

I.

A BRIEF HISTORY

The problem of mental causation in its old form originates with René Descartes' *Meditations* (1641). In his Sixth Meditation, Descartes distinguishes two substances, mind and matter, and claims that these interact. But Descartes is also the champion of 17th century mechanistic thinking, and substance dualism (at least the Cartesian interactionist kind) does not sit well with mechanism. For if the motions of our bodies can be explained mechanistically, then what work is there for that strange, non-extended, invisible substance that is the mind? Substance dualism also requires at least one spot where mind and body get in touch: a Cartesian pineal gland. But as far as we know, our real pineal glands are not like that, nor is any other organ in our bodies. Finally, substance dualism implies that mind-interactions continuously perturb the laws of physics, in doing so undermining the very possibility of such laws, and making a complete mystery of the indubitable victory march of physical science ever since Newton.

Little wonder, then, that Cartesian substance dualism is now a marginal view and that contemporary philosophy of mind is largely physicalistic. Physicalism has put a ban on mind-stuff, immaterial forces, or the 'ghost in the machine,' as Gilbert Ryle called the Cartesian mind (1949). But banning ghosts is not yet banning minds. Minds need not be thought of as stuff- or force-like to begin with: with Ryle we may well regard doing so a category mistake. Very few philosophers are ready to deny that in our world of physical objects and forces there are thoughts, desires, hopes, assumptions, feelings, and so on. There may not be mental substances or energies, it is said, but certainly there are mental properties, or states of affairs. One may be a substance monist while being a property dualist, or pluralist.

Still, mental causation is waiting to be accounted for. If our actions did not have mental causes, most of our ordinary assumptions about human action would be radically false. Only by assuming that something like intentions cause actions can we distinguish, say, a vile kick from an involuntary fit, or a wink from a blink.¹ Donald Davidson has plausibly pointed out (1963) that when we explain someone's actions by citing her putative reasons (in terms of her beliefs, expectations, wishes, etc.)—reasons that justify,

¹ Example taken from Juarrero 1999.

or at least motivated the actions—we give a causal explanation. A question like ‘Why did she do it?’, Davidson claims, concerns a cause, and not just a rationalization. And indeed, why should we be interested in someone’s motives, will, beliefs, etc., if they do not make a difference to that person’s behaviour?

Davidson has put forward a doctrine, called Anomalous Monism,² that seemed to take mental causes into account. The ‘monism’ amounts to the claim that there are only (causally related) physical events. The ‘anomalism’ amounts to the claim that while causation is a matter of strict physical law, there are no psychophysical and no psychological laws. Still, according to Anomalous Monism, mental events can be causes of other mental events and of actions, because they are, after all, physical events. We can subsume such events under a causal law as long as we describe them in physical terms; and we can evaluate such events as reasons as long as we describe them in mental terms. But the events described are just the same events.

Anomalous Monism split with the reductionistic mind-brain identity theories put forward in the years around 1960 by Feigl, Place, and Smart. Identity theory had already been under criticism for a time, especially since Hilary Putnam (1968) had argued that mental events like pains might well occur in brains of wildly heterogeneous physical constitutions. What Anomalous Monism did was to rescue the mind-body identity claim while rejecting mind-body, or mind-brain reductionism.

Anomalous Monism seemed to many an attractive (that is, respectably physicalist) alternative to reductionism. Like reductionism it seemed to avoid the dead end of Cartesian dualism, but in contrast to reductionism it gave the mind an autonomous status in the order of things. In particular, it recognized the fact that the domain of the mental is essentially normative, and that agents choose their actions by weighing reasons rather than blindly conforming to regularities. Anomalous Monism thus seemed to allow one to have physicalism and mental autonomy at the same time.

Davidson’s critics, however, most notably Jaegwon Kim,³ have pointed out that Anomalous Monism does not in fact leave any causal autonomy for the mind. They accused it of implying that mental properties are epiphenomenal: produced by physical reality, but causally redundant themselves. For consider: if mental events are physical events, and the physical properties of those events suffice to cause other events—because, after all, these are what the laws of physics are about—, then what causal work is there left to do for the mental properties?

² See his 1970, 1973 and 1974.

³ See, for instance, his 1984a, 1985 and 1995a.

Davidson himself has always refused to see a problem here. He is a nominalist about properties, and a realist only about particulars (here: events). So the question whether it is 'in virtue of' physical or mental properties that one event causes another does not even arise for him.⁴ Most others, however, do believe that the question makes good sense.

We might consider Descartes and Davidson as two classic authors around whose claims the problem of mental causation has traditionally been centered; that is why I have started with them. But the current debate on mental causation is very complex and many-sided, and mental causation is considered problematic in more than one way. In the rest of this chapter, my aim is to map the problem area, to formulate the challenges that have to be met, and to give an outline of the book.

Here is already the general idea. As we will see below, there are no less than five problems of mental causation. These are distinct problems, but they are also closely related, which is why we cannot hope to solve them one at a time. Only a comprehensive theoretical framework that is able to cope with all five issues will help us out. Such a framework has to cover issues in ontology, the philosophy of causality, and the philosophy of mind, and these fields correspond to the three parts of my project.

Thus, in Part I, I develop the ontological foundations for my main arguments; in Part II, I develop a view of causality; and in Part III, I discuss how the results apply to mind and action. Because from the literature we can conclude that causality is a confusing and ill-understood phenomenon, I give particular attention to this topic. In my view, misconceptions about causality are the main obstacle to an understanding of mental causation. Yet, in received views about mental causation we encounter a theory of mind that is closely intertwined with specific ideas about causality, and only by undoing the entire knot can we come to see things in the right perspective. This is why I will also criticize a number of current views in the philosophy of mind.

My general aim, then, is to develop a theoretical framework that can serve as an alternative to the received one, a framework in which mental causation is at home, rather than an unwelcome guest. All this may seem rather pretentious, but as will become clear, most of the work has already been done by others. The great ideas that help us understand the issue are already on the air; I combine and adjust them, and add a little of my own.

⁴ See in particular his 1995.

II.

THE FIVE PROBLEMS OF MENTAL CAUSATION

Instead of one, there are at least five distinct problems of mental causation: five distinct though closely related ways in which standard assumptions and intuitions about mind (on the one hand) and causality (on the other hand) seem to clash. They are problems, because they suggest that much of our philosophical, scientific and ordinary-life discourse is in radical error: that, contrary to what any sane person assumes, it does not matter one bit what we think. Before trying to do something about this embarrassing situation, let us have a look at each problem in turn.

II.I. MENTAL ANOMALISM

Our first problem of mental causation is the fact that, while our beliefs and desires are among the causal explanantia of our actions, in ascribing these intentional states to each other in order to explain such actions (i.e., in making sense of each other as agents) we are doing something very different from applying laws, as in physics.

Let us first take a look at causation. Causation and natural law seem to be intimately linked. Suppose that *c* caused *e*. If so, it cannot be the case that *e* merely followed *c*, for mere temporal succession just is not causation. It seems that in the circumstances, *e* would not have happened if *c* had not happened, barring pre-emption. That is, causation seems to be a matter of counterfactual dependence among distinct events,⁵ the only possible source of which seems to be natural law. True, many empiricist philosophers mistrust counterfactuals and believe, in the wake of Hume, in regularities only. On their view, *e*'s being caused by *c* depends on the principle of 'same cause, same effect.' But note that just like the counterfactualist approach, the regularity approach treats causation as a matter of natural law: it only has a different conception of natural law! And indeed, something must distinguish

⁵ Counterfactual dependence is among propositions. But following Lewis (1973, 199), I will below speak of counterfactual dependence among states of affairs, events, etc.: 'Counterfactual dependence among events is simply counterfactual dependence among the corresponding propositions.'

causal from noncausal successions, the notion of natural law being the only 'something' plausibly invoked.

Now let us look at intentional states, the causes of our actions. Davidson has argued that the ascription of such states is essentially a holistic affair. For beliefs and desires can only be ascribed in large clusters that are constituted by normative connections such as 'coherence, rationality, and consistency' (1974, 231). Only insofar as specific intentional states (such as, for instance, the belief that Jones is a good sport) are embedded in such clusters, it seems, can we determine whether they are true, appropriate, or reasonable, and whether it makes sense to ascribe them in the first place.

This mental holism has been Davidson's main consideration in his rejection of psychological or psychophysical laws: the rational connections among intentional states do not seem to have a counterpart in physical theory. But apart from that, rational thought and responsible agency are not easily squared with blind, anonymous and indifferent laws of nature anyhow. Do we not have some degree of autonomy and freedom? When the telephone rings, most of us can be reliably predicted to go and answer it; yet, we could all just decide not to do so at any time. We mostly drift along with the daily course of events, but we do not seem to be at its mercy (if we come to feel that we are, we had better find help). If we cannot act according to choices that are genuinely our own, our very status as agents is compromised. But if decisions and actions are to be genuinely our own, they are not supposed to be determined by mere natural law, as it were behind our backs.

It might be thought that the way we deliberate is still highly predictable, because sane deliberation is subject to rational constraints, or rules, and as agents we all want to be rational. Might then lawfulness and acting for reasons not still coincide? Little chance, for practical wisdom is typically not codifiable. If it were, we could all use handbooks or pocket computers for practical thinking (for instance to solve ethical dilemmas), and we would not have to decide for ourselves. But that would be something for brainwashed zealots rather than for responsible agents.

But is it suggested that our minds elude the laws of nature? Does the above imply that a science of human beings impossible? No, such claims are preposterous. The relation between natural law on the one hand and thought and agency on the other hand is bound to be subtle and indirect. But it remains a fact that while our beliefs and desires are causes of our actions, they do not seem to be such in virtue of nomically necessitating them (or nomically raising their probability). But given that cause and law are inseparable, how can this be? This is our first problem of mental causation.