Consciousness

Jon Opie*

School of Humanities, University of Adelaide, SA, Australia
*Correspondence: jon.opie@adelaide.edu.au

Understanding consciousness and its place in the natural world is one of the principal targets of contemporary philosophy of mind. Australian philosophers made seminal contributions to this project during the twentieth century which continue to shape the way philosophers and scientists think about the conceptual, metaphysical and empirical aspects of the problem. After some scene setting, I will discuss the main players and their work in the context of broader developments in the philosophy of mind.

Towards the end of the nineteenth century, scientific psychology set itself the task of systematically exploring the mind, understood as the conscious activity that accompanies perception and thought. Labs in Germany and the United States began the tedious work of determining the structure of experience via the reports of trained subjects operating under carefully controlled stimulus conditions. The hope was that the phenomena revealed by this means might eventually be correlated with activity in the central nervous system.

Many philosophers considered this project misguided. The logical positivists, who insisted that a statement is only meaningful if one can specify observable conditions that would render it true or false, rejected the view that psychological predicates such as 'pain' have any subjective content. A statement like 'Paul has a toothache' is merely an abbreviation for a list of physical events (such as Paul weeping, Paul's blood pressure rising, etc.) which collectively exhaust the meaning of the statement (Hempel 1980).

Ryle (1949) and Wittgenstein (1953) regarded the so called 'mind-body problem' as the result of a misuse of ordinary language. According to Ryle, it is a "category mistake" (1949: 16) to treat the mind as part of the body, because psychological and physical language follow different rules. The former provides a mentalistic short-hand for characterising behavioural dispositions, but does not pick out the internal causes of behaviour (unlike physiology).

It was in this climate that Ullin Place and Jack Smart, both working at the University of Adelaide, first proposed their pioneering idea that conscious states and processes are none other than states and processes of the brain: the so-called 'identity theory'. In philosophy circles this was widely regarded as an outlandish proposal. One English philosopher is said to have reacted: "A touch of the sun, I suppose" (reported in Armstrong 1993: xiii). Place, who first proposed the theory, thought the dispositional analysis of mental concepts such as 'believing' and 'intending' was sound, but claimed that there is "an intractable residue of concepts clustering around the notions of consciousness, experience, sensation and mental imagery, where some kind of inner process story is unavoidable" (1956: 44). He emphasised that the identity theory is not an analysis of statements about sensations into statements about the brain, but a defeasible scientific hypothesis (*ibid.*, p.45).

Smart, initially a skeptic (see his 2008), soon came to the theory's defence. Following Place, he compared the identity of sensations and brain processes to the relationship between lightning and electrical discharges. The latter is not a matter of definition; one can understand statements about lightning without any knowledge of electricity. Nor is it a matter of lightning and electrical discharges being contemporaneous and co-spatial (as when two gases are mixed). Rather, it is an "identity in the *strict* sense" (Smart 1959: 145). Lightning is nothing more than an electrical discharge. Likewise, the identity theory asserts that sensations are not merely *correlated* with brain processes, as psychologists had supposed; they are one and the same thing.

To the objection that we attribute different kinds of properties to experiences and brain processes — sensations are private, brain processes are public; sensations can be intense or unpleasant, brain processes cannot — Smart responded that our linguistic conventions are not

fixed, but will undoubtedly change with future science. We may one day be able to state objective physiological criteria for the application of expressions such as 'Smith experiences a strong sweet taste', and will have no qualms about describing experience in physical terms (*ibid.*, pp.152-53).

To the objection that 'raw feels' such as the yellow of a lemon, or the sweetness of sugar, are irreducibly mental, Smart offered an account of colour, taste, etc., as "powers…to evoke certain kinds of discriminatory responses in human beings" (*ibid.*, p.149). Such powers belong to the objects we perceive, not to our sensations. Thus, in describing sugar as "sweet" we are not referring to a non-physical quality of a sensation, but to a power of sugar to produce certain effects in us. After-images are a problem here, since they have no object. However, Smart noted that a report such as 'I see a yellow after-image' can easily be expressed in a *topic-neutral* way (i.e., in terms that are neutral between materialism and dualism), for example: 'Something is going on with me that is like what goes on when I look at a lemon' (*ibid.*, pp.148-50).

An important player in these early developments was Charlie Martin, who was at the University of Adelaide between 1954 and 1966. Although Martin did not publish a great deal at the time, his influence is frequently acknowledged by Place and Smart (see Place 1989 for an account of Martin's contribution to "Is Consciousness a Brain Process?", and Martin 2007 for an overview of his distinctive approach to dispositions, emergence, and mind).

David Armstrong (1968, 1977) and Brian Medlin (1967) were part of a second wave of Australian identity theorists. They extended the theory by offering a causal analysis of mental states as "[states] of the person apt for bringing about a certain sort of behaviour" or "apt for being brought about by a certain sort of stimulus" (Armstrong 1968: 82). A desire for food, for example, is a state of a person that typically produces food-seeking and food-consuming behaviour; a belief that it is raining is a state that is typically caused by rainfall, and so on. Both Armstrong and Medlin argued that the states which play these roles in us are states of the brain. Their account, known as 'central-state materialism' (Feigl 1967), differs from the Place/Smart theory in that it identifies *all* mental states, not just conscious ones, with brain states.

Consciousness appears in two guises in Armstrong's work: as introspective awareness, and as the qualities of sensations and mental images. Armstrong regards introspection as analogous to perception, but whereas perception informs us about objects in the physical environment, introspection is a kind of inner sense whereby we acquire information about our own mental states. It produces these special states of awareness via some kind of self-scanning process in the brain (1968: 323-38). As remarked above, perception presents us with the problem of *qualia*, Locke's secondary qualities. Like Smart, Armstrong argued that these apparently qualitative features of experience do not belong to conscious states at all, which are "transparent" (1993: xxii). What our perceptual states reveal, when they don't deceive us, is certain complex micro-physical properties of external objects (1968: 270-90).

The causal analysis of mind, independently worked out by Lewis (1966), contributed to the development of functionalism (Putnam 1967). Functionalism identifies a mental state with a *causal role*: the kinds of stimuli that produce it, the kinds of behaviour it produces, and the way it interacts with other mental states. One of the advertised strengths of functionalism is that it imposes very few limits on the nature of the physical states that can play such roles, thus allowing for mentality in organisms, or even machines, that are physically very different from us. Armstrong originally held that a given *type* of mental state is identical to a corresponding *type* of brain state (1993: xv). This 'type-type' theory is vulnerable to the possibility of mental states that are realised by more than one kind of brain state. Functionalism only insists that a mental state should be realised by some physical state or other — mental states are *multiply realisable*.

Despite its advantages, functionalism has come in for some serious flack. Frank Jackson, a self-confessed "qualia freak" (1982: 127), is among a number of philosophers who have expressed dissatisfaction with physicalism in both its type-type and functionalist forms. The problem, as Jackson sees it, is that no amount of physical information about the structure, function, or causal history of brain states can capture the phenomenal qualities of experience. Consequently, physicalism must be false. In support of this claim Jackson asks us to imagine a neuroscientist,

Mary, who has spent her whole life in a black and white room, her only access to the outside world provided by a black and white monitor. By assumption, Mary knows everything there is to know about the neurophysiology of colour vision, despite never having seen a coloured object. What will happen when she exits her room? Jackson takes it to be obvious that Mary will learn something new about visual experience, and thus that physicalism leaves something out (*ibid.*, p.130). This 'knowledge argument' has generated a great deal of critical reaction (see, e.g., Churchland 1985, Lewis 1988). For his part, Jackson no longer buys the conclusion of the argument, on the grounds that it is self-defeating: If Mary learns something new, then qualia are non-physical; but if qualia are non-physical then they are causally inert, and can't influence our beliefs; so Mary's new qualia can't possibly lead her or us to conclude that qualia are non-physical (Braddon-Mitchell and Jackson 1996: 134).

David Chalmers is another Australian philosopher who takes issue with earlier treatments of qualia. Chalmers divides the mystery of consciousness into an 'easy problem' and a 'hard problem'. The easy problem is to explain how the brain processes stimuli, integrates information, and reports on our internal states. The hard problem is the further question: "Why is all this processing accompanied by an experienced inner life?" (1996: xii). Chalmers believes that the hard problem goes beyond what can be explained in terms of the structural and dynamical properties of physical processes, because "the existence of my conscious experience is not logically entailed by my functional organization" (*ibid.*, p.97). We can see this, he claims, by recognising the conceptual possibility of *phenomenal zombies*: creatures that are physically and functionally identical to us, but which lack experience. Although phenomenal zombies may not exist in our world, "the mere intelligibility of the notion is enough to establish the conclusion" (*ibid.*, p.96). Chalmers advocates what he calls 'naturalistic dualism' according to which conscious experience has phenomenal (or proto-phenomenal) properties that are not entailed by physical properties, but which may be lawfully related to those properties (*ibid.*, pp.124-29).

Jackson and Chalmers offer modal arguments for their anti-physicalist positions: they argue from certain possibilities — the existence of phenomenal zombies; that an omniscient scientist might know all the physical facts, yet learn something new via experience — to conclusions about experience. Daniel Stoljar has recently developed a general response to this style of argument. He defends what he calls 'the ignorance hypothesis' (2006: 6), the claim that we are ignorant of a type of non-experiential fact that is relevant to the nature of experience. Such ignorance undermines our capacity to imagine the scenarios described by Chalmers and Jackson (ibid., pp.67-86). One simply can't imagine a phenomenal zombie, for example, if one is not in possession of all the relevant facts. One can imagine an organism that lacks experience even though it is identical to us in all the non-experiential respects we know about. But this is no zombie, because our imagining has perforce omitted some of the physical facts (those of which we are ignorant). Although Stoljar doesn't offer a positive account of such facts, he argues for the plausibility of the ignorance hypothesis on the basis of historical precedent and general observations about our epistemic situation (ibid., pp.87-141).

O'Brien and Opie (1999), swimming against the functionalist tide, defend a connectionist approach to consciousness. Their *connectionist vehicle theory* identifies conscious states with stable patterns of firing in the brain's many neural networks. Connectionists argue that such firing patterns are a crucial class of representing vehicles in the brain, whose interactions are the causal basis of cognition (Rumelhart, McClelland and the PDP Research Group 1987). O'Brien and Opie's account thus does justice to both the representational role of consciousness and its causal impact on behaviour. It is an identity theory in the original sense, because it identifies phenomenal consciousness with a particular type of neural activity. The theory explains consciousness not in terms of what the brain's representing vehicles *do*, as a functionalist would, but in terms of *what they are* (1999: 138). The prospects for a vehicle theory of consciousness depend on the ability of disciplines such as cognitive neuroscience and psychophysics to establish a detailed mapping between the rich, multi-layered structure of consciousness, and the equally rich, multi-level organisation of neural activity, as envisaged by early experimental psychologists.

(Thanks to Greg O'Hair and Philip Gerrans for a number of helpful comments and suggestions.)

References

Armstrong, D. (1968) *A Materialist Theory of the Mind*. New York: Routledge and Kegan Paul.

Armstrong, D. (1993) *A Materialist Theory of the Mind*, revised edition. London: Routledge.

Armstrong, D. (1977) "The Causal Theory of the Mind", *Neue Hefte für Philosophie*, nr. 11. Gottingen: Vendenhoeck and Ruprecht.

Braddon-Mitchell, D. and Jackson, F. (1996) *Philosophy of Mind and Cognition*. Oxford: Blackwell.

Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. New York: Oxford University Press.

Churchland, P. (1985) "Reduction, Qualia, and the Direct Introspection of Brain States", *Journal of Philosophy* 82: 8-28.

Feigl, H. (1967) *The 'Mental' and the 'Physical': The Essay and a Postscript* Minneapolis: University of Minneapolis Press.

Hempel, C. (1980) "The Logical Analysis of Psychology". In N. Block (ed.), *Readings in the Philosophy of Psychology, Volume 1*. Cambridge, MA: Harvard University Press, 14-23.

Jackson, F. (1982) "Epiphenomenal Qualia", *Philosophical Quarterly* 32: 127-36.

Lewis, D. (1966) "An Argument for the Identity Theory", *Journal of Philosophy* 63: 17-25.

Lewis, D. (1988) "What Experience Teaches", *Proceedings of the Russellian Society* 13: 29-57.

Martin, C. (2007) *The Mind in Nature*. Oxford: Clarendon Press.

O'Brien, G. and Opie, J. (1999) "A Connectionist Theory of Phenomenal Experience", *Behavioral and Brain Sciences* 22: 127-48.

Place, U. (1956) "Is Consciousness a Brain Process?", *British Journal of Psychology* 47: 44-50.

Place, U. (1989) "Low Claim Assertions". In J. Heil (ed.), *Cause, Mind and Reality: Essays Honoring C. B. Martin*. Dordrecht: Kluwer Academic Publishers, 121-135.

Putnam, H. (1967) "Psychological Predicates". In W. Capitan and D. Merrill (eds), *Art, Mind and Religion*. Pittsburgh: University of Pittsburgh Press, 37-48.

Rumelhart, D., McClelland, J., and the PDP Research Group (1987) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, vol. 1: Foundations*. Cambridge, MA: MIT Press.

Medlin, B. (1967) "Ryle and the Mechanical Hypothesis". In C. Presley (ed.), *The Identity Theory of Mind*. St. Lucia: Queensland University Press, 94-150.

Ryle, G. (1949) *The Concept of Mind*. New York: Barnes and Noble.

Smart, J. (1959) "Sensations and Brain Processes", *Philosophical Review* 68: 141-56.

Smart, J. (2008) "The Identity Theory of Mind". In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/archives/fall2008/entries/mind-identity/.

Stoljar, D. (2006) Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness. Oxford: Oxford University Press.