Sven Walter
University of Osnabrueck, Germany
Institute of Cognitive Science
s.walter@philosophy-online.de
www.philosophy-online.de

Review of *Mental Causation: The Mind-Body Problem* by Anthony Dardis. New York: Columbia University Press 2008.

That mental states can have a causal impact upon our behaviour—and *via* our behavior upon our environment—is an integral part of what one could, following Wilfrid Sellars, call our 'manifest image of the world.' It is a manifest fact of our everyday experience that we are agents who act *because* of our beliefs, desires, sensations, intentions, perceptions etc.: a headache makes us frown, the intention to make a bid at an auction causes us to raise our hand, and the desire to hear a loved one's voice leads us to make a phone call. *That* our mental life is causally efficacious is thus hard to deny. But exactly *how* can our mind make a causal difference to the movements of our physical body? Four centuries have passed since Descartes' pioneering discussion of this so-called 'problem of mental causation,' but we still lack a satisfying account of how our mental life fits into the causal structure of the physical world. *That* there is mental causation is part and parcel of our self-conception, but the *How* of mental causation constitutes a serious philosophical problem.

Solving the problem of mental causation requires showing how the mental fits into the causal structure of an otherwise physical world in such a way that we are able to exert a genuine causal influence upon it. However, there are at least three philosophical problems that seem to render such an account impossible. First, causation seems to require *laws*, while there are grounds for denying the existence of appropriate laws connecting the mental and the physical (the 'Argument from the Anomaly of the Mental'). Second, causation is arguably a *local* or *intrinsic* affair, while in the case of, for instance, beliefs and desires, those aspects constitutive of them *qua* mental are arguably *relational* or *extrinsic* (the 'Argument from Anti-Individualism'). Third, we do not understand how the mental can be causally relevant without coming into conflict with other parts of the causal structure we suspect to play an indispensable causal role in the production of physical effects (the 'Argument from Causal Exclusion').

In his recent book *Mental Causation: The Mind-Body Problem*, Anthony Dardis discusses the first and the third problem. His key claim is that the mental—mental *properties*, in particular—can make a causal difference to the physical world in virtue of the fact that there are *laws of nature* connecting the mental and the physical.

Chapter 1 briefly introduces the problem of mental causation and connects it to other issues from the philosophy of mind like the mind–body problem or the problem of free will. Chapters 2 to 4 contain a–highly interesting and illuminating–digression into the historical roots of the problem of mental causation. Chapter 2 covers Plato and Aristotle. Chapter 3 deals with Descartes (who firmly believed in mental causation but could not explain how it could possibly be squared with his dualism) and Huxley (also a convinced dualist who, in contrast to Descartes, saw that dualism and mental causation are hard to reconcile and therefore sought refuge in *epiphenomenalism*—i.e. the claim that the mental is distinct from and caused by the physical, but does not itself cause anything). Chapter 4 addresses the way the problem of mental causation has shaped the twentieth century debate about the mind–body problem. It covers the Wittgensteinian and Rylean view that mentalistic explanations of a person's behavior are, despite appearances, not causal explanations at all, as well as that identity theory and functionalism. According to Dardis, neither of these three positions is fully convincing: the Wittgensteinian and Rylean view denies the causal relevance of mental

properties and thus contradicts a basic fact of experience; the identity theory denies ontological independence of mental properties, which is equally untenable, and while functionalism preserves the ontological independence of mental properties, it is unable to account for their causal relevance.

While the first two points are rather obvious, Dardis' argument against functionalism's (understood as *role*-functionalism; what is known as '*filler*-functionalism' collapses, he maintains, into a version of the identity theory) ability to account for the causal relevance of mental properties is interesting. If, Dardis argues, according to functionalism a given mental property is *defined* as the property that stands in a web of causal relations to other properties, then functionalism cannot, in principle, provide a substantial account of why that property enters into these causal relations. (As Dardis acknowledges, this argument is reminiscent of the so-called 'logical connection argument' raised by followers of the Wittgensteinian and Rylean tradition against proponents of the causal theory of action.)

In order to solve the problem of mental causation, Dardis argues, one must show that mental properties, albeit ontologically independent, are not 'screened off' or 'excluded' from causal relevance by physical properties. This is the third of the three problems identified above—the Argument from Causal Exclusion. The problem, allegedly, is this: mental properties appear to be distinct from physical properties; they may depend upon physical properties, but at least from a first person point of view they seem to be something 'over and above' these properties, something not reducible to them. Yet, a scientific worldview seems to require that the physical world is *causally closed* in the sense that physical effects can be fully explained without recourse to non-physical entities or forces. Since our bodies are a part of the physical world, this also holds for the alleged bodily effects of our mental states. Yet, if it is always possible–at least in principle–to account for our behavior in purely physical terms, and if the mental is distinct from the physical, then, it seems, the mental does not contribute to the production of our behavior. There just seem to be 'no gaps' in the physical causal nexus that the mental could use to unfurl its own causal relevance—as a result, it is 'screened off' or 'excluded' from causal relevance by the physical.

On the one hand, as long as one acknowledges that mental and physical properties are distinct, the Argument from Causal Exclusion seems to deprive mental properties of causal relevance. On the other hand, denying the ontological independence of mental properties seems equally untenable. How can this tension be resolved? Dardis' solution is to argue that mental properties are indeed ontologically independent of, and not reducible to, physical properties, but that they are not entirely distinct from them either, not at least in a way that would allow the intuition of 'screening off' or 'exclusion' to get off the ground. In order to defend this claim, chapters 5 to 7 develop a metaphysical account of properties, causation, and laws. According to this account, mental properties *supervene* upon physical properties, and are thus asymmetrically necessitated by them: they are, although supervenient upon physical properties, ontologically independent and yet not entirely distinct from them; rather, mental and physical properties *overlap* in quite a literal sense.

According to Dardis, to every set of actual and possible entities there corresponds a property. Properties thus reduce to sets, and they can be said to overlap because sets can be said to overlap, viz., they can have a non-empty intersection. On Dardis' account, a physical property upon which a given mental property supervenes turns out to be a proper subset of the mental property. Hence—and this is the solution to the Argument from Causal Exclusion developed in chapter 8—although mental properties are ontologically independent, they are not completely distinct from the physical properties in their supervenience base. Therefore, mental properties cannot be said to compete with the physical properties in

their supervenience base for causal relevance, and without competition there can be no 'screening off' or 'exclusion'.

So far, as Dardis observes entirely correctly, this shows only that mental properties are not 'screened off' or 'excluded' from causal relevance by physical properties. Whether mental properties actually *are* causally relevant is left open. To establish that mental properties actually are causally relevant is the burden of chapter 9. There Dardis argues that mental properties are causally relevant in virtue of the same fact that renders physical properties causally relevant (cf. chapters 6 and 7): they figure in laws of nature. If Dardis were right, this would not only show that mental properties actually are causally relevant, it would also solve the first of the problems of mental causation mentioned in the beginning, viz., the one based on the Argument from the Anomaly of the Mental.

The gist behind this (Davidsonian) argument is that the essentially holistic and rational nature of the mental entails that there can be no strict laws connecting mental properties with other mental properties or with physical properties because mentalistic laws are necessarily hedged by *ceteris paribus* clauses. However, if causation requires strict laws, and if there are no strict mentalistic laws, then mental causation seems to be impossible. In chapter 9, Dardis argues that the fact that mentalistic laws are *ceteris paribus* laws does not prevent mental properties from figuring in laws of nature. In a nutshell, his solution is the following: the fact that mental properties supervene upon physical properties ensures that for any possible exception which would render a mentalistic law non-strict, one can find physical factors that—once included in the antecedent of the law—prevent exactly that exception from happening. In other words: we can render a mentalistic law strict by explicitly ruling out all the possible conditions that could render the antecedent of the law true and its consequent false. Hence, there is no reason to think that mental properties cannot occur in strict laws, i.e., laws of nature. Together with the claim that figuring in a law of nature is what renders a property causally relevant (cf. chapter 6), this entails, Dardis argues, that mental properties actually are causally relevant.

I have three major qualms with Dardis' argumentation. First, even if it is granted that we can somehow transform mentalistic laws into strict laws, Dardis' attempt to vindicate the causal relevance of mental properties by appeal to such laws seems bound to fail. There are causal laws and non-causal laws, and it is a legitimate question to ask what distinguishes causal from non-causal laws. For those who have an independent notion of causal relevance (i.e., one that does not attempt to characterize causal relevance in terms of laws), this question has an easy answer: laws that mention causally relevant properties are causal laws, laws that do not are non-causal laws. This answer, however, is not available to someone who, like Dardis, tries to analyze causal relevance in terms of laws, and it is hard to see what other answer might do the job. Yet, unless one has a criterion that distinguishes causal from non-causal laws, it is obviously futile to argue that properties are causally relevant in virtue of figuring in laws—since those properties that figure in non-causal laws are not causally relevant.

Second, Dardis' response to the Argument from Causal Exclusion seems also unconvincing: even if mental properties partially overlap with physical properties and are thus not entirely distinct from them, it remains a fact that there is nothing to do, causally speaking, for mental properties, once the physical properties of the objects in the world are fixed, and that seems to render mental properties rather irrelevant.

Third, I cannot see why Dardis first offers a (not at all uncontroversial) metaphysical account of properties in order to show that mental and physical properties do not compete for causal relevance, only to then argue that mental properties are causally relevant in virtue of figuring in appropriate laws. In my eyes, the second step makes the first step unnecessary: if mental properties can

indeed be causally relevant in virtue of figuring in appropriate laws, then, if there are such laws, of course they cannot be 'screened off' or 'excluded' by physical properties. For laws at the physical level do obviously not 'screen off' or 'exclude' laws at the mental level. Hence, arguing for the nomological conception of causal relevance would be sufficient to drive home the point Dardis tries to make, even without his independence-without-distinctness-argument.

That said, let me stress that—although the devil is, as so often, in the details—Dardis has written an extremely clear and well-argued book which is valuable reading to both the uninitiated and the expert. I highly recommend it to anyone interested in the metaphysics of the mind in general and the problem of mental causation in particular.