mark.sprevak@ed.ac.uk

Realism and instrumentalism

Mark Sprevak
University of Edinburgh

20 December 2009

The choice between realism and instrumentalism is at the core of concerns about how our scientific models relate to reality: Do our models aim to be literally true descriptions of reality, or is their role only to be useful instruments for generating predictions? Realism about *X*, roughly speaking, is the claim that *X* exists and has its nature independent of our interests, attitudes, and beliefs. An instrumentalist about *X* denies this. She claims that talk of *X* should be understood as no more than a locution for generating predictions; such talk should not be understood as taking on a commitment to the existence of *X*. According to an instrumentalist, we should either flatly not believe that *X* is out there, or else suspend judgement about the existence of *X*. The most we need acknowledge is that talk of *X* is useful in making predictions.

The question of realism vs. instrumentalism can be asked about almost any theoretical entity in science. It is likely, and seems reasonable, that different answers will be given in different cases. Someone may wish to be a realist about certain theoretical entities (e.g. *electrons*), but an instrumentalist about others (e.g. *centres of gravity*). Not every noun-phrase in a scientific theory should be taken as expressing an ontological commitment. Psychological theories are no exception. Almost every theoretical entity posited by psychology has been questioned as to whether it is *really out there* or just a *useful theoretical fiction*. In this entry, I will focus on two major theoretical entities posits: (a) *propositional attitudes* (e.g. *beliefs, desires*) and (b) *conscious states* (*qualia*).

1 Propositional attitudes

Psychological theories, both in science and our folk conceptions, often use propositional attitudes (*beliefs*, *desires*, *hopes*, *assumptions*, *fears*, etc.) to explain and predict how people think. These mental representations seem to figure as causal agents in our best explanations of how agents behave and reason (e.g. 'If one believes x and desires y, that causes one to z…'). Should psychological theories using propositional attitude terms be taken at face value: as referring to concrete entities that have an objective existence, that cause action, combine with one another, are caused by sensation, and so on? Or should talk of propositional attitudes be understood as no more than a theoretical fiction that allows our psychological theories to achieve their predictive success, but does not correspond to entities that are really out there?

Jerry Fodor argues for a robust form of realism about propositional attitudes. His reasoning is based on the empirical success of psychological theories that employ propositional attitude terms. We do astonishingly well at predicting how people behave if we are allowed to talk in terms of beliefs, desires, and other propositional attitudes. There are no rival accounts of human psychology that enjoy similar predictive success and that do not make use of propositional attitudes. According to Fodor, there is therefore at least a presumptive inference that talk of beliefs and desires latches onto real entities that pull the strings behind our behaviour. Just as the success of our physical theories gives us reason to infer that their theoretical terms refer to real entities (e.g. *electrons*), so the success of our psychological theories gives us reason to infer that their propositional attitude terms latch onto objective features of the world.

If propositional attitudes are real, what sort of entities are they? According to Fodor, in order adequately to account for the explanatory success of psychology, one has to understand propositional attitudes as having a sentence-like structure. Fodor consequently posits a language of thought in which beliefs, desires, etc. appear as sentences. Just as expressions in a computer's machine code control a machine's behaviour, and cause the occurrence of further expressions of machine code inside the machine, so sentences in our language of thought enter into causal relations, control our behaviour, and cause the occurrence of new sentences in our language of thought (new beliefs, desires, etc.). Like a computer's machine code, sentences in our language of thought exist as a pattern of physical activity inside our heads. Hence, propositional attitudes such as beliefs and desires are discrete, reoccurring, entities with causal powers. They are the causal agents behind our behaviour, just as patterns of electrical activity instantiating a computer's machine code instructions are the causal agents behind a computer's behaviour.

In contrast, Paul Churchland argues for a robust form of instrumentalism about

propositional attitudes. According to Churchland, belief and desire terms fail to latch onto any entities in the world, and at best serve as a useful way of talking for ordinary folk. Churchland concedes that psychological theories employing propositional attitudes enjoy some predictive success, however he thinks that Fodor overestimates the degree of that success in light of the potential of future neuroscientific theories to explain behaviour without reliance on propositional attitudes. Churchland points out that many theories are instrumentally useful yet false. Ptolemaic astronomy, which posited celestial spheres, makes many true predictions, but is nevertheless false. Churchland claims that beliefs and desires will go the way of celestial spheres. Churchland's main argument for the non-existence of propositional attitudes can be broken into two steps. The first step is to argue that folk psychology is a theory, and that the meaning of propositional attitude terms (expressions like 'belief' and 'desire') is fixed by their role in that theory. What 'belief' and 'desire' mean is wholly, and exclusively, specified by folk psychology: what it means to be a *desire* is to be something that combines with beliefs and causes action in precisely the way described by folk psychology. Churchland's second step is to argue that folk psychology is false. If the folk psychology theory is false, then *nothing* satisfies the role ascribed to beliefs and desires, and consequently beliefs and desires, as traditionally conceived, do not exist. Both steps in Churchland's argument have been questioned. Against the first, Ronald Mallon and colleagues have argued that it is far from clear the extent to which the meaning of propositional attitude terms like 'belief' and 'desire' ride on the fortunes of folk psychology. It is not obvious how much, or indeed if any, of folk psychology needs to be true in order to fix the meanings of propositional attitude terms. Against the second step, Terence Horgan and James Woodward have argued that Churchland underestimates the success of folk psychology and overestimates the demands we should place in order to be justified in believing it is true.

There are many ways of developing the instrumentalist thought. Daniel Dennett offers a milder form of instrumentalism about propositional attitudes than Churchland. On Dennett's view (unlike Churchland's), talk of beliefs and desires is *true*, but (unlike Fodor's) such talk does not succeed in referring to entities that have an objective existence or representational content independent of our interests. According to Dennett, what is involved in an agent having a belief or desire is not that there is a discrete entity inside the agent—*the belief that p*—with causal powers pulling the strings behind the agent's behaviour, but merely that there is a predictive pay-off in describing the agent *as if* it were controlled by such an entity. To describe an agent in terms of propositional attitudes is to adopt what Dennett calls the 'intentional stance': a mode of explanation that attributes to the agent the beliefs and desires that a rational being placed in its shoes *ought* to have. According to Dennett, if the intentional stance is reliable as a method of predicting the behaviour of a system

S, then *ipso facto* system *S* has those beliefs and desires. All that is required for a system to have a belief is for it to be useful in predicting the behaviour of that system to assume that it has that belief. A consequence of Dennett's instrumentalism is that beliefs are easy to achieve. It is often helpful to predict the behaviour of cats, robots, washing machines, computers, plants, bacteria, cars, and thermostats by treating them *as if* they have beliefs and desires. According to Dennett, there is no difference between this *as if* and genuine possession of beliefs. It also becomes harder for some systems to achieve beliefs on Dennett's view. Patients suffering from mental illness often cannot be profitably viewed as rational agents when it comes to predicting their behaviour. Therefore, on Dennett's view those patients lack beliefs and desires. Their behaviour would have to be explained in some other way than intentional psychology: for example, by dropping down to the level of their neurophysiology.

Dennett's instrumentalism about propositional attitudes raises questions about exactly how one should draw the line between realism and instrumentalism. The simple characterisation of realism and instrumentalism given at the beginning of this entry fragments into a number of different theses that can, in principle, be affirmed or denied separately. Dennett denies two key realist theses about propositional attitudes: (i) *Mind-Independence*: propositional attitudes of an agent have their existence and nature independent of the interpreting interests of observers; (ii) *Discrete Causal Powers*: propositional attitudes are discrete, reoccurring, entities inside the head with the causal powers to produce behaviour. However in contrast to Churchland, Dennett affirms a realist intuition about propositional attitudes: (iii) *Existence*: propositional attitudes exist—they are 'really out there', unlike celestial spheres. According to Dennett, propositional attitudes exist as patterns that are available to an interpreter to be used for prediction. These patterns are 'real' and 'objective' in the sense that there are objective facts about what is and what is not predictively successful to assume within the intentional stance. In other words, some belief and desire attributions pay off in that they yield successful predictions of behaviour, and others do not. These real patterns of predictive success and failure are the facts in the world that make claims about propositional attitudes true or false.

2   Qualia

We often report that there is a phenomenal aspect to our experience: seeing red *feels* a certain way, having a mouse cupped in one's hand *feels* a certain way, and the qualitative aspects of different experiences are different. Distinctive qualitative experiences accompany large parts of our mental life. Do our reports of qualitative

experiences describe really existing entities (phenomenal properties or states inside our head), or do they serve some other purpose? Are qualia real, or does talk of qualia fail to refer to anything in the world?

An immediate problem that realism and instrumentalism about qualia face is that it is hard even to describe what qualia *are*, and hence hard to say what one is or is not being a realist about. Often the best one can do is point to examples of qualia, such as those above. Daniel Dennett develops a strong instrumentalist line against qualia. His target is the widespread assumption that qualia have certain special properties: they are *ineffable*, *intrinsic properties of experience*, *private* and *directly accessible to the experiencer*. Dennett argues that nothing satisfies this specification, and hence there are no such things as qualia. Dennett's position is similar to Churchland's strong line against propositional attitudes: just as the falsity of Ptolemaic astronomy justifies the inference that there are no celestial spheres, so the falsity of philosophical claims about qualia justifies the inference that there are no qualia. Talk of qualia still serves a purpose according to Dennett in that it provides a shorthand summarising our ability to detect certain properties in the world, such as colour properties, that lack a compact description in any other terms.

Dennett's instrumentalism has drawn heavy criticism, not least because it runs up against the robust impression that there *are* real qualitative properties of our mental states that have at least some of the properties mentioned above. However, realism comes at a price. If one grants realism about qualia, then the question arises of what kinds of entities qualia are. Are qualia represented features of the world encoded by our nervous system, similar to our unconscious encoding of features of the world like position, size, and shape information? If so, what makes conscious 'felt' representations different from unconscious representations? Or are qualia intrinsic physical properties of our nervous system, independent of our ability to represent? Or are qualia something different entirely, requiring properties that float free from the physical world and any representations it encodes? No consensus currently exists to answer this question of the nature of qualia under a realist understanding—the 'hard question' of consciousness.

Bibliography

Chalmers, D. J. (1996). *The Conscious Mind*. Oxford: Oxford University Press.

Churchland, P. M. (1981). "Eliminative materialism and the propositional attitudes". In: *The Journal of Philosophy* 78, pp. 67–90.

Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

Dennett, D. C. (1991). "Real patterns". In: *The Journal of Philosophy* 88, pp. 27–51.

— (2002). "Quining qualia". In: *Philosophy of Mind: Classical and Contemporary Readings*. Ed. by D. J. Chalmers. Oxford: Oxford University Press. Chap. 26.

Fodor, J. A. (1975). *The Language of Thought*. Sussex: The Harvester Press.

Horgan, T. and J. Woodward (1985). "Folk psychology is here to stay". In: *Philosophical Review* 94, pp. 197–226.

Mallon, R. et al. (2009). "Against arguments from reference". In: *Philosophy and Phenomenological Research* 79, pp. 332–356.

Putnam, H. (1975). *Mathematics, Matter and Method, Philosophical Papers, volume 1*. Cambridge: Cambridge University Press.

Tye, M. (1995). *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.