

Epiphenomenalism

First published Mon Jan 18, 1999; substantive revision Sat May 11, 2019

Epiphenomenalism is the view that mental events are caused by physical events in the brain, but have no effects upon any physical events. Behavior is caused by muscles that contract upon receiving neural impulses, and neural impulses are generated by input from other neurons or from sense organs. On the epiphenomenalist view, mental events play no causal role in this process. Huxley (1874), who held the view, compared mental events to a steam whistle that contributes nothing to the work of a locomotive. James (1879), who rejected the view, characterized epiphenomenalists' mental events as not affecting the brain activity that produces them "any more than a shadow reacts upon the steps of the traveller whom it accompanies".

Ancient theories of the soul gave rise to debates among Aristotle's successors that have a strong resemblance to some contemporary discussions of the efficacy of mental events (Caston, 1997). The modern discussion of epiphenomenalism, however, traces back to a 19th century context, in which a dualistic view of mental events was assumed to be correct. The first part of our discussion — Traditional Arguments — will be phrased in a style that reflects this dualistic presupposition. By contrast, many contemporary discussions work within a background assumption of the preferability of materialist monism. One might have supposed that this position would have put an end to the need to investigate epiphenomenalism; but, as we shall see under Arguments in the Age of Materialism, such a supposition is far from being the case. A brief outline of both discussions follows.

- 1. Traditional Arguments (A) Pro
 - 2. Traditional Arguments (B) Con (with Epiphenomenalists' Responses)
 - 2.1 Obvious Absurdity
 - 2.2 Natural Selection
 - 2.3 Knowledge of Other Minds
 - 2.4 Self-stultification
 - 3. Arguments in the Age of Materialism
 - 3.1 Three Routes to Puzzlement: Externalism, Anomalous Monism and Realization
 - 3.2 Responses to Exclusion Arguments
 - 3.3 Mental Properties and Distinctive Causal Powers
 - 3.4 Epiphenomenalism and Intrinsic Properties
 - 3.5 Empirical Considerations
 - 4. Historical Note on Automatism and the Term "Epiphenomenalism"
 - Bibliography
 - Academic Tools
 - Other Internet Resources
 - Related Entries
-

1. Traditional Arguments (A) Pro

Many philosophers recognize a distinction between two kinds of mental events. (A) The first goes by many names, e.g., phenomenal experiences, occurrences of qualitative consciousness, the what-it-is-like of experience, *qualia*. Pains, afterimages, and tastes can serve as examples. (B) Mental events of the second kind are occurrent propositional attitudes, e.g., (occurrent) beliefs and desires. Arguments about epiphenomenalism may concern either type of mental event, and it should not be assumed that an argument given for one type can be rephrased without loss for the other. The two types can often be connected, however, through beliefs that one has one's experiences. Thus, if it is held that pains have no physical effects, then one must say either (i) pains do not cause beliefs that one is in pain, or (ii) beliefs that one is in pain are epiphenomenal. For, if pains caused beliefs that one is in pain, and the latter had physical effects, then pains would, after all, have effects in the physical world (albeit indirectly). But epiphenomenalism says mental events have *no* effects in the physical world.^[1]

The central motivation for epiphenomenalism lies in the premise that whenever there is a sufficient cause of a physical event, there is a sufficient *physical* cause of that event. If a mental event is something other than a physical event, then for it to make any causal contribution of its own in the physical world would require a violation of physical law. Descartes' (1649) interactionist model proposed that nonphysical events could cause small changes in the shape of the pineal gland. But such nonphysical effects, however slight, would mean that the physical account of motion is *false* — for that account says that there will be no such change of shape unless there is a physical force that causes it.

One may try to rescue mental efficacy by supposing that whenever there is a mental effect in the physical world there is *also* a physical force that is a sufficient cause of the effect. This view, however, both offends Occamist principles and fails to satisfy the leading anti-epiphenomenalist intuition, namely, that the mental *makes a difference* to the physical, i.e., that it leads to behavior that would not have happened in absence of the mental. The view also leads to an epistemological problem: If there is always a sufficient physical cause for whatever a mental event is supposed to produce, then one could never be in a position where one *needs* to suppose there is anything non-physical at work, and thus there could never be any reason to introduce mental causes into one's account of neural events or behavior.

Many contemporary thinkers would respond to the central motivation for epiphenomenalism by denying its dualistic presupposition, i.e., by holding that mental events are identical with physical events, and may therefore have physical effects. Questions that remain for such physicalistic views will be explained in section 3. For now, it should be noted that the argument stated in the previous two paragraphs is not supposed to be an argument for dualism, but only for adopting epiphenomenalism, once dualism is accepted.

Further support for epiphenomenalism can be derived from the fact, noted by Wilhelm Wundt (1912), that “each simple sensation is joined to a very complicated combination of peripheral and central nerve processes”, together with the fact that the causes of behavior are likewise complex neural events. This latter fact makes it natural to look for complex events throughout the causal chain leading to behavior; and these can be found in the neural events that are required for the occurrence of simple sensations. The sensations themselves could not contribute to behavior without first having neural effects that are more complex than themselves. Thus an anti-epiphenomenalist stance would require us to prefer the hypothesis that simple sensations cause (relatively) complex neural events over the

hypothesis that complex neural events (that are required in any case for the causation of sensations) are adequate to cause the neural events required for the causation of behavior.

2. Traditional Arguments (B) Con (with Epiphenomenalists' Responses)

2.1 Obvious Absurdity

Epiphenomenalism is absurd; it is just plain obvious that our pains, our thoughts, and our feelings make a difference to our (evidently physical) behavior; it is impossible to believe that all our behavior could be just as it is even if there were no pains, thoughts, or feelings. (Taylor, 1963 and subsequent editions, offers a representative statement.)

This argument is surely the briefest of those against epiphenomenalism, but it may have been more persuasive than any other. Epiphenomenalists, however, can make the following reply. First, it can never be obvious what causes what. Animated cartoons are full of causal illusions. Falling barometers are regularly followed by storms, but do not cause them. More generally, a regularity is causal only if it is not explained as a consequence of underlying regularities. It is part of epiphenomenalist theory, however, that the regularities that we observe to hold between mental events and actions can be explained by underlying regularities. Schematically, suppose physical event P_1 causes both mental event M and physical successor P_2 , as in Figure 1.

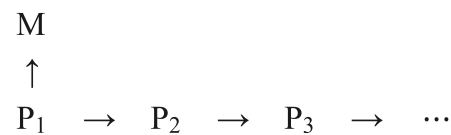


FIGURE
1.

Suppose there is no other cause of M , and no other cause of P_2 . Then every M will be followed by P_2 , yet the cause of P_2 will be adequately found in P_1 . It is true that, under the assumptions stated, the counterfactual, “If M had not occurred, then P_2 would not have occurred” holds; but then, so may “if the barometer had not fallen, the storm would not have occurred.” The moral to be drawn is that causation may imply that certain counterfactuals hold, but the holding of counterfactuals is not enough to show causation. Thus, it is true that some of our actions would not have occurred, under normal conditions, unless we had had certain mental events. But this fact cannot show that those actions are caused by our mental events (rather than being caused by the physical causes of those mental events).

It is often said that pains cause withdrawals of affected parts of the body. In extreme cases, however — for example in a case of touching a hot stove — it can be observed that the affected part is withdrawn *before* the pain is felt. These cases cannot show that pain never causes withdrawals, but they do show that pain is not necessary as a cause of withdrawals. In less extreme cases, it is open to the epiphenomenalist to hold that the causal order is the same as in the extreme cases (i.e., some physical event, P_1 , causes both withdrawal and pain) but is not ordinarily recognized to be so.

A variant of the obvious absurdity objection is that epiphenomenalism leads to a feeling of loss of self, or a sense that we can no longer regard our actions as *ours*. (See Hyslop, 1998.) Epiphenomenalists

may, however, reply that there would be different brain states corresponding to the difference between cases of moving our bodies and cases of having our body parts moved by something else, and their view allows different brain states to have different phenomenological effects. Epiphenomenalism is also compatible with there being a distinction between actions that proceed from normal brain processes in normal conditions, and actions that proceed from manipulation by others, disease, or situations that overwhelm normal reasoning abilities.

2.2 Natural Selection

The development of consciousness must be explainable through natural selection. But a property can be selected for only if it has an effect upon organisms' behavior. Therefore, consciousness (both qualia and intentional states) must have effects in behavior, i.e., epiphenomenalism is false. (See Popper and Eccles, 1977; James (1879); Romanes, 1896)

According to the same biology that embraces natural selection, however, behavior has muscular causes, which in turn have neural causes. Barring neural events that are inexplicably in violation of biological constraints on their conditions of activation, there must be an adequate physical cause of every link in the causal chain leading to behavior. Thus, it is easily understood how certain kinds of neural events can be selected for. Epiphenomenalists hold that conscious events are effects of (certain) neural events. Thus, it fits well in their view that we have the conscious events we do because the neural causes of these events have been selected for. Indeed, if neural causes of behavior are selected for, and are sufficient causes, there *cannot* be any *further* effect attributed to natural selection.

William James (1879; 1890; see also Bradley, 2011) offered an intriguing variant of the argument from natural selection. If pleasure and displeasure have no effects, there would seem to be no reason why we might not abhor the feelings that are caused by activities essential to life, or enjoy the feelings produced by what is detrimental. Thus, if epiphenomenalism (or, in James's own language, automaton-theory) were true, the felicitous alignment that generally holds between affective valuation of our feelings and the utility of the activities that generally produce them would require a special explanation. Yet on epiphenomenalist assumptions, this alignment could not receive a genuine explanation. The felicitous alignment could not be selected for, because if affective valuation had no behavioral effects, misalignment of affective valuation with utility of the causes of the evaluated feelings could not have any behavioral effects either. Epiphenomenalists would simply have to accept a brute and unscientific view of pre-established harmony of affective valuation of feelings and the utility of their causes.

Epiphenomenalists can respond to James's argument by offering support for the following two views. (I) There is a distinction between (a) neural events in sensory systems, which cause feelings, and (b) neural events in a reward system (where an event is a member of a reward system if and only if it contributes to continuance (or, discontinuance) or repetition (or avoidance) in similar circumstances, of any behavior that leads to that event). (II) What "pleasure" refers to in any possible world is the effect in consciousness of the events in (b) that contribute to continuance or repetition. Since both (a) events and (b) events have neural and, ultimately, behavioral effects, they can be selected for, and so can their combination. The alignment of feelings caused by useful stimuli and pleasure (and, by parallel reasoning, alignment of feelings caused by harmful stimuli and displeasure) would then follow from alignment of neural events of the kinds described in (a) and (b). This latter alignment is independently plausible. A reward system that can lead to quick decisions and a sensory system that provides discriminations for use in longer term planning would both confer advantages; and these

systems would, in general, have to work together in a successful organism. (For fuller discussion of James's argument see Robinson 2007, 2019; and for the structure of pleasure, see Robinson 2006b, 2019.)

2.3 Knowledge of Other Minds

Our reason for believing in other minds is inference from behavioral effects to mental event causes. But epiphenomenalism denies such a causal connection. Therefore, epiphenomenalism implies the (exceedingly implausible) conclusion that we do not know that others have mental events. (Jackson, 1982, replies to this and several other arguments against epiphenomenalism. The argument is stated, and accepted, by Benecke, 1901.)

The first premise of this argument is a widely held dogma, but it can be denied without absurdity. (See Robinson, 1997.) It is perfectly obvious to everyone that the bodies of human beings are very much alike in their construction, and it requires no sophisticated reasoning to infer that if others are made like me, they probably hurt when affected like me, e.g., when their bodies are stuck with pins, beaten, cut and so on. There is no principle that makes an inference from similar effects to similar causes more secure than an inference from similar causes to similar effects; on the contrary, the latter inference is more secure, because there can sometimes be quite different causes of extremely similar effects. Thus, an inference to other minds that is allowed by epiphenomenalism must be at least as strong as the inferential route to other minds with which it is incompatible.

2.4 Self-stultification

The most powerful reason for rejecting epiphenomenalism is the view that it is incompatible with knowledge of our own minds — and thus, incompatible with knowing that epiphenomenalism is true. (A variant has it that we cannot even succeed in referring to our own minds, if epiphenomenalism is true. See Bailey (2006) for this objection and Robinson (2012) for discussion.) If these destructive claims can be substantiated, then epiphenomenalists are, at the very least, caught in a practical contradiction, in which they must claim to know, or at least believe, a view which implies that they can have no reason to believe it. Moreover, unless epiphenomenalists can consistently claim to know their own minds, they cannot offer the response to the other minds objection given in 2.3 above. (See Walton, 1989.) (Many authors give some version of this objection. For a full statement of this argument, and several others concerning epiphenomenalism, see Chalmers, 1996. For strong recent versions, see DeBrigard, 2014; Moore, 2012 (responded to by Robinson, 2013); and Moore, 2014.)

The argument that is given to support the destructive claims is that (i) knowledge of one's mental events requires that these events cause one's knowledge, but (ii) epiphenomenalism denies physical effects of mental events. So, either we cannot know our own mental events, or our knowledge of them cannot be what is causing the plainly physical event of our saying something about our mental events. Thus, suppose S is an epiphenomenalist, and that S utters "I am in terrible pain." S is committed to the view that the pain does not cause the utterance. But then, it seems, S would be making the same utterance whether or not a pain were occurring. If this is so, then S's testimonies about S's own pains are worthless — both to us and to S. They cannot be taken to represent any knowledge about pains on S's part (if S's epiphenomenalist view is true). In fact, on an epiphenomenalist view, all the arguments for epiphenomenalism and rebuttals to counterarguments we have reviewed might be given even if we were all zombies — i.e., even if we were all possessed of physical causes of our utterances and completely devoid of any mental life whatsoever.

The argument that epiphenomenalism is self-stultifying in the way just described rests on the premise that knowledge of a mental event requires causation by that mental event. But epiphenomenalists may reject that premise without absurdity. One way of seeing how to do this involves considering the interactionist diagram in Figure 2, which shows P_1 as directly causing M but not P_2 , and M directly causing P_2 . (Directly causing is an intransitive relation. Causation (when used without modifier) is transitive: events are causally related if there is a chain of direct causes, however long, that connects them.)

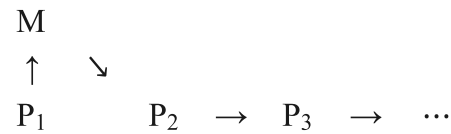


FIGURE
2.

Now consider P_3 , which is directly caused by P_2 and which we will assume to cause (directly or indirectly) further behavior such as S 's utterance of "I am in terrible pain". P_3 is not directly caused by M . Does it convey knowledge of M ? If we answer negatively, on the ground that P_3 is not directly caused by M , we will be rejecting interactionism for virtually the same reason that epiphenomenalism is thought to be unacceptable. Since this is an extremely implausible stance, let us take it that P_3 does convey knowledge of M . But what property does P_3 actually have that makes it a case of conveying knowledge of M ? Epiphenomenalists will wish to point out that P_3 does not have any property that contains information as to how it was caused. Looking backward from P_3 , so to speak, one cannot tell whether it was indirectly caused by M (as in the interactionistic Figure 2) or indirectly caused by M 's cause (as in the epiphenomenalistic Figure 1). There is, however, a property that P_3 does have that is intuitively strongly connected to its conveying knowledge of M — namely, that it would not be occurring unless M had recently occurred. But P_3 has this property on epiphenomenalist and interactionist views alike. Thus, if not occurring unless M has recently occurred is the property that is responsible for P_3 's conveying knowledge of M , epiphenomenalists have as much right as anyone to claim that P_3 conveys knowledge of M , and they are not debarred from knowing what they claim to know.

Critics of epiphenomenalism can of course point out that there is a property that interactionism, but not epiphenomenalism, assigns to P_3 — namely, the property of being indirectly caused by M . Epiphenomenalists, however, are likely to think that the intuitive connection between this property and knowledge is much weaker than that between knowledge of M and the fact that P_3 would not be occurring unless M had recently occurred. In fact, they may hold that the relevance of indirect causation is exhausted by its ensuring that P_3 would not occur unless M had recently occurred. They can then reiterate that there is another way of ensuring that this condition holds, namely, the set of relations, diagrammed in Fig. 1, that is affirmed by epiphenomenalism.

The foregoing way of responding to the self-stultification argument is further explained and defended in Robinson (1982b, 2013; see also 2006a). An alternative response can be found in Chalmers (1996; see also Nagasawa, 2010). Chalmers's property-dualistic view holds that there is more to a person than just a brain and a body. It allows for persons to be directly acquainted with experiences, and it is this direct acquaintance, rather than any causal relation, that justifies our beliefs about experiences. On this view, experiences are partially constitutive of beliefs about experiences, and "the justification of my belief [about experiences] accrues not just in virtue of my physical features but in virtue of some of

my nonphysical features — namely the experiences themselves” (Chalmers, 1996, p. 198). In supplying non-causal relations to support the claim to knowledge of experiences, this view disconnects the knowledge question from the question of how things stand causally, and thus avoids the self-stultification argument.

A third response to self-stultification begins with the observation that terms for sensory qualities (words for colors, flavors, pitches and timbres, etc.) must be learned. Very plausibly, this learning depends (a) on correlations between properties of present objects and production of words for those properties by surrounding adult speakers; and, consequently, (b) on correlations between brain event types that are caused by the properties of those objects, and brain event types that are caused by the hearing of adults’ words. It also plausibly depends (c) on the brain’s sensitivity to the correlation of brain events in (b). It is part of epiphenomenalist theory that the first brain events mentioned in (b) include causes of sensations of characteristic kinds, and that, after learning, those same brain events causally determine the production of predicates in sensation reports. The combination of these relations ensures that, so long as conditions in the brain are normal, a report of a sensation is guaranteed to correctly identify the kind of sensation that is normally caused in the speaker by objects that are standardly described by the predicate in the report. This guarantee underwrites a claim to knowledge of what kinds of sensation we are having.

Conditions may, of course, sometimes be abnormal. But that cannot be a defeater for epiphenomenalist knowledge of our sensations, because all views must allow for the possibility of cognitive breakdowns such as unnoticed slips of the tongue or linguistic difficulties brought on by disease.

In a 2006 paper, Michael Pauen has given a particularly persuasive argument that offers a new and deeper twist to the problem of what epiphenomenalists can consistently claim to know about qualia. Pauen’s focus is on the epiphenomenalists’ claim that there are laws linking the occurrence of neural events of certain kinds with occurrences of qualitative events. Since epiphenomenalists deny the identity of qualia and physical properties, such laws are not only contingent, but independent of physical laws. Thus, epiphenomenalists must concede the possibility of worlds in which the physical laws are as they are here, but there are no laws connecting physical events to qualitative events. If such worlds are possible, we might actually be in such a world, and so the epiphenomenalist can be asked to provide evidence that we are not in such a world. But the epiphenomenalist denial of efficacy for qualia precludes the possibility of providing any such evidence. Even conceding that we know at each moment what qualia we are having, we cannot exclude the possibility that our memories seem to tell us of qualitative events that never occurred (or that were of different kinds than we seem to remember); yet we would have to rely on such memories in order to have evidence of causal laws holding between physical and qualitative events. (Another argument that turns on memory is given by Swinburne, 2011.)

Alexander Staudacher (2006) has given a critical discussion of Pauen’s paper. Among several points, perhaps the strongest is a comparison between Pauen’s challenge and the challenge to prove I am not a brain in a vat. Inability to decisively answer this latter challenge is not generally taken to impugn our knowledge concerning tables and chairs, nor are physicists expected to lay it to rest before proceeding with their science. Analogously, it may be that worlds without causal laws are not “relevant alternatives” for epiphenomenalism, i.e., not alternatives that epiphenomenalists are fairly required to rule out. Staudacher points out that if we are allowed to raise skeptical challenges of the kind Pauen raises, it will be possible to construct analogous, unanswerable skeptical challenges for interactionism.

And, although Staudacher does not provide an analogous case for physicalism, it seems that parallel difficulties are constructable. For example, according to physicalism, it would seem to be a conceptual possibility that qualia are identical, not with neural properties, but with a combination of neural properties and phases of the Moon, while memory traces depend only on neural properties. For reasons parallel to Pauen's, it seems we could never have evidence that would rule this possibility out. But it is not clear that physicalists need to admit the necessity of providing evidence against such a view, in order to responsibly affirm their view. If that is right, parallel reasoning may be able to support Staudacher's way of resisting the force of Pauen's challenge.

J. Megill (2007) has raised a problem for epiphenomenalism based on the assumption that properties must be individuated by their causal relations. Since epiphenomenalists deny efficacy to phenomenal qualities, the only causal relations to which they could appeal to individuate them would be causes of events with those qualities. However, different causes may produce the same effect. Thus, appeal to difference of causes alone is an insufficient basis for individuation of phenomenal qualities.

Epiphenomenalists can be expected to deny the necessity of causal relations for individuation. They may consistently hold, for example, that the unique hues are qualities that simply differ and have no metaphysical need for a principle of individuation. An epistemological problem, however, may appear to remain: How can epiphenomenalists know that they experience different phenomenal qualities? This, however, is the self-stultification problem in a different guise, and epiphenomenalists can refer to the responses already considered.

3. Arguments in the Age of Materialism

One might have thought that if the mental and the physical are identical, there could be no room for epiphenomenalistic questions to arise. Behavior is caused by muscular events, and these are caused by neural events. Mental events will be identical with some of these neural events; so whatever effects these neural events have will be effects of mental events, and mental events will make a causal contribution to, i.e., will "make a difference" to our behavior.

Questions about epiphenomenalism, however, arise the moment any distinction is made between the mental properties and the physical properties of an event. Section 3.1 will explain three ways in which this can be done within a broadly materialist monism. The third of these ways is still the subject of lively debate, and some of the issues will be explained in sections 3.2 and 3.3.

It should be noted that most recent writers take a somewhat dogmatic position against epiphenomenalism. They presume that epiphenomenalism is to be avoided, and they go to great lengths to try to show that they have avoided incurring that anathema, despite maintaining the sufficiency of physical causation in conjunction with some kind of distinction between the mental and the physical.

3.1 Three Routes to Puzzlement: Externalism, Anomalous Monism and Realization

(1) Many philosophers hold an *externalist* view of intentionality, according to which intentionality requires representation, and representation depends on circumstances external to the body of the representing subject. (To illustrate with Putnam's (1975) famous example from Twin Earth, what a

thought that S might express by “This is water” is actually about depends on what the transparent, tasteless liquid in S’s environment actually is.) It seems, however, that the causal determinants of S’s behavior can depend only on events occurring inside S’s body. Thus, if externalism is right, what S does cannot depend on the intentional content of S’s thoughts.

This conclusion is compatible with holding that a proper *description* of S’s behavior should refer to circumstances external to S. For example, describing S as *reaching for a glass of water* may not be appropriate unless S believes the glass contains water; and that this is what S believes may depend on circumstances external to S. But then, it is at least tempting to conclude that it cannot be the intentional character of S’s belief that is causing the extension of S’s arm toward the glass. After all, S’s twin-earth double has, in some sense, a different belief (one that refers to XYZ) but the internal bodily story of the causation of the double’s arm-extension will be exactly the same as the story for S (up to substitution of XYZ in the double’s body wherever S has an H₂O molecule).

(2) Donald Davidson’s (1970) anomalous monism held that (i) each mental event is identical with a physical event, but (ii) there are no psychophysical laws. Davidson accepted the view that causation involves laws and, in view of (ii), held that the laws into which mental events entered related physical properties (or, mental events under their physical descriptions). Many philosophers regarded this view as tantamount to epiphenomenalism, i.e., to the view that causation of our behavior involves only the physical properties of our parts, and that the mental properties as such have no efficacy.

(3) A similar discussion rages today, but begins with a background assumption of multiple realization. It is thought to be highly implausible that belief in a particular proposition, or desire for a particular state of affairs, is identical with the same state of a brain (or a part of a brain) in different people, or even in the same person at different times. Instead, the same belief or desire is taken to be “realized” by different neural structures or states of activation on different occasions.

Since the physical properties of different neural events are different, transitivity of identity prohibits claiming identity of mental properties with any of the physical properties whose instantiation realizes those mental properties. In a body of work stretching over many years, Jaegwon Kim (1993, 2005, e.g.) has argued that this division of mental from physical properties leads to exclusion of mental properties from playing a causal role. After all, it is widely accepted that the physical properties of neural events, together with a person’s state of neural connectivity, are sufficient to produce later neural events. Since behavior depends on muscle contractions that depend on neural innervation, it seems that physical events and structures, and the physical laws that apply to them, are quite enough to bring about our behavior. There is nothing left for mental properties to do.

There are, of course, some cases in which there is more than one sufficient cause of an event. Caesar’s death seems likely to have been overdetermined, i.e., it is probable that more than one attacker’s stroke would have been sufficient by itself for his death. But instances of mental properties are not distinct from instances of physical properties that are their realizers in the way that stabbings are distinct from each other, and systematic overdetermination of our behavior is generally regarded as implausible.

3.2 Responses to Exclusion Arguments

Many writers have held that Kim-style exclusion arguments depend on an understanding of causation for which there are preferable alternatives. (See, e.g., Woodward, 2008, 2017; Kroedel, 2016; Beebe, 2017.) This kind of response is sometimes put forward as providing a more correct account of

causation. Other writers allow that exclusion arguments can succeed if causation is understood as productive (or ‘oomphy’) causation, but hold that there are other “conceptions of causation” (Beebe, 2017) on which it remains acceptable to say that our mental states cause our behavior.

Epiphenomenalists may concede that there is some such sense, but still think their view is vindicated if all the productive causation is provided by the physical activities of our brains.

Epiphenomenalists may also emphasize that examples in exclusion arguments are usually beliefs, desires or intentions. These mental properties are closely associated with behavioral dispositions, and can easily be understood in functional terms for which multiple realizability is immediately plausible. As Kim has explicitly recognized, however, (1993, p. 366), there are properties that have long caused difficulties for functionalism, namely, the qualities of phenomenal experiences such as pains, itches, tastes, smells, afterimages, and so on. The “explanatory gap” (Levine, 1983) or “unintelligibility” (Robinson, 1982a) or the “hard problem” (Chalmers, 1996) concerning the connection between neural events and phenomenal qualities can be otherwise expressed as our inability to see any necessity in that connection. We are unable to understand why it should be that a series of neural activations occurring in various degrees of intensity and temporal relations should always be accompanied by pain, or itch, or, indeed, by any phenomenal quality whatever. Inability to see any such necessity is, of course, not a proof that such a necessity does not obtain. Nonetheless, absent insight into the necessity of the connection between neural properties and qualitative properties, we are arguably in an *explanatory* position similar to traditional epiphenomenalism. That is, we will have a sufficient explanation of behavioral reactions to stimuli that invokes exclusively neural properties. In addition, we may hold the view that these neural properties are necessarily connected to qualitative properties; but, lacking explanation of this necessity, this connection will contribute no understanding of how qualitative properties could make a difference to behavior. Because this difficulty has not been removed in the case of qualia, the success or failure of the previously discussed Traditional Arguments remains relevant to contemporary thinking about epiphenomenalism.

3.3 Mental Properties and Distinctive Causal Powers

L. A. Shapiro and E. Sober (2007) have developed an alternative line of argument that they hold to apply just as well to functional properties and properties that resist functional analysis. The distinction between common causation and realization plays a key role. Common causation, they note, can lead to an empirically based denial that one co-effect causes another. For example, tail length in mice offspring is caused by their parents’ genes, not by their parents’ tail length — and this can be empirically shown by cutting parental tails (which, of course, leaves their genotype constant) and observing no effect on tail length of offspring.

It is impossible, however, to suppress a realized condition while leaving its realizer constant, so there is no sense to the demand that beliefs should have causal powers that are additional to the causal powers of their neural state realizers. Nor is it possible to suppress a phenomenal quality while holding constant a state with its (allegedly) identical neural property. Thus, saying that “there is nothing left for the property of being a belief (or a phenomenal quality) to do, once the neural realizer (or identical neural property) has done its work” does not undercut the efficacy of the beliefs or phenomenal qualities. It shows only that beliefs (or phenomenal qualities) do not satisfy a demand (for an ‘extra’ causal contribution) that is incoherent in itself.

J. A. Baltimore (2010) has responded by arguing that there are reasons other than that illustrated by the mice tails example for denying efficacy to mental events. M. Baumgartner (2010; see Woodward,

2015 for response) suggests that Shapiro and Sober's approach depends on a certain (interventionist-inspired) view of causation, but that attributions of causation in virtue of mental properties cannot be counted as making sense without departing from the terms of that view. Whether this strong critique is accepted or not, it is evident that removing a reason for objecting to efficacy of mental events in virtue of their mental properties does not by itself provide a positive reason for asserting such efficacy.

Such a reason may, however, be found in the subset view proposed by Shoemaker (2007). To briefly illustrate a key idea, suppose that a certain belief, M , can be realized by alternative physical states P_1, \dots, P_n . Suppose further that this belief is involved in an inference, and that its possessor arrives at a new belief, i.e., arrives at another state that has a new mental property, M^* . Suppose, finally, that M^* can be realized by alternative physical states P_1^*, \dots, P_n^* , and that each P_i is a sufficient physical cause for some P_j^* . A particular P_i causes a particular P_j^* by a particular causal route that need not be shared by other P to P^* instances, and so P_i has causal powers not shared by all instances of M . But any instance of M would have been realized by a sufficient cause of some realizer of M^* . Thus, M 's causal powers include just those that are relevant to bringing about M^* (assuming that other conditions, e.g., belief in other premises required for the inference, are held constant). Given these assumptions, it is reasonable to conclude that it is M rather than its particular realizers that causally contribute to M^* , even though each instance of M is realized by some P_i that is a sufficient cause of a realizer of M^* , and thus is sufficient for M^* .

The subset view is the subject of lively debate; see, e.g., Antony (2010), Audi (2012), Melnyk (2010), Shoemaker (2010). One objection is that the subset view assumes rather than shows that M has any causal powers at all, and that this is the key question at issue (Kim, 2010). Reasons have been offered for doubting this assumption, i.e., accepting epiphenomenalism, at least for beliefs and desires. Segal (2009) argues that being a belief and being a desire are dispositional properties, and that dispositional properties are not efficacious. Tammalleo (2008) supports an exclusion argument, and invokes mechanisms independently established by cognitive psychology to explain why we have the mistaken intuition that mental states are causes.

3.4 Epiphenomenalism and Intrinsic Properties

Frank Jackson (1982) has given an epiphenomenalistic argument that has spawned lively responses from many quarters. This argument turns on the concept of physical information, where "physical information" is information about ourselves and our world of the kind that is obtainable in the physical, chemical, and biological sciences. In Jackson's argument, a brilliant scientist, Mary, has learned all the physical information there is about color vision. Having been confined to a black and white room, however, Mary has never had a color experience. Jackson asks whether Mary will learn anything when she is released from her confinement and thus comes for the first time to have color experiences. It seems compelling that she would learn something; but as she already has all the physical information there is, what she learns must be some other kind of information, which we may call "phenomenal information". This "knowledge argument" has been regarded as a strong reason to accept a dualistic view of our experiences. When combined with the traditional arguments (Pro) given above, it becomes a potent source of support for epiphenomenalism.

David Lewis (1988) undertakes a thorough response to the knowledge argument. Among Lewis's many considerations, there is one that seeks to enforce a connection between phenomenal information *per se* and epiphenomenalism. According to Lewis's argument, even if one says that phenomenal *events* are identical with physical events, and even if one says that phenomenal events

produce physical effects in violation of physical laws, one will still be led to a form of epiphenomenalism if one says there is phenomenal information that is irreducibly different from physical information. To put the argument in ruthlessly summary form, let V_1 and V_2 be two possibilities for the phenomenal information that one acquires by, and only by, tasting Vegemite. Suppose that P_1 is a physical state produced by the taste of Vegemite. That the taste of Vegemite has this physical effect is a piece of physical information. But this same physical information is compatible with two possibilities, (a) V_1 is related by a law, L_1 , to P_1 ; and (b) V_2 is related by a different law, L_2 , to P_1 . Now, either of these possibilities is compatible with all the physical information we have; i.e., their difference makes no physical difference. Thus, that the phenomenal information in the taste of Vegemite is, say, V_1 rather than V_2 can make no difference to anything physical, i.e., V_1 is epiphenomenal. Lewis's point here is not to argue for or against epiphenomenalism; rather, he assumes epiphenomenalism is false, and uses the fact that the hypothesis of phenomenal information leads to it as an argument against that hypothesis.

Denis Robinson (1993) raises the possibility that Lewis's argument can be extended to produce a far-reaching and puzzling result. Suppose that I_1 and I_2 are two possibilities for an *intrinsic* property of a basic physical entity, e.g., a quark. Everything relevant to physics can be expressed by the lawlike relations in which quarks stand to fundamental physical objects and properties. Let this set of relations be S . It appears that there are two possibilities, (a) I_1 is related by a set of laws, L_1 , to S ; or (b) I_2 is related by a different set of laws, L_2 , to S . Either of these possibilities is compatible with all the physics we have, i.e., their difference makes no physical difference. Thus, that the intrinsic property of quarks is, say, I_1 rather than I_2 can make no difference to physics, i.e., I_1 is epiphenomenal. The generalization of this point is that the intrinsic properties of the fundamental objects of physics must be epiphenomenal.

It thus appears that we must either (1) deny that fundamental objects of physics have any intrinsic properties, or (2) deny that Lewis's argument for the connection of phenomenal information with epiphenomenalism is sound, or (3) deny that Lewis's argument can be paralleled in the suggested way for the case of intrinsic physical properties, or (4) admit an epiphenomenalism of intrinsic properties into our view of the basic structure of physical reality.

Bertrand Russell (1927, p. 382) held the view that physical theory can reveal only causal structure, or "formal properties" of matter, and that "by examining our percepts we obtain knowledge which is not purely formal as to the matter of our brains." This idea is taken up sympathetically (with substantial reworking in a quantum mechanical context) by Lockwood (1993). Chalmers (1996) offered a useful discussion of the view (now known as "Russellian monism"), and expressed some sympathy for it — a sympathy which has increased in intervening years (see, e.g., Chalmers, 2010). Denis Robinson (1993), however, regards intrinsic similarity of fundamental physical entities as different from similarity of phenomenal properties.

If phenomenal properties are intrinsic properties of fundamental physical objects, and the latter stand in lawlike relations, then lawlike relations will hold between phenomenal properties and some physical occurrences. This conclusion appears to give a causal role to phenomenal properties and thus to suggest a way out of epiphenomenalism. But if intrinsicity carries epiphenomenality, as D. Robinson's extension of Lewis's argument suggests, then this way out of epiphenomenalism would be blocked. Moreover, since there is no phenomenal quality that we are always experiencing, no instantiation of a quality by a fundamental physical particle can, by itself, be one of our sensations. It

is thus not clear that Russellian monism gives any more causal role to our sensations than does epiphenomenalism (see Robinson, 2018 for elaboration).

3.5 Empirical Considerations

There are some empirical results that engage epiphenomenalism in several ways. These empirical results do not cut as deeply as the foregoing arguments, for they do not claim to show that consciousness is completely inefficacious. In particular, they do not show, and are not aimed at showing, that episodes of consciousness do not causally contribute to reports of such episodes. They suggest instead that the reported consciousness is not causally related to our nonlinguistic behavior in ways that we ordinarily suppose.

An early and very well known argument derives from the work of B. Libet (1985, 2004). Libet asked participants in his experiments to make an unplanned movement when they felt the urge to do so, and to report (using a specially designed clock-like instrument) the time when they felt the urge to move. Recordings of participants' brain waves were made during these experiments, and Libet found a signal (the "readiness potential", RP) that correlated with the movements. The key finding was that (after allowing for corrections due to neural transmission times) the RP occurred about 350 ms earlier than the time that participants reported having the urge to move. Libet himself allowed for the possibility of a veto after the urge that would block the movement. However, many have drawn an implication that the conscious urge to make a movement comes too late to causally contribute to a movement that has already been initiated by brain processes (that were detected by the RP measurement). A corollary of this implication is that, to the extent we have an intuition that our urge to move caused the movement, we are subject to a causal illusion. The possibility of causal illusion, in turn, weakens the intuitive argument against epiphenomenalism explained in section 2.1 above.

More recently, work led by J-D. Haynes (2013; see also Soon, *et al.*, 2008) has found brain events that are somewhat correlated with decisions, and that occur several seconds before the decision is made. This work seems unsettling for the following reason. If we believe we were weighing options on a matter that was still open at, say, one second before our conscious decision, but our action was already predictable from neural events before that, then we would be under some illusion about the extent to which our conscious decision has an effect on our behavior.

Work by Wegner and Wheatley (1999), Wegner (2002), and Linser and Goschke (2007), describes participants who have judged themselves to have partial control of movements over which they in fact lacked control. But if we can demonstrably have illusions about being in control, our sense of being in control of our actions cannot be taken as evidence that our conscious intentions actually have effects in our (nonlinguistic) behavior. Such illusions further imply that we sometimes lack knowledge of the actual course of causation of our actions.

In well-regarded studies, Nisbett and Wilson (1977a, 1977b) showed that people sometimes confabulate, that is, they give reasons for evaluative judgments that do not reflect the actual causes of making those judgments. This work again carries two implications: that people are sometimes mistaken in accepting efficacy of what they believe to be their reasons for a judgment, and that they are sometimes mistaken about the real reasons for their behavior.

The implications drawn from these experiments have been criticized on several grounds. In some cases, there are technical criticisms of methods or statistical analyses (see Dennett, 1991 for critique of Libet's conclusions, and S. Walter's 2014 discussion of the Wegner and Wheatley (1999) experiment).

The force of the work led by Haynes is mitigated by the fact that predictions of behavior from several seconds before its occurrence, while statistically significant, are far from perfect. Many other points have been made (see Nahamias, 2008; Mele, 2014; Shepherd, 2017; Baumeister *et al.*, 2018). Two basic criticisms are offered by many writers. First, many of the experimental conditions involve meaningless setups (ouija-like boards, for example) or meaningless decisions (e.g., when to move a finger or which of two buttons to press) that bear little resemblance to real-life decisions (e.g., whether to accept a job, or go to a particular place for a vacation) where people bring substantive reasons and personal preferences to bear. These oddities and simplifications may be thought to render ineffective psychological mechanisms that would be operative in more realistic cases and that would forestall illusions about the relation between our consciously entertained reasons and our behavior. Second, the fact that there are some cases in which unconscious influences have a noticeable effect on our behavior does not show that we are never, or even not usually, acting in a manner that would be rational, given our particular longstanding beliefs and preferences.

4. Historical Note on Automatism and the Term “Epiphenomenalism”

James Ward’s *Encyclopedia Britannica* (tenth edition, 1902) article, “Psychology”, contains the following summary of T. H. Huxley’s view: “physical changes are held to be independent of psychical, whereas psychical changes are declared to be their ‘collateral products’. They are called *collateral* products, or ‘epiphenomena’ to obviate the charge of materialism” McDougall (1911) roundly declares, referring to Huxley, that “to him [the doctrine] owes the name by which it is generally known; for he it was who suggested that the stream of consciousness should be called epiphenomenal, or the epiphenomenon of the brain-process.” In Carington (1949), H. H. Price expresses his belief that the term “epiphenomenalism” was introduced by T. H. Huxley.

It is interesting, therefore, that the term “epiphenomenalism” does not occur in Huxley’s (1874) essay on our topic; nor have I been able to find it elsewhere in his published work. (Neither does Huxley use the terms “stream of consciousness” or “brain-process”.) Of course, it is possible that Huxley made oral use of “epiphenomenalism” in lecturing. This seems unlikely, however, as he had at his disposal another brief term for the view he was concerned to promote, the meaning of which would have been more immediately accessible to most audiences, namely, “automatism”. This is the term that occurs in his 1874 essay, which bears the title “On the Hypothesis that Animals are Automata, and its History”. Besides containing the analogy of the steam-whistle that contributes nothing to the locomotive’s work, this essay compares consciousness to the sound of the bell of a clock that has no role in keeping the time, and treats volition as a symbol in consciousness of the brain-state cause of an action. As Ward correctly noted, nonefficacious mental events are referred to in this essay as “collateral products” of their physical causes. The essay is not solely concerned with animals: to the best of Huxley’s judgment, “the argumentation which applies to brutes holds equally good of men”.

Huxley and his contemporaries seem to have been impressed by preparations in which frogs had had various portions of their brains removed. Reasoning by analogy with humans lesioned by disease or battle, Huxley finds it plausible that the frogs are not conscious, or not exercising volition; yet when thrown into water, for example, they swim just as well as undamaged frogs. Huxley also discusses at some length the case of a Sergeant F., who had sustained a shot that fractured his left parietal bone. Once or twice a month, this soldier would have a day-long bout in which he exhibited complex

behavior (e.g., singing, writing a letter, “reloading”, “aiming”, and “firing” his cane with motions exactly appropriate to a rifle in a skirmish) while being plausibly unconscious, as evidenced by insensitivity to pins and shocks, sound, smell and taste, and to a great extent, vision. Huxley allows that there can be no direct evidence showing that the soldier is conscious or not conscious; but he concludes that he may be devoid of consciousness, while performing his complex and apparently purposeful movements.

Huxley was not alone among 19th century figures who gave vigorous and clear expositions of an epiphenomenalistic view. S. Hodgson (1870), W. K. Clifford (1874) and H. Maudsley (1886) were exponents of the view. Romanes’ posthumous (1896) contains an excellent statement of the view, which was first published in the early 1880s; and William James (1879) likewise offers an early clear statement of it. Both Romanes and James follow their statements of the view with arguments against its acceptance.

None of the works just mentioned include the term “epiphenomenalism”. I have located three articles in *Mind* in the 1890s that do use the term (the earliest, in 1893, hyphenates it as “epi-phenomenalism”). The earliest occurrence of the term for referring to automatism that I have been able to locate is in William James’s *The Principles of Psychology*, first published in 1890. It occurs in his chapter “The Automaton-Theory” once, in scare quotes; the rest of the time, the view is referred to as the “automaton-theory” or the “conscious automaton-theory”. James attributes the origination of the view to Shadworth Hodgson, in *The Theory of Practice* (1870). A section of this work titled “Dependence of consciousness on nerve movement” does indeed contain a forthright statement of the view (without “epiphenomenalism”, “automatism” or any other “-ism” tag).

Early in his discussion of automatism, James (1890) includes some remarks about his intellectual development, and refers to his early study of medicine. “Epiphenomenon” has a use in this field, meaning a symptom concurrent with, but not causally contributory to, the course of a disease. Some early twentieth century dictionaries list only this meaning of the term; by mid-twentieth century, the focal philosophical meaning is standardly given. My present surmise is that the term “epiphenomenalism” came into philosophy from medicine in the late nineteenth century, possibly, though less certainly, through William James’s use of the term in his influential *Principles of Psychology* (1890).

Bibliography

Two extensive bibliographies are available on line under entries for epiphenomenalism in (1) the PhilPapers’s bibliography (Other Internet Resources) and (2) *The Philosopher’s Index*. (The latter restricts entry to subscribers.) The following list contains all items referred to in the foregoing article, and a few other sources that offer particularly helpful discussions.

Antony, L., 2010, “Realization Theory and the Philosophy of Mind: Comments on Sydney Shoemaker’s *Physical Realization*”, *Philosophical Studies*, 148: 89–99.

Audi, P., 2012, “Properties, Powers, and the Subset Account of Realization”, *Philosophy and Phenomenological Research*, 84(3): 654–674.

Bailey, A., 2006, “Zombies, Epiphenomenalism, and Physicalist Theories of Consciousness”, *Canadian Journal of Philosophy*, 36: 481–509.

- Baltimore, J. A., 2010, "Defending the Piggyback Principle Against Shapiro and Sober's Empirical Approach", *Synthese*, 175(2): 151–168.
- Baumeister, R. F., Lau, S., Maranges, H. M. and Clark, C. J., 2018, "On the Necessity of Consciousness for Sophisticated Human Action", *Frontiers in Psychology*, 9: 1925. doi: 10.3389/fpsyg.2018.01925
- Baumgartner, M., 2010, "Interventionism and Epiphenomenalism", *Canadian Journal of Philosophy*, 40(3): 359–384.
- Beebe, H., 2017, "Epiphenomenalism for Functionalists", in H. Beebe, C. Hitchcock and H. Price (eds.), *Making a Difference: Essays on the Philosophy of Causation*, Oxford: Oxford University Press.
- Benecke, E. C., 1901, "On the Aspect Theory of the Relation of Mind to Body", *Aristotelian Society Proceedings* (New Series), 1: 18–44.
- Bieri, P., 1992, "Trying Out Epiphenomenalism", *Erkenntnis*, 36: 283–309.
- Bradley, M., 2011, "The Causal Efficacy of Qualia", *The Journal of Consciousness Studies*, 18(11–12): 32–44.
- Carrington, W., 1949, *Matter, Mind, and Meaning*, New Haven: Yale University Press. [H. H. Price edited this work, and wrote an introduction and notes for it.]
- Caston, V., 1997, "Epiphenomenalisms, Ancient and Modern", *The Philosophical Review*, 106: 309–363.
- Chalmers, D. J., 1996, *The Conscious Mind: In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- , 2010, *The Character of Consciousness*, Oxford: Oxford University Press.
- Clifford, W. K., 1874, "Body and Mind", lecture originally given to the Sunday Lecture Society, Nov. 1, 1874. Published in *The Fortnightly Review*, 16 (December): 714–736; reprinted in L. Stephen & F. Pollock (eds.), *Lectures and Essays of the late W. K. Clifford*, London: Macmillan, 1879.
- Davidson, D., 1970, "Mental Events", in Lawrence Foster and J. W. Swanson (eds.), *Experience and Theory*, London: Duckworth. Reprinted, with other relevant papers, in D. Davidson, *Actions and Events*, Oxford: Clarendon, 1980.
- De Brigard, F., 2014, "In Defence of the Self-Stultification Objection", *The Journal of Consciousness Studies*, 21(5–6): 120–130.
- Dennett, D. C., 1991a, "Real Patterns", *The Journal of Philosophy*, 88: 27–51.
- , 1991b, *Consciousness Explained*, Boston: Little, Brown.
- Descartes, R., 1649, *The Passions of the Soul*, Part I, art. xxxiv.
- Fodor, J. A., 1989, "Making Mind Matter More", *Philosophical Topics*, 17: 59–79.
- Haynes, J-D., 2013, "Beyond Libet", in A. Clark, J. Kiverstein and T. Vierkant (eds.), *Decomposing the Will*, Oxford: Oxford University Press.
- Hodgson, S., 1870, *The Theory of Practice*, London: Longmans, Green, Reader, & Dyer.
- Honderich, T., 1982, "The Argument for Anomalous Monism", *Analysis*, 42: 59–64.
- , 1983, "Anomalous Monism: Reply to Smith", *Analysis*, 43: 147–149.

- , 1984, “Smith and the Champion of Mauve”, *Analysis*, 44: 86–89.
- Horowitz, A., 1999, “Is There a Problem in Physicalist Epiphenomenalism?”, *Philosophy and Phenomenological Research*, 59: 421–434.
- Huxley, T. H., 1874, “On the Hypothesis that Animals are Automata, and its History”, *The Fortnightly Review* 16 (New Series): 555–580. Reprinted in *Method and Results: Essays by Thomas H. Huxley*, New York: D. Appleton and Company, 1898.
- Hyslop, A., 1998, “Methodological Epiphenomenalism”, *Australasian Journal of Philosophy*, 76: 61–70.
- Jackson, F., 1982, “Epiphenomenal Qualia”, *The Philosophical Quarterly*, 32: 127–136.
- James, W., 1879, “Are We Automata?” *Mind*, 4: 1–22.
- James, W., 1890, *The Principles of Psychology*, New York: H. Holt.
- Kim, J., 1993, *Supervenience and Mind: Selected Philosophical Essays*, Cambridge: Cambridge University Press.
- , 2005, *Physicalism, or Something Near Enough*, Princeton: Princeton University Press.
- , 2010, “Thoughts on Sydney Shoemaker’s *Physical Realization*”, *Philosophical Studies*, 148: 101–112.
- Kroedel, T., 2016, “Dualist Mental Causation and the Exclusion Problem”, *Nous*, 49(2): 357–375.
- Lalor, B. J., 1997, “It Is What You Think: Intentional Potency and Anti-individualism”, *Philosophical Psychology*, 10: 165–178.
- LePore, E. & Loewer, B., 1987, “Mind Matters”, *The Journal of Philosophy*, 84: 630–642.
- Lewis, D., 1988, “What Experience Teaches”, *Proceedings of the Russellian Society*, J. Copley-Coltheart (ed.), University of Sydney. Reprinted in W. G. Lycan, (ed.), *Mind and Cognition*, Cambridge, MA: MIT Press, 1990.
- Libet, B., 1985, “Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action”, *Behavioral and Brain Sciences*, 8: 529–566.
- , 2004, *Mind Time: The Temporal Factor in Consciousness*, Cambridge, MA: Harvard University Press.
- Linser, K and Goschke, T., 2007, “Unconscious modulation of the conscious experience of voluntary control”, *Cognition*, 104: 459–475.
- Lockwood, M., 1993, “The Grain Problem”, in H. Robinson (ed.), *Objections to Physicalism*, Oxford: Oxford University Press.
- Lyons, J. C., 2006, “In Defense of Epiphenomenalism”, *Philosophical Psychology*, 19: 767–794.
- Maudsley, H., 1886, *Body and Mind: An Inquiry into Their Connection and Mutual Influence, Specially in Reference to Mental Disorders*, New York: D. Appleton and Co.
- McDougall, W., 1911, *Body and Mind: A History and Defense of Animism*, London: Methuen.
- Megill, J., 2007, “Naturalism, Physicalism and Epiphenomenalism”, *Philosophical Psychology*, 20(6): 681–686.
- Mele, A. R., 2014, *Free: Why Science Hasn’t Disproved Free Will*, Oxford: Oxford University Press.

- Melnyk, A., 2010, "Comments on Sydney Shoemaker's *Physical Realization*", *Philosophical Studies*, 148: 113–123.
- Moore, D., 2012, "On Robinson's Response to the Self-Stultifying Objection", *The Review of Philosophy and Psychology*, 4: 627–641.
- , 2014, "The Epistemic Argument for Mental Causation", *The Philosophical Forum*, 45(2): 149–168.
- Nagasawa, Y., 2010, "The Knowledge Argument and Epiphenomenalism", *Erkenntnis*, 72: 37–56.
- Nahamias, E., 2008, "Is Free Will an Illusion? Confronting Challenges from the Modern Mind Sciences", in W. Sinnott-Armstrong (ed.), *Moral Psychology*, Cambridge, MA: MIT Press/Bradford.
- Nisbett, R. E. & Wilson, T. D., 1977a, "Telling More Than We Can Know: Verbal Reports on Mental Processes", *Psychological Review*, 84: 231–259.
- , 1977b, "The halo effect: evidence for unconscious alteration of judgments", *Journal of Personality and Social Psychology*, 35: 250–256.
- Pauen, M., 2006, "Feeling Causes", *Journal of Consciousness Studies*, 13 (1–2): 129–152. (This combined issue of the *Journal of Consciousness Studies* contains eight papers presenting a variety of positions regarding epiphenomenalism.)
- Pockett, S., Banks, W. P. & Gallagher, S. (eds.), 2006, *Does Consciousness Cause Behavior?*, Cambridge, MA: MIT Press.
- Popper, K. & Eccles, J., 1977, *The Self and Its Brain*, New York: Springer-Verlag.
- Putnam, H., 1975, "The Meaning of 'Meaning'", in K. Gunderson (ed.), *Language, Mind and Knowledge* (Minnesota Studies in the Philosophy of Science, Volume VII), Minneapolis: University of Minnesota Press. Reprinted in H. Putnam, *Mind, Language and Reality: Philosophical Papers*, Volume 2, Cambridge: Cambridge University Press.
- Robinson, D., 1993, "Epiphenomenalism, Laws and Properties", *Philosophical Studies*, 69: 1–34.
- Robinson, W. S., 1982a, "Why I Am a Dualist", in E. D. Klemke, A. D. Kline & R. Hollinger (eds.), *Philosophy: The Basic Issues*, New York: St. Martin's Press.
- , 1982b, "Causation, Sensations and Knowledge", *Mind*, 91: 524–540.
- , 1997, "Some Nonhuman Animals Can Have Pains in a Morally Relevant Sense", *Biology and Philosophy*, 12: 51–71.
- , 2006a, "Knowing Epiphenomena", *The Journal of Consciousness Studies*, 13: 85–100.
- , 2006b, "What Is It Like to Like?", *Philosophical Psychology*, 19: 743–765.
- , 2007, "Evolution and Epiphenomenalism", *The Journal of Consciousness Studies*, 14: 27–42.
- , 2012, "Phenomenal Realist Physicalism Implies Coherency of Epiphenomenalist Meaning", *The Journal of Consciousness Studies*, 19(3–4): 145–163.
- , 2013, "Experiencing is Not Observing: A Response to Dwayne Moore on Epiphenomenalism and Self-Stultification", *The Review of Philosophy and Psychology*, 4(2): 185–192.
- , 2018, "Russellian Monism and Epiphenomenalism", *Pacific Philosophical Quarterly*, 99: 100–117

- , 2019, *Epiphenomenal Mind: An Integrated Outlook on Sensations, Beliefs, and Pleasure*, New York and London: Routledge.
- Romanes, G. J., 1896, *Mind and Motion, and Monism*, London: Longmans, Green, and Co. [This book is an edition of material that first appeared in 1882 through 1886.]
- Russell, B., 1927, *The Analysis of Matter*, New York: Harcourt, Brace.
- Segal, G., 2009, “The Causal Inefficacy of Content”, *Mind and Language*, 24(1): 80–102.
- Shepherd, J., 2017, “Neuroscientific Threats to Free Will”, in K. Timpe, M. Griffith, and N. Levy (eds.), *The Routledge Companion to Free Will*, London and New York: Routledge.
- Shoemaker, S., 2007, *Physical Realization*, Oxford: Oxford University Press.
- , 2010, “Reply to My Critics”, *Philosophical Studies*, 148: 125–132.
- Smith, P., 1982, “Bad News for Anomalous Monism?”, *Analysis*, 42: 220–224.
- , 1984, “Anomalous Monism and Epiphenomenalism: A Reply to Honderich”, *Analysis*, 44: 83–86.
- Shapiro, L. A. and Sober, E., 2007, “Epiphenomenalism: The Do’s and the Don’ts”, in G. Wolters and P. Machamer (eds.), *Thinking About Causes: From Greek Philosophy to Modern Physics*, Pittsburgh: University of Pittsburgh Press.
- Soon, C., Brass, M., Heinze, H.-J., and Haynes, J.-D., 2008, “Unconscious determinants of free decisions in the human brain”, *Nature Neuroscience*, 11: 543–545.
- Staudacher, A., 2006, “Epistemological Objections to Qualia-Epiphenomenalism”, *Journal of Consciousness Studies*, 13(1–2): 153–175.
- Stich, S., 1983, *From Folk Psychology to Cognitive Science*, Cambridge, MA: MIT Press.
- Swinburne, R., 2011, “Could anyone Justifiably Believe Epiphenomenalism?”, *The Journal of Consciousness Studies*, 18(3–4): 196–216.
- Taylor, R., 1963, *Metaphysics*, Englewood Cliffs, NJ: Prentice Hall.
- Van Rooijen, J., 1987, “Interactionism and Evolution: A Critique of Popper”, *British Journal for the Philosophy of Science*, 38: 87–92.
- Walter, S., 2014, “Willusionism, epiphenomenalism, and the feeling of conscious will”, *Synthese*, 191(10): 2215–2238.
- Walton, M., 1989, “The Knowledge Argument Against the Knowledge Argument”, *Analysis*, 49: 158–160.
- Ward, J., 1902, “Psychology”, *Encyclopedia Britannica*, 10th edition, Volume 32. [Material quoted above appears in a section titled “Relation of Body and Mind: Psychophysical Parallelism” which did not appear in the 9th (1883) edition.]
- Wegner, D. M. and Wheatley, T., 1999, “Apparent mental causation: sources of the experience of will”, *American Psychologist*, 54: 480–492.
- , 2002, *The Illusion of Conscious Will*, Cambridge, MA: MIT Press/Bradford.
- Woodward, J., 2015, “Interventionism and causal exclusion”, *Philosophy and Phenomenological Research*, 91(2): 303–347.

—, 2017, “Intervening in the Exclusion Argument[Special-character]”, in H. Beebe, C. Hitchcock and H. Price (eds.), *Making a Difference: Essays on the Philosophy of Causation*, Oxford: Oxford University Press.

Wundt, W., 1912, *An Introduction to Psychology*, translated from the second German edition by R. Pintner, London: George Allen.

Academic Tools

[How to cite this entry.](#)

[Preview the PDF version of this entry at the Friends of the SEP Society.](#)

[Look up topics and thinkers related to this entry at the Internet Philosophy Ontology Project \(InPhO\).](#)

[Enhanced bibliography for this entry at PhilPapers, with links to its database.](#)

Other Internet Resources

- [Browse papers in Epiphenomenalism](#), at PhilPapers.org.

Related Entries

[consciousness](#) | [Descartes, René](#) | [dualism](#) | [emergent properties](#) | [functionalism](#) | [mental causation](#) | [multiple realizability](#) | [qualia](#) | [supervenience](#)

Copyright © 2019 by
William Robinson <wsrob@iastate.edu>