

# Introduction

It seems undeniable that our mental life makes a difference, sometimes a big difference, to our bodily life. What we think, what we believe, what we want, what we feel affects what we do with our bodies. I add salt to the sauce because I think that will make it taste better, I water my plants because I want them to grow, I take off my shoes because my feet feel sore, I wince because I remember an embarrassing mistake, I speak hesitantly because I feel nervous. Ordinary experience seems to suggest that what we do with our bodies causally depends, somehow, on what's going on in our minds.

How to understand the causal aspect of the mind–body connection is the subject of this book. Many philosophers have thought that our ordinary experience shows that there is causal interaction between mind and body, or that changes in one cause changes in the other. However, problems start to arise when we try to understand how this could be, given certain assumptions about the nature of reality. For example, suppose you thought, as Descartes did, that the mind is not a material thing. Instead, it is the immaterial part of ourselves that thinks and which is joined with our body but nevertheless distinct from it. If you also assumed that causal interaction could only occur between material things, perhaps because you thought all causal interaction required some kind of physical contact, then it becomes hard to see how mind–body causal interaction is possible. How can the mind have causal effects in the material world if it is not itself material?<sup>1</sup>

In contemporary philosophy of mind, putting together a plausible account of the mind–body connection remains a significant challenge. Philosophers of mind strive to give an account of what the mind is that allows mentality to have causal relevance but which also fits with the most plausible views of what causation in the actual world must be like. This is the problem of mental causation.

---

<sup>1</sup> This is the most famous objection levelled at Descartes's dualist metaphysics. See Shapiro (2007: 62) for Princess Elisabeth's version of this objection.

---

## How to cite this book chapter:

White, Andrea. 2024. *Understanding Mental Causation*. Pp. 1–9. York: White Rose University Press, DOI: <https://doi.org/10.22599/White.a>. License: CC BY-NC 4.0

Physicalism is the modern anti-Cartesian theory of what the mind is. This view says that everything that exists is either itself a physical entity or somehow constituted by, composed by, or exhaustively determined by physical entities. The main draw of physicalism about the mind is that it seems, at first, to easily solve the problem of mental causation. Physicalism says that when we talk about someone's mental life we are actually talking about physical states, properties or events, so mental causation reduces to causation by certain physical states, properties or events. In its crudest form, this kind of physicalism says that mental states and events are neural states and events, and mental causation is causation by neural states and events.

As it turns out, the most popular kind of physicalism has difficulty dealing with the problem of mental causation. Most contemporary philosophers of mind who call themselves physicalists accept some form of 'non-reductive physicalism'. On this view, that which is mental is not *identical* with anything physical; nevertheless, physical states, events and properties realise, constitute, compose or exhaustively determine mental states, properties and events. This kind of physicalism is thought to be difficult to reconcile with the principle of the causal closure of the physical world, which says that 'at every time at which a physical event has a cause it has a sufficient physical cause' (Gibb 2013: 2). As Jaegwon Kim (2005) argues, if some physical events have mental causes and those mental causes are not identical with any physical entities (as non-reductive physicalism maintains), then these physical effects must be overdetermined or the principle of causal closure must be false.<sup>2</sup> Since this objection was raised, non-reductive physicalists have offered many counterarguments aiming to show that their version of physicalism can save the phenomenon of mental causation while respecting causal closure. For example, Karen Bennett (2003), Sydney Shoemaker (2013) and Steinvör Thöll Árnadóttir and Tim Crane (2013) argue that both mental entities and the physical entities that realise them can be causally efficacious without this being a case of 'double-causing' anything like the paradigmatic cases of overdetermination. The debate about whether non-reductive physicalism can solve the problem of mental causation or if a fully reductive version of physicalism is required is ongoing.

The aforementioned debate notwithstanding, physicalism remains a popular metaphysics of mind because it appears to be the only metaphysics of mind that can (a) permit mental causation and (b) respect plausible principles about what actual causation is like, such as the principle of causal closure. This argumentative strategy underlies the main argument for physicalism about the mind, which is known as the causal argument for physicalism. Debates within philosophy of mind tend to centre on which kind of physicalism gives the best reconciliation between (a) and (b). Non-physicalist alternatives are generally thought incapable of giving any kind of reconciliation at all. In this way, contemporary philosophy of mind is shaped by this question: how is it possible for

---

<sup>2</sup> See also Crane (1995) and Heil (2013).

the mental to causally interact with the physical, especially given the apparent physicality of causation?

However, I believe that this is the wrong question to ask. I believe that contemporary philosophy of mind labours under a misapprehension of what mental causation is.

In most discussions of the problem of mental causation, mental causation is presented as a cause–effect relation between mental and physical entities. In many cases, mental causation is presented as a causal relation between mental and physical *events*. Sometimes mental causation is presented as a causal relation that can hold between *states*. Less frequently, mental *processes* are mentioned. Usually, events, states and processes are thought of as being very similar in nature, so that there is no need to treat mental events, mental states and mental processes differently when considering their candidacy as causal relata. In most discussions of mental causation, events, states and processes are thought of as three subclasses of the same general ontological category. Members of this general ontological category—I will call them *items*—are typically thought of as *particulars*, where particulars are unrepeatable, concrete individuals. So, even where mental causation is not presented as a causal relation between events—or not only between events—it is still presented as a causal relation between items that are mental and items that are physical. I call this understanding of mental causation the relational understanding of mental causation:

**Relational understanding of mental causation:** mental causation is mental items (events, processes or states) standing in causal relations to physical items (e.g. movements of a person's body).

Central to the relational understanding of mental causation is the idea that mental causation is a cause–effect relation between mental and physical items; mental phenomena are thought of as links in causal chains. This is the understanding of mental causation that has become standard in philosophy of mind but which I think is misconceived.

I believe that the relational understanding of mental causation is presupposed in many debates within philosophy of mind because of a triad of philosophical theories: (1) physicalism, (2) causal theories of intentional action and (3) relational approaches to causation. Although these theories are logically independent and about distinct philosophical questions, in practice they are mutually reinforcing. The relational understanding of mental causation presupposed by most arguments for physicalism is made to seem indispensable because of causal theories of intentional action, which in turn owe much of their apparent plausibility to relational assumptions about causation, assumptions that physicalists are likely to make. I believe this triad of views has limited our thinking about mental causation and therefore prevented us from exploring more diverse accounts of the relationship between our mind and body. I call this triad of views *the physicalist triad* because the upshot of endorsing each

element of the triad is that physicalism becomes the only acceptable metaphysics of mind as it appears to be the only view that has a chance of saving the phenomenon of mental causation. The aim of this book is to try to break out of this triad in order to open up new ways of understanding mental causation and thereby refresh debates within philosophy of mind.

I am not the first to suggest that there are connections between philosophy of action, philosophy of causation and physicalism. E. J. Lowe (2008) argues that physicalist consensus in philosophy of mind prevents and undermines a powerful account of rational agency. Jennifer Hornsby (2015) also argues that neo-Aristotelian theories of action—the main rivals to causal theories of action—call into question the existence of the kind of mental causation that forms the subject of debate in philosophy of mind, and hence have consequences for causal arguments for physicalism. However, Hornsby points out that ‘none of this work appears to have made any impression upon work in mainstream philosophy of mind’ (2015: 133). I suspect this is because the connection between physicalism, causal theories of intentional action and relational approaches to causation has not been sufficiently explicit to those working within philosophy of mind. Furthermore, no-one has provided reasons to persuade someone dissatisfied with the causal argument for physicalism that their best strategy for resisting the conclusion of this argument is to use lessons from philosophy of action and causation to question the foundational assumption of the causal argument. This is what I intend to do.

The arguments I put forward here will be of interest to those who are sceptical of physicalism as a metaphysics of mind but also feel dissatisfied with the standard counterarguments to physicalism. What I offer here is a distinctive non-physicalist approach to the problem of mental causation. However, I will not argue directly against physicalism. Ultimately, it is the relational approach to causation, and not physicalism itself, that does the most harm to our understanding of mental causation. Nevertheless, I hope to provide reasons to question physicalism’s hegemony as the metaphysics of mind that best accommodates mental causation.

In my view, the dominance of physicalism in philosophy of mind is not indicative of physicalism’s veracity. Instead, it ought to be something to make us suspicious. Physicalism is commonly thought of as the only naturalistic metaphysics of mind. Alternatives to physicalism are quickly criticised for rendering our mental lives inefficacious or for being at odds with scientific understanding. Physicalism has also become the theoretical backdrop for many of the kinds of questions discussed within contemporary philosophy of mind, such as: how do thoughts cause behaviours? what are the neural correlates of consciousness? how are mental entities and physical entities related if they are not identical? In this way, physicalism has prescribed what kinds of questions we ask about action, mental causation and the mind–body connection. This suggests to me that we need to interrogate the ideas about mental causation that contemporary philosophy of mind is taking for granted and which make physicalism seem like the only option.

In this book, I argue that physicalism's dominance, and the dismissal of non-physicalist alternatives as unnaturalistic or unscientific, depends on an understanding of mental causation that is not as theory-neutral as it first appears and relies upon (as it turns out) questionable assumptions about causation. My aim with this book is to provide a different, hopefully more philosophically neutral, description of the mental causation associated with intentional action. In this way, I hope to give us a fresh starting point for developing an alternative metaphysics of mind and for asking new questions about action, mental causation and the mind-body connection.

This book is divided into two parts. The first part explores the views that make up the physicalist triad. I explain how the three views are interconnected and provide evidence that, while logically independent, the views are mutually reinforcing. I also explain how these three views are responsible for the widespread acceptance of the relational understanding of mental causation. The philosophical topics discussed in these chapters will probably be familiar to the reader. However, it is my hope that by examining the interconnections between physicalism, causal theories of intentional action and relational approaches to causation I can reveal some important, but often unstated, assumptions made by these theories.

In Chapter 1, I outline physicalism in more detail and explain how arguments for physicalism presuppose the relational understanding of mental causation. I also explain how physicalism is supported by the other two elements of the physicalist triad.

In Chapter 2, I outline causal theories of intentional action. These theories have their roots in work by Donald Davidson. Davidson (1963) argues that when we say someone acted as they did because they wanted to do something, or because they believed that something was the case, we are giving a causal explanation. From this, Davidson concludes that states of desiring and states of believing—or at least events suitably related to states of desiring and believing, such as the onset of the desire or belief—are causes of the actions they explain. This argument has inspired the view that intentional actions are events that are caused by mental items. I explain how this view is used to justify the relational understanding of mental causation. I also argue that causal theories of intentional action owe much of their plausibility to relational approaches to causation.

In Chapter 3, I explain what a relational approach to causation is. A theory of causation is relational if and only if it is committed to the following thesis:

**Relationalism:** causation is always and everywhere a relation; the worldly phenomenon that is referred to by our concept 'causation' is not ontologically diverse in this respect.

The regularity theory of causation and David Lewis's (1973a; 1973b) counterfactual theory of causation are paradigm examples of relational theories of causation. However, there are many other examples.

In Chapters 4 and 5 I explain why I think we ought to challenge the physicalist triad. I do not argue directly against any of the theories that make up the triad. I do not argue that physicalism fails on its own terms, or that the causal theory of action cannot tell us what intentional action is, or that a relational theory of causation is impossible. Instead, I focus on what I take to be the weakest point of the triad, which is the account of agency it provides. In Chapter 4, drawing on arguments presented in philosophy of action, I argue against the physicalist/event-causalist description of agency provided by the physicalist triad. In Chapter 5, I offer a critical examination of some existing alternative theories of agency that appeal to the concept of agent causation or substance causation. I suggest that the chief failing of these theories is that they do not go far enough when it comes to rejecting the relational approach to causation.

In the second part of this book I show how broadening our understanding of causation, and more specifically incorporating the concept of *process* into our understanding of causation, opens up new ways of understanding intentional action and the mental causation associated with it. In this way, I hope to describe what a theory of mental causation can look like if the physicalist triad is rejected. I provide reasons to think that this alternative approach to causation allows us to develop a better understanding of intentional action and the mental causation associated with it.

In Chapter 6, I present my own non-relational approach to causation. My approach denies that causation is always a relation and holds instead that causation can be a process rather than a relation, of which processes like breaking, crushing, bending etc. are more determinate species. My proposal is that causation is on display not only when events make the difference to the occurrence of other events but also when substances engage in processes. I suggest that engaging in a process is analogous to instantiating a property, and that events are instances of processes.

In Chapters 7 and 8, I challenge Davidson's argument that states of desiring and states of believing are causes of the actions they explain. This argument has been challenged before. Philosophers such as Ludwig Wittgenstein (1958) and Elizabeth Anscombe (1957) rejected the idea that beliefs and desires stand to actions as causes to effects. They argued that concepts like *belief*, *desire* and *intention* do not refer to items that can stand in causal relations to actions or physical events. Similarly, Gilbert Ryle (1949) argued that 'mental conduct verbs'—like 'knowing', 'believing', 'intending' and 'desiring'—do not signify or denote inner causal events, so when such verbs are employed to explain why an agent acted they do not designate inner causes of the action they explain. This view, which I call the *non-causalist* view, denies that intentional action entails the existence of causal relations between mental items and physical events.

However, non-causalists reach this conclusion by arguing that explanations of intentional actions that cite beliefs or desires are not usually causal explanations at all, whereas I believe that explanations of intentional actions that cite the agent's beliefs or desires do give causal information. Fortunately, this kind

of intermediary view is made possible if one rejects the relational approach to causation. In Chapter 7, I argue that it is not necessary for an explanation to be causal that its explanandum designate an effect and its explanans designate an item that is the cause of that effect. My non-relational theory of causation implies that facts about causal relations are not the only causal facts that causal explanations could answer to. I suggest that some causal explanations are made true by the non-relational aspect of causal reality, that is, by facts about substances engaging in processes.

In Chapter 8, I argue that explanations of intentional action that cite the agent's reasons for acting are the kind of causal explanation that are not made true by causally related events. The most important consideration favouring this view is that it saves two strong intuitions: (a) that reason-giving explanations are causal and (b) that the mental states cited in reason-giving explanations do not denote items that stand in causal relations to the actions they explain. This view has important consequences for how we ought to think about the nature of intentional action. Most importantly, it casts doubt on the view that intentional actions are distinguished from non-intentional actions by their causes.

In Chapter 9, I propose an alternative view of intentional action. I propose that to act intentionally is to engage in a process, and as such is to exercise a power—but a power of a special sort. The power to act intentionally is a power to structure one's own activities so that they demonstrate a pattern—a pattern that is only revealed by attributing mental states to the agent. So, when an agent acts intentionally, they engage in the process of causation. The process they engage in counts as *mental* causation in virtue of the fact that the agent is manifesting a special power to organise their activities so that they instantiate a certain structure, a structure that is made comprehensible by the agent's mental states.

In Chapter 10, I revisit the problem of mental causation. If the arguments of the previous chapters are successful, then the existence of intentional action does not entail that mental items stand in causal relations to physical items. When we say that someone acted intentionally because of what she believed, desired, intended or decided, these mental concepts need not refer to items that stand in causal relations to physical events. Instead, it is possible to think of the mentality of the causal processes human beings engage in when they act intentionally to consist in the fact that these processes are part of a larger pattern of meaningful, or interpretable, activity. This means that the standard way mental causation is set up as a problematic subject in philosophy of mind may not be right. As explained above, debates within philosophy of mind tend to centre on which metaphysics of mind best reconciles the claim that mental items stand in causal relations to physical items with plausible principles about what actual causation is like, such as the principle of causal closure. However, if realism about mental causation does not require the relational understanding of mental causation at all, then the problem of mental causation as it is usually understood is a pseudo-problem. In Chapter 10, I discuss alternative ways



to understand mental causation and the consequences this has for philosophy of mind.

I think it is undeniable that our mental life makes a difference to our bodily life. I agree that what we do with our bodies causally depends on what's going on in our minds. However, I think it has been a mistake to assume that the causal aspect of the mind–body connection ought to be understood as causal interaction between mind and body. Descartes was wrong, I believe, to divide human beings into two distinct substances, mind and body. Modern philosophers of mind are similarly wrong to divide mental causation into a causal exchange between distinct aspects of ourselves.

### References<sup>3</sup>

- Anscombe, G E M 1957 *Intention*. Oxford: Basil Blackwell.
- Árnadóttir, S T and Crane, T 2013 There is no exclusion problem. In: Gibb, S C, Lowe, E J and Ingthorsson, R *Mental causation and ontology*. Oxford: Oxford University Press. pp. 248–266.
- Bennett, K 2003 Why the exclusion problem seems intractable and how, just maybe, to tract it. *Noûs*, 37(3): 471–497. DOI: <https://doi.org/10.1111/1468-0068.00447>
- Crane, T 1995 The mental causation debate. *Aristotelian Society Supplementary Volume*, 69(1): 211–236. DOI: <https://doi.org/10.1093/aristoteliansupp/69.1.211>
- Davidson, D 1963 Actions, reasons, and causes. *Journal of Philosophy*, 60(23): 685–700. DOI: <https://doi.org/10.2307/2023177>. Reprinted in Davidson 2001 pp. 3–20.
- Davidson, D 2001 *Essays on actions and events*. 2nd ed. Oxford: Clarendon Press.
- Gibb, S 2013 Introduction to mental causation and ontology. In: Gibb, S C, Lowe, E J and Ingthorsson, R *Mental causation and ontology*. Oxford: Oxford University Press. pp. 1–17.
- Heil, J 2013 Mental causation. In: Gibb, S C, Lowe, E J and Ingthorsson, R *Mental causation and ontology*. Oxford: Oxford University Press. pp. 18–35.
- Hornsby, J 2015 Causality and 'the mental'. *HUMANA.MENTE Journal of Philosophical Studies*, 8(29): 125–140.
- Kim, J 2005 *Physicalism, or something near enough*. Princeton, NJ: Princeton University Press.
- Lewis, D K 1973a Causation. *The Journal of Philosophy*, 70(17): 556–567. DOI: <https://doi.org/10.2307/2025310>

<sup>3</sup> Author note: some references to Davidson are formatted (1963/2001). This indicates the initial date of publication of the paper (in this case 1963) but references the paper as it appears in the 2001 collection of his essays, with the page numbers relating to that volume.



- Lewis, D K 1973b *Counterfactuals*. Oxford: Basil Blackwell.
- Lowe, E J 2008 *Personal agency: The metaphysics of mind and action*. New York: Oxford University Press.
- Ryle, G 1949 *The concept of mind*. London: Hutchinson's University Library.
- Shapiro, L (ed.) 2007 *The correspondence between Princess Elisabeth of Bohemia and René Descartes*. Chicago, IL: University of Chicago Press.
- Shoemaker, S 2013 Physical realization without preemption. In: Gibb, S C, Lowe, E J and Ingthorsson, R *Mental causation and ontology*. Oxford: Oxford University Press. pp. 35–57.
- Wittgenstein, L 1958 *The blue and brown books*. Oxford: Blackwell.