

EPIPHENOMENALISM

Keith Campbell and Nicholas J.J. Smith

December 1993

Written for the *Routledge Encyclopedia of Philosophy*.

Epiphenomenalism is a theory concerning the relation between the mental and physical realms, regarded as radically different in nature. The theory holds that only physical states have causal power, and that mental states are completely dependent on them. The mental realm, for epiphenomenalists, is nothing more than a series of conscious states which signify the occurrence of states of the nervous system, but which play no causal role. For example, my feeling sleepy does not cause my yawning — rather, both the feeling and the yawning are effects of an underlying neural state.

Mental states are real, and in being conscious we are more than merely physical organisms. Nevertheless, all our experiences, thoughts and actions are determined by our physical natures. Mental states are actually as smoke from a machine seems to be, mere side effects making no difference to the course of Nature.

1 Epiphenomenalism

Epiphenomenalism is a version of dualism, rejecting reduction of the mental to the physical (see Dualism, Reductionism in the philosophy of mind). Unlike other dualist theories, however, it denies that conscious mental states are ever causes. It is never pain that makes us wince, nor anger that makes us shout. Remembering our childhood plays no part in the writing of our

memoirs. What has caused philosophers to propose a theory which is such an affront to common sense?

On the one hand, the rise early in the seventeenth century of the conception of the physical realm as a closed system, in which the forces of material nature are the only influences that determine the course of events, when combined with the naturalistic view that human beings are a part of material nature, and governed by its laws, seems to leave no room for a realm of mental states having a role in fixing the course of events. With the demise of vitalism (see Vitalism) regarding the forces governing animate life, the case for the physical causal closure of the material realm has seemed compelling.

On the other hand, philosophers have held that it is a fact of experience that we *do* enjoy conscious states, whose features are incompatible with a purely physical nature. The appeal of epiphenomenalism lies in its capacity to resolve this dilemma.

The theory can, as in its classical form, be applied to all mental states. Other versions can admit physical, and effective, subconscious mental states. Even among conscious states, a theory can be epiphenomenalist about some, such as the phenomenal qualia of sensory awareness (see Qualia), and yet reductively materialist regarding memories or thoughts.

The term *epiphenomenon*— meaning a secondary symptom — was first applied to consciousness in 1890 by William James, but the position which he was attacking had already existed for some time.

Simmiias, in Plato's *Phaedo*, asserts that body stands to mind as a musical instrument stands to its *harmonia* (85e3 —86d4). If we interpret the latter as meaning the music produced by an instrument, Simmiias's theory has epiphenomenalist overtones.

In the eighteenth century Charles Bonnet discussed in his *Essai de Psychologie* (1735) a theory according to which the soul is a mere spectator of the movements of its body though it believes itself to be the author of them while the body performs of itself all that series of actions which constitutes life

In 1865 Shadworth Hodgson's *Time and Space* provided the first full formulation of epiphenomenalism. He wrote, "States of consciousness are not produced by previous states of consciousness, but both are produced by the action of the brain; and, conversely, there is no ground for saying that states of consciousness react upon the brain or modify its action." In 1870 Hodgson became epiphenomenalism's first explicit supporter (in *Theory of Practice*). Thomas Huxley soon followed; and his 1874 essay *On the Hypothesis that Animals are Automata, and its History* with its famous phrase "We are conscious automata" is the classic statement of the theory.

In the twentieth century epiphenomenalism was not widely supported, although George Santayana (1905) and C.D. Broad (1925) both have epiphenomenalist leanings, and John Lachs (1963) vigorously defended the theory. In 1970 Keith Campbell proposed a *New Epiphenomenalism* which combines aspects of epiphenomenalism with the view that mental states are brain states (see *Mind, identity theory of*). Frank Jackson later defended a similar view. Where classical epiphenomenalism asserts that mental states are non-physical and causally inert, the new epiphenomenalism asserts that mental states are causally potent physical states of the

brain, but that in addition to their physical properties some of these states possess phenomenal properties or *qualia* which are non-physical and non-causal.

For epiphenomenalism to be a doctrine distinct from both dualism and materialism, it must involve a very strong conception of causality as productive power or efficacy. No Humean or regularity theory of causation will be sufficient. For the epiphenomenalist admits that many conscious states are regularly followed by other conscious states or by actions, yet denies that the former ever cause the latter.

It will not do, furthermore, to move from mere de facto regularities to *necessary* sequences. For conscious states figure in necessary sequences just as much as neural ones do. If necessary sequence is necessary and sufficient for the causal link, then the neural states will be followed, of necessity, both by the conscious states and by the actions to which the neural states give rise. So the conscious states will also be linked necessarily with the actions, and so will be causes of the actions.

It will not be sufficient, either, to offer a merely counterfactual account of the causal link. Provided the conscious states are effects of the neural states alone, if the conscious states had not existed, nor would the neural states. Then the actions would not have existed either. So if the conscious states had not existed, nor would the actions, and so once again the conscious states count as causes of the actions.

The epiphenomenalist needs to insist that the sequences, and necessary sequences, and counterfactual linkages are indeed there, connecting conscious states with actions, but that these links are not present *on account of any causal activity in the conscious states*. Only insofar as

such a strong conception of causation can be sustained, can epiphenomenalism be maintained as a distinctive position in the theory of mind.

2 The principal objections to epiphenomenalism

The evidence problem. The standard way in which we obtain evidence for the existence and nature of something is from its effects on the mind, either directly, or through traces. If epiphenomenalism is true, then conscious states leave no effects, not even on the mind in which they occur. How then can we ever have any reason to believe that we are, or were, in any conscious state?

The epiphenomenalist can appeal here to indirect causal chains. The neural state which produces the conscious epiphenomenon also produces awareness, and recall, of that conscious state. Because the neural state is a reliable indicator of the existence of the conscious state, the evidence we have for the existence of the neural state is *ipso facto* evidence for the existence of the conscious one.

A further problem concerns evidence for epiphenomenalism as a theory. Epiphenomenalism would be refuted — and perhaps a strong dualism established — by the discovery of mental efficacy which is *not* neural efficacy. But to confirm the theory requires proving the absence of mental causal power, and a negative is notoriously difficult to establish conclusively. The case for epiphenomenalism is the case for physical closure — a global proposition supported only by overall theoretical power — combined with the case against the reduction of mental states to physical ones.

The evolution problem. One of the most persistent criticisms of epiphenomenalism, associated especially with Sir Karl Popper (see Popper and Eccles 1977), is the claim that it is incompatible with the theory of evolution.

According to the epiphenomenalist, creatures with consciousness will behave in exactly the same way as creatures with the same neural organization, but without consciousness. So consciousness can confer no reproductive advantage. Accordingly, there can be no natural mechanism which selects for consciousness, no matter what the processes that drive selection may be. So consciousness cannot emerge in an evolutionary development.

The argument as it stands is faulty; for it may be that although there can be no selection *for consciousness*, there may well be selection for something else — complex neural organization — of which consciousness is an inevitable by-product. The argument does retain some force, however, as an impotent consciousness seems to be such a gratuitous and inexplicable addendum. If consciousness is genuinely impotent, it is hard to see why neural organizations that match conscious ones in power, but lack consciousness, should not have evolved alongside the existing ones. And if epiphenomenalism is true, we have a further epistemic problem here. How could we know that consciousness does always accompany neural complexity (see *Other minds*)?

The problem of meaning. The thought is widespread that, at least in fundamental cases, the referential part of the meaning of reports of the existence of a state or an event depends on a causal chain from the state or event in question to the report (see *Reference* /4). Can epiphenomenalists accept this principle? According to them mental states and events do not feature in the causal history of any utterance. Thus if the principle were true, the vocabulary of

mental descriptions would be critically defective, and all reports concerning mental states and events would be meaningless.

While the epiphenomenalist's strong notion of causation prevents an appeal to a direct causal link in explaining the meaning of mental reports, an appeal might be made to counterfactual dependencies. A certain neural state causes me to feel like a million dollars, and also to utter the words "I feel like a million dollars." The feeling does not cause the utterance, but if I did not feel this way then the neural state would not have existed, and neither, therefore, would the utterance have occurred.

The problem here is the symmetry of the counterfactual dependency in question. Neither the feeling nor the utterance would have existed without the other. Yet the epiphenomenalist needs to hold that the utterance is *about* the feeling, while the feeling is not *about* the utterance.

Faced with the options of no semantic relation between utterance and mental item, or a semantic relation obtaining both ways, the epiphenomenalist chooses the former. All mental reports are meaningless — but only when considered as purely physical noises. For the epiphenomenalist, meaning and reference are *mental* phenomena, and hence are *epiphenomena*: when you say to me "I had an idea," I act upon your words purely mechanically, as a computer at a telephone exchange responds to a sequence of tones. Meaning does not enter here, and were none of us conscious we would all behave — linguistically and non-linguistically — exactly as we do now. However, in addition to the causal relations involved in this automatic process there are semantic relations obtaining between the mental effect of the neural state caused in me when I hear your words, and the idea which you have had.

For the epiphenomenalist, meaning does not span the gap between mental and physical; meaning remains on the mental side, causation on the physical. While this is a consistent position, it does underline the counterintuitive nature of epiphenomenalism.

References and further reading

Bonnet, C. (1755) *Essai de Psychologie*, London, ch. 32. (Sympathetic discussion, using the possibility of an automaton to replicate human behaviour.)

Broad, C.D. (1925) *The Mind and its Place in Nature*, London: Routledge & Kegan Paul, esp. chaps 3, 10 and 12. (Clear and judicious, but ultimately non-committal, exposition.)

Campbell, K. (1984) *Body and Mind*, Notre Dame, IN: University of Notre Dame Press, 2nd edn. (Classic source for new epiphenomenalism.)

Hodgson, S. (1865) *Time and Space: A Metaphysical Essay*, London: Longmans, Green, part 1, ch. 5, /30. (Earliest explicit presentation of epiphenomenalism.)

Hodgson, S. (1870) *Theory of Practice: An Ethical Enquiry*, London: Longmans, Green, book I, ch. 3. (First explicit defence of the theory.)

Huxley, T. (1874) *On the Hypothesis that Animals are Automata, and Its History* *On Collected Essays*, London: Macmillan, 1893, vol. 1, *Method and Results*, 199—250. (Classic presentation of the theory and the grounds for it.)

Jackson, F. (1982) "Epiphenomenal Qualia" *Philosophical Quarterly* 32: 127—36. (Influential article defending the legitimacy of epiphenomena.)

James, W. (1890) *The Principles of Psychology*, New York: Holt, ch. 5. (Extended, ultimately critical discussion.)

Lachs, J. (1963) "The Impotent Mind" *Review of Metaphysics* 17: 187—99. (Defends the theory by responding to many leading objections.)

Plato (*c.* 370 BC) *Phaedo*, trans. D. Gallop, Oxford: Clarendon Press, 1975. (Suggests, but does not support, the dependence of mind on matter.)

Popper, K. and Eccles, J. (1977) *The Self and its Brain*, London: Springer International. (Adopts a dualist position and presents the evolution objection to.)

Santayana, G. (1905) *The Life of Reason*, London: Constable, vol. 1. (Expresses the view that consciousness cannot produce, but only confer, meaning on events.)