. We can ask what makes the difference between conscious mental states and unconscious mental states. Is there, for example, something distinctive about the neural underpinnings of conscious mental states? If we had a plausible answer to that question, there would be the further question *why* mental states with that distinctive neural nature are *conscious* mental states. This question is apt to seem puzzling and even unanswerable. But, according to analytical functionalism, there would be no puzzling ‘why?’ question about consciousness. All the mental facts, including the facts about consciousness and phenomenology, are entailed a priori by the physical facts.

A powerful intuition thus speaks against the conceptual commitments of analytical functionalism. For it seems that even the full physical story about the world would not settle a priori the question whether a creature was in a conscious mental state. It seems to be conceivable (not ruled out a priori) that there could be a creature physically just like one of us yet lacking consciousness – a *zombie* – or even a complete physical duplicate of our world from which consciousness was totally absent – a *zombie world*.^[14] Between the physical sciences and the facts of consciousness there seems to be an **explanatory gap**.^[15]

Thomas Nagel has drawn attention to a difference between two kinds of conception. Conceptions of conscious mental states are **subjective**; they are accessible from some, but not all, points of view. The conscious mental states that we can *conceive* are limited to relatively modest imaginative extensions from the conscious mental states that we ourselves *undergo*. In contrast, the conceptions deployed in grasping theories in the physical sciences are **objective**; they are accessible from many different points of view. The physical theories that we can grasp are limited, not by our sensory experience, but by our intellectual powers.^[16] Many contemporary philosophers of consciousness argue that the explanatory gap is a product of this duality of conceptions. There is no a priori entailment from the physical and functional facts *objectively conceived* to the phenomenal facts *subjectively conceived*.

The majority of these philosophers maintain that a duality of conceptions does not require an ontological dualism of substances, states, or properties and that the explanatory gap is consistent with physicalism as an ontological doctrine. But there is an important minority view that acceptance of an explanatory gap must lead to a rejection of physicalism. David Chalmers, beginning from the intuition of an explanatory gap, recommends a return to some form of dualism.^[17] Others argue in the opposite direction, embracing physicalism, denying that there is an explanatory gap, and accepting the counterintuitive conceptual reductionism of analytical functionalism.

Personal and Subpersonal Levels of Description and Explanation

The mind-brain relation is an aspect of a more encompassing relationship between persons and the physical systems of which they are constituted, including systems of neural information processing.

Our conception of persons as such is a conception of subjects and agents. At the personal level of description and explanation, we describe what people feel, think, want and do, and we explain what people do in terms of their sensations, beliefs and desires. As the case of conscious mental states illustrates, our personal-level *descriptions* are not always entailed a priori by physical and functional descriptions of the systems that constitute us. Personal-level descriptions involve subjective and normative concepts that are different from the objective and descriptive concepts that figure in the physical sciences.

Our personal-level *explanatory practices* seem to be different in kind from our scientific practices of explaining the operation of mechanical systems. John McDowell describes personal-level explanations as ‘explanations in which things are made intelligible by being revealed to be, or to approximate to being, as they rationally ought to be’.^[18] In a similar spirit, Daniel Dennett describes them as ‘non-mechanistic’. A mechanistic account of what happens when a person feels, thinks, wants and acts would belong at a quite different level of description and explanation, not the ‘level of people and their sensations and activities’, but ‘the *sub-personal* level of brains and events in the nervous system’.^[19]

One extreme view of the relationship between the personal and subpersonal levels highlights what is distinctive about the personal level and regards it as substantially independent from the subpersonal level. This view might encourage the interpretationist account of personal-level psychological descriptions, minimising the ontological and

causal commitments of personal-level discourse to avoid constraints on that discourse from the subpersonal level of neuroscience.

The opposite extreme view is the conceptually reductionist view of analytical functionalism. The personal level is the level of mental states whose causal roles are revealed by conceptual analysis while the subpersonal level is the level of neural states that play those roles. There are no explanatory gaps. All that is true at the personal level is entailed a priori by physical truths at the subpersonal level.

According to an attractive view that is intermediate between these two extremes, the relationship between the personal and subpersonal levels is one of *interaction without reduction*.^[20] As against the first extreme view, the personal level is not independent of the subpersonal level but constrained by it, because our personal-level descriptions – cast in terms of experience, thought, planning and agency – carry commitments about causal structure in the brain. But, as against the second extreme view, there are also explanatory gaps that reveal themselves when we try to construct illuminating accounts of those personal-level notions using only the subpersonal-level resources of neuroscience.

Conclusion

Descartes's ontological dualism of mind and body made it difficult for him to describe the phenomenology of embodiment, the way we experience our own body. Contemporary theories of the mind-brain relation are predominantly physicalist, rather than dualist, in their ontology. But the duality of objective and subjective conceptions still presents a challenge for the sciences of the mind. Persons understood as such, partly from the first-person perspective – persons conceived as subjects and agents, with their experiences, thoughts, plans and actions – will not be visible in a purely objective, scientific story of the physical world.

Further Reading

Anthology

Chalmers, D.J. (ed.) (2002). *Philosophy of Mind: Classical and Contemporary Readings*. Oxford: Oxford University Press.

The items marked with an asterisk on the list of references below are reprinted, in whole or in part, in this anthology.

Textbook

Braddon-Mitchell, D. and Jackson, F.C. (2007). *Philosophy of Mind and Cognition: An Introduction* (Second Edition). Oxford: Blackwell Publishing Ltd.

References

- [1] *Descartes, R. (1641). *Meditations on First Philosophy*. In Descartes, R. (1988). *Selected Philosophical Writings* (trans. and eds. J. Cottingham, R. Stoothoff and D. Murdoch). Cambridge, UK: Cambridge University Press.
- [2] Papineau, D. (2002). The history of the completeness of physics. In *Thinking About Consciousness*, Appendix. Oxford: Oxford University Press.
- [3] *Ryle, G. (1949). Descartes' myth. In *The Concept of Mind*, Chapter 1. Harmondsworth: Penguin Books.
- [4] *Place, U.T. (1956). Is consciousness a brain process? *British Journal of Psychology*, **47**, 44–50.
- [5] *Smart, J.C.C. (1959). Sensations and brain processes. *Philosophical Review*, **68**, 141–56.
- [6] Block, N. (1995). The mind as the software of the brain. In E.E. Smith and D.N. Osherson (eds), *An Invitation to Cognitive Science, Volume 3: Thinking* (Second Edition), 377–425. Cambridge, MA: MIT Press.
- [7] *Putnam, H. (1967). The nature of mental states. Originally titled 'Psychological predicates', in W.H. Capitan and D.D. Merrill (eds), *Art, Mind, and Religion*, 37–48. Pittsburgh, PA: University of Pittsburgh Press.
- [8] *Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, **50**, 249–58.

- [9] *Block, N. (1978). Troubles with functionalism. In C. Wade Savage (ed.), *Minnesota Studies in the Philosophy of Science, Volume 9*, 261–325. Minneapolis, MN: University of Minnesota Press.
- [10] Jackson, F.C. (2005). The case for a priori physicalism. In C. Nimtz and A. Beckermann (eds), *Philosophy – Science – Scientific Philosophy: Main Lectures and Colloquia of GAP.5, Fifth International Congress of the Society for Analytical Philosophy, Bielefeld, 22–26 September 2003*, 251–65. Paderborn: Mentis.
- [11] Wittgenstein, L. (1981). *Zettel* (eds. G.E.M. Anscombe and G.H. von Wright; trans. G.E.M. Anscombe), Sections 608–9. Oxford: Basil Blackwell.
- [12] Davidson, D. (1973). Radical interpretation. *Dialectica*, **27**, 313–28. Reprinted in D. Davidson, *Inquiries into Truth and Interpretation*, 125–39. Oxford: Oxford University Press, 1984.
- [13] *Dennett, D.C. (1987). True believers: The intentional strategy and why it works. In *The Intentional Stance*, 13–35. Cambridge, MA: MIT Press.
- [14] Chalmers, D.J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- [15] *Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly*, **64**, 354–61.
- [16] *Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, **83**, 435–50. Reprinted in T. Nagel, *Mortal Questions*, 165–80. Cambridge, UK: Cambridge University Press, 1979.
- [17] Chalmers, D.J. (2007). Naturalistic dualism. In M. Velmans and S. Schneider (eds), *The Blackwell Companion to Consciousness*, 359–68. Oxford: Blackwell Publishing.
- [18] McDowell, J. (1985). Functionalism and anomalous monism. In E. LePore and B.P. McLaughlin (eds), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, 387–98; at p. 389. Oxford: Basil Blackwell.
- [19] Dennett, D.C. (1969). *Content and Consciousness*, at pp. 91–4. London: Routledge and Kegan Paul.
- [20] Davies, M. (2000). Interaction without reduction: The relationship between personal and sub-personal levels of description. *Mind and Society*, **1**, 87–105.