CHAPTER 3

The Relational Approach to Causation

I turn now to the third aspect of the physicalist triad: the relational approach to causation. 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 between distinct entities ('cause' and 'effect'); the worldly phenomenon that is referred to by our concept 'causation' is not ontologically diverse in this respect.

We have seen how the relational approach to causation lends plausibility to both physicalism and causal theories of intentional action. The driving force behind arguments for physicalism is the problem of mental causation, but the way mental causation is understood in these debates is heavily influenced by background assumptions about the nature of causation. Specifically, philosophers writing on the problem of mental causation assume that mental causation is a cause–effect relation where the cause relatum or effect relatum, or both, is a mental item (the relational understanding of mental causation). It is very difficult to imagine an alternative understanding of mental causation if you take a relational approach to causation. On this approach, 'cause' is an unequivocal term. All causation everywhere is the same, so the only thing that can discriminate between different categories of causation is the nature of the relata involved. The relational approach to causation also entails that causal reality is nothing more than a chain of causally related events, so, if intentional action is a causal phenomenon at all, it must be located within this worldview. This lends support to causal theories of intentional action that reduce the agent's role in bringing about what she intends to causation by mental events.

The relational approach to causation is not argued for by physicalists or those who propose a causal theory of intentional action. Instead, it is often taken for granted or treated as a harmless background assumption or

How to cite this book chapter:
White, Andrea. 2024. *Understanding Mental Causation*. Pp. 55–71. York: White Rose University Press. DOI: https://doi.org/10.22599/White.d.

pre-theoretical notion. I think this is incorrect, because I think the assumption that causation is always and everywhere a relation is not as innocuous as it seems. It is a substantive claim about the nature of causation. The purpose of this chapter is to show that the relational approach to causation is a substantive philosophical position, and not merely a harmless background assumption or pre-theoretical notion.

3.1 Hume's legacy

The relational approach to causation is not recognised as a substantive philosophical position because most philosophers working on causation accept relationalism, at least implicitly. Most have assumed that providing a theory of causation is a matter of explaining what a relation must be like to be a causal relation. In his Stanford Encyclopedia article on the metaphysics of causation, Jonathan Schaffer introduces this philosophical project with the following question: 'What must a world be like, to host causal relations?' (2016). He goes on to state that '[q]uestions about the metaphysics of causation may be usefully divided into questions about the causal relata, and questions about the causal relation' (2016). The majority of work on the metaphysics of causation proceeds as if Schaffer's taxonomy of questions concerning causation are the only questions we can ask about what reality must be like when causal statements are true. In J. Dmitri Gallow's Stanford Encyclopedia article on the metaphysics of causation, which replaced Schaffer's article, the metaphysics of causation is still described as the project of finding out 'what kind of relation [causal] claims are about' (2022). The possibility that causation may not fit into a single ontological category is rarely taken seriously.[11]

Relationalism is widely accepted in part due to the lasting influence David Hume has had on the philosophy of causation. Briefly examining Hume's influence on the philosophy of causation will help make it clear that, far from being pre-theoretical, relationalism has its roots in Humean theories of causation.

During the early modern period, the concept 'cause' underwent a transformation. Earlier Aristotelian and Scholastic ideas about causation were challenged, replaced, and abandoned, including the Aristotelian view that there are four modes of causation, or that 'cause' has four distinct senses. Hume concluded that 'all causes are of the same kind, and that in particular there is no foundation for that distinction, which we sometimes make betwixt efficient causes and causes *sine qua non*; or betwixt efficient causes, and formal, and material, and exemplary, and final causes' (1964: 171). Since Hume, philosophers of causation have come to regard efficient causation as the only mode of

[11] One notable exception is Helen Steward (2012: 212–216), whose consideration of the ontological heterogeneity of causal reality informs her understanding of agency.

causation there is: all causation is a matter of causes being that which produced effects. Aristotle's other modes of causation are not really causation at all; they are more accurately described as modes of explanation or modes of 'because' (Hocutt 1974). Following Hume, contemporary philosophy of causation rarely entertains the idea that 'cause' might be ambiguous. The univocality of the concept 'cause' is a key tenet of relationalism. Relationalism entails that, when we inquire about what reality must be like when true causal statements are made, there is just one sort of thing we are looking for—it is not the case that the reality causal statements answer to might vary depending on the context within which those statements are made.

Aristotelian ideas about substances and powers and how these concepts figure in causation were also challenged during the early modern period. As Walter Ott (2009) describes it, Aristotelian ideas about substances and powers were gradually replaced by laws of nature and a mechanist ontology (albeit in a messy, often piecemeal way), a development that abetted Hume's scepticism about the existence of a mind-independent necessary connection between cause and effect.

Hume wanted to know what the source or origin of our idea of necessary connection was and argued forcefully that we gain no impression of it when we observe a single instance of one type of event being followed by another. Hume argued that we experience 'one event follows another; but we never can observe any tie between them. They seem conjoined, but never connected' (1975: 74). Hume drew a similar conclusion with regard to powers: we observe 'an uninterrupted succession' but not any 'power or force which actuates the whole machine' (1975: 63); we can perceive what a thing is like but not what it is capable of doing. Hume argued further that we cannot perceive the operation of power even in cases where we ourselves are doing something or making something. Even in these cases, all we observe is a sequence of events. However, Hume argued, when we *repeatedly* experience events of one type being followed by events of another type, we come to *expect* an event of the second type when we experience an event of the first, and this internal feeling of expectation is the impression from which this idea of necessitation between cause and effect arises. On one interpretation, Hume's conclusion is that the idea of causation as necessary connection or the exercise of power is a product of our own minds, and what exists in mind-independent reality are unconnected events within which we can discern patterns of regularity.

This admittedly controversial interpretation of Hume has had a lasting influence over modern theories of causation.[12] The principle that cause and effect are distinct events and so there can be no metaphysically necessary connections between them, a principle sometimes known as 'Hume's dictum' (Wilson 2010), presents a challenge. If cause and effect are not joined by a necessitating

[12] See Beebee (2007) and Millican (2007) for good discussions on how Hume's claims about causation should be interpreted.

relation, how are they joined? *This* is the challenge contemporary theories of causation have focused on. As a result, the project of giving an account of the metaphysics of causation has become a matter of specifying the nature of the relation that joins cause and effect together. Relationalism is taken for granted by many contemporary theories of causation, and the ontologically richer views of causation entertained by Aristotelians and Scholastics rarely surface in modern theories of causation. However, the fact that Aristotelian ideas, such as the idea that there are different kinds of cause, stand opposed to relationalism shows that relationalism is not a pre-theoretical assumption about the nature of causation. Relationalism is the dominant theoretical position within contemporary philosophy of causation, but it is still a theoretical position.

3.2 Two relational theories of causation

The regularity theory of causation and David Lewis's counterfactual theory of causation are paradigm examples of relational theories of causation. Both theories hold that causation is a special type of relation between cause and effect. Both theories also attempt to spell out what this special relation is in non-causal terms. In that way, both theories offer a reductive account of causation. Briefly examining the metaphysical commitments of these theories will help make it clear what beliefs about causation are consistent with relationalism. It will also make it easier to articulate the alternative to relationalism in later chapters.

The regularity theory holds that causation, as it exists in the world independently of our thinking about it or knowledge of it, is exhaustively constituted by certain relations of spatiotemporal contiguity that obtain with regularity. More specifically, the regularity theory holds that causation is a relation of spatiotemporal contiguity between two events, *c* and *e*, where *c* occurs before *e*, and where all events of the same type as *c* are regularly followed by events of the same type as *e*. The regularity theory as stated above faces problems and, in response, more sophisticated versions of the theory have been proposed.[13] However, the simplest version of the regularity theory will suffice for my purposes here.

The main argument for adopting a regularity theory is that it offers a reductive account of causation where, as Stathis Psillos puts it, 'causal talk becomes legitimate, but it does not imply the existence of a special realm of causal facts that make causal talk true, since its truth conditions are specified in non-causal terms, that is, in terms of spatiotemporal relations and actual regularities' (2002: 4). The idea is that the regularity theory of causation—or at least a suitably worked-up version of it—provides everything we would want from a theory of causation, without positing the existence of powers or a *sui generis* kind of

[13] For example, Baumgartner (2008), Mackie (1974) and Mill (1843) have all offered more sophisticated versions of the regularity theory.

necessity. According to the regularity theory, what ascriptions of power, or statements about what a thing can do, actually *mean* (if they are not false or nonsense) is that the behaviour of the object to which the 'power' is attributed is regular in a certain way. That is, it might be true to say some object has a power, but what makes such a statement true will be some fact about the arrangement of the spatiotemporal mosaic of instantiations of intrinsic, qualitative, categorical properties.

The mosaic metaphor is how Lewis describes the metaphysics presupposed by the regularity theory. In more detail, this metaphysics says:

> [I]n a world like ours, the fundamental relations are exactly the spatiotemporal relations: distance relations, both spacelike and timelike, and perhaps also occupancy relations between point-sized things and spacetime points. And it says that in a world like ours, the fundamental properties are local qualities: perfectly natural intrinsic properties of points, or of point-sized occupants of points. Therefore it says that all else supervenes on the spatiotemporal arrangement of local qualities throughout all of history, past and present and future. (1994: 474)

As Lewis puts it in the introduction to his *Philosophical Papers (vol. II)*, 'all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another' (1986: ix). Jonathan Schaffer describes this worldview slightly differently: Schaffer writes that the world is 'history' i.e. 'the fusion of all events throughout space-time' (2007: 83).

Lewis's (1973a; 1973b) counterfactual theory of causation analyses causation in terms of counterfactual dependence. This theory exploits the intuition that causes are that which made the difference to the occurrence of the effect; that is, had the cause not occurred, the effect would not have occurred either. Lewis developed this idea by analysing the causal relation as the ancestral of a counterfactual dependence relation. So, an event *c* stands in a causal relation to another event *e* if and only if *e* counterfactually depends on *c*, or *e* counterfactually depends on an event that counterfactually depends on *c*, or *e* counterfactually depends on an event that counterfactually depends on an event that counterfactually depends on *c*, etc. As with the regularity theory, Lewis's counterfactual theory has been modified in light of objections raised against the original version, but again the simplest version of the counterfactual account will suffice for now.[14]

Lewis's counterfactual theory's status as reductive depends, in part, on Lewis's theory of modality. Lewis opts for a possible world semantics for counterfactuals. So, a counterfactual like 'if *c* had not occurred, then *e* would not have occurred' is true if and only if *e* does not occur at the closest possible

[14] See, for example, Ganeri, Noordhof and Ramachandran (1996), Lewis (2000), McDermott (2002) and Sartorio (2005).

world where *c* does not occur. How close a possible world is to the actual world depends on how *similar* that world is to the actual world. For Lewis, similarity between two possible worlds is determined by what particular states of affairs obtain at the two worlds and what the laws of two worlds are. So, world w^1 is more similar to world w^2 the more states of affairs w^1 has in common with w^2 and the more laws w^1 has in common with w^2.

If one went along thus far with Lewis's semantics for counterfactuals, but thought that laws of nature were brute facts about what powers things have, or facts about primitive 'necessitation' relations holding between universals, then even if one opted for an account of causation where causation is reduced to counterfactual dependence, the resultant theory of causation would not be reductive. This is because, on such a view, the truth of counterfactual conditionals depends on similarity rankings of possible worlds, which in turn depends on brute facts about powers, or a *sui generis* form of necessity. However, Lewis gives an account of laws of nature that does not presuppose the existence of powers or anything over and above the spatiotemporal mosaic of instantiations of intrinsic, qualitative, categorical properties. For Lewis, laws of nature are simply regularities that are deducible from axioms in an explanatory system that best balances simplicity and strength. An explanatory system picks as few general truths as possible to serve as axioms—the fewer, the simpler—then deductively derives further general truths from these. The more general truths the system deductively entails, the stronger the system. As Helen Beebee (2006) points out, because Lewis seeks to analyse causation without assuming the existence of any kind of worldly necessitation, and ends up turning to regularities in order to fulfil that mandate, Lewis's counterfactual theory of causation has a lot in common, metaphysically speaking, with the regularity theory. On both theories, the worldly structures that make true causal claims are, in the end, regularities. And, just like the regularity theory, Lewis's counterfactual theory does not posit any kind of entity or deeper fact (like facts about what powers things have or what is a natural necessity) that grounds or explains why regularities hold, or why certain counterfactual conditionals are true.

What is important to notice about these two theories is that they reject the idea that causation is (at least sometimes) the exercise of power or the making-happen of an effect in favour of describing causation in more ontologically sanitised terms, which they assume means describing causation in terms of a relation between elements of the 'spatiotemporal mosaic'. Many rivals to the regularity theory or the counterfactual theory of causation challenge the *reductive* aspects of these theories. That is, rival theories of causation challenge the principle that causation, as it exists in the world independently of our thinking about it or knowledge of it, is exhaustively constituted by non-causal states of affairs. However, the principle that causation is always a relation between cause and effect is not challenged.

For example, Galen Strawson articulates a conception of causation that he calls Causation with a capital 'C'. To believe in the existence of Causation is to believe: 'a) that there is something about the fundamental nature of the world in virtue of which the world is regular in its behaviour; and b) that that something is what causation is, or rather it is at least an essential part of what causation is' (1989: 84–85). Strawson thus advocates a view that takes causation to be an entity that grounds the world's regularities but cannot be reduced to regularities, or indeed any aspect of the 'spatiotemporal mosaic'. Although Strawson (1989) argues that causation, as it is in reality, is regular succession plus something extra, which explains why events unfold in a regular way, he is noncommittal on what this extra element is. Strawson (1987) suggests that this additional element could be the presence of 'objective forces—e.g. the "fundamental forces" postulated by physics' that 'govern the way objects behave and interact' (1987: 254), and adds:

> I will avoid speaking of 'natural necessity', or of 'laws of nature' (understood in a strong, non-Regularity-theory sense), or of the 'causal powers' of objects. It is very difficult to keep control of these rival terminologies. But here the notion of objective forces is being understood in such a way that accounts of causation given in terms of these other notions may be supposed to reduce naturally to the account in terms of forces. For example: (1) if objects have causal powers, they have the powers they do wholly in virtue of the nature of the forces informing (and so governing) the matter of which they are constituted. (1987: 255)

Michael Tooley (1990a) similarly argues against views that hold that 'causal relations are … logically supervenient upon non-causal properties and relations' (1990a: 217). The sort of causal realism that Tooley endorses treats 'causal concepts as theoretical concepts, so that causal relations can only be characterised, indirectly, as those relations that satisfy some appropriate theory' (1990a: 234). The appropriate theory, Tooley (1990b) proposes, is one that includes claims about the formal properties of causal relations, and which tells us what a law must be like to be a causal law. Causal relations are thus relations that have the right formal properties and 'whose presence in a law makes that law a causal one' (1990b: 303). Tooley shares Armstrong's view about laws of nature (of which causal laws are a subset); that is, he thinks that laws are necessitation relations between universals. So, it would seem that Tooley's account of causation, in virtue of its appeal to causal laws, makes use of a *sui generis* form of necessity.

The point I wish to emphasise is that both Strawson and Tooley are arguing specifically against attempts to reduce the causal *relation* to some non-causal *relation*. Strawson (1989) is concerned with showing that we should believe there is something more to the relation between cause and effect than regular

succession. Similarly, it is specifically 'realism with regard to causal relations' that Tooley considers (1990a: 233).

Interestingly, Tooley cites Elizabeth Anscombe as a philosopher who upholds a realist view of causation where causal relations are directly observable 'not only in the everyday sense of that term, but in a much stronger sense which entails that concepts of causal relations are analytically basic' (1990a: 233–234). Anscombe (1971) suggested that we come by our primary knowledge of causality when we learn to speak and come to associate the linguistic representation of a causal concept with its correct application. An example of such a causal concept that Anscombe provides is 'infect'. Others include 'scrape, push, wet, carry, eat, burn …' (1971: 9). She suggests that causal activities like scraping and pushing (though perhaps not infecting) are activities that we can directly perceive. Tooley ultimately rejects this form of realism. He argues that, even if Anscombe is right that we know by observation that one thing is pushing another (for example), this does not show that what it is about the events we are seeing that means they are causally related is something irreducible we can nevertheless observe. It might be that we infer, from what we perceive, that causation is there.

However, I think that Tooley has misconstrued what Anscombe is claiming in her 1971 lecture 'Causation and Determination', from which he cites. What Anscombe suggests we directly perceive is not a special *relation* between cause and effect but substances exerting causal power over other substances. We do not observe a cause causing an effect; we observe an agent acting on a patient. Anscombe is suggesting that an agent acting on a patient is causation, and this is in spite of the obvious truth that agent and patient are not related to each other as cause and effect. Anscombe's point is that we come by knowledge of causality when we directly perceive agents pushing patients and correctly associate what we see with the inherently causal concept 'pushing'. Tooley might be right that the fact that we directly perceive agents pushing patients (for example) may not be enough to show that we directly perceive a connection between the events that makes it the case that they are causally related. But why can't the fact that we directly perceive an interaction be enough to show that we directly perceive causation? Tooley construes Anscombe's claim incorrectly, I think, because of his commitment to a version of relationalism that says that causation is a relation between events.

3.3 Manipulability accounts of causation

Another important family of theories of causation is manipulability accounts of causation. Manipulability accounts of causation explore the intuition that causes are things in nature that we can manipulate and thereby alter outcomes. These theories connect causation to our sense of agency, to the idea of ourselves as beings which alter the course of events. Indeed, some manipulability

accounts explicitly define causation in terms of agency. For example, Georg Henrik von Wright argues that an event *c* is the cause of event *e* if and only if bringing about *c* is a way for an agent to bring about *e*, that is, only if *e* can be considered the result of the action of bringing about *c*:

> [T]o think of a relation between events as causal is to think of it under the aspect of (possible) action. It is therefore true, but at the same time a little misleading to say that if *p* is a (sufficient) cause of *q*, then if I could produce *p* I could bring about *q*. For *that p* is the cause of *q*, I have endeavoured to say here, *means* that I could bring about *q*, if I could do (so that) *p*. (1971: 74)

Similarly, Peter Menzies and Huw Price argue that 'an event [*c*] is cause of distinct event [*e*] just in case bringing about the occurrence of [*c*] would be an effective means by which a free agent could bring about the occurrence of [*e*]' (1993: 187) and an event *c* is an effective means by which a free agent could bring about occurrence of *e*, just in case the probability of *e* occurring given that *c* was brought about by a free agent is greater than the unconditional probability of *e* occurring.

In assigning a central role to human agency, these theories might seem to offer a richer account of causation, one that leaves room for the idea that causation could be something other than a relation between cause and effect; instead, it might be an activity (manipulation) that agents perform, or it might be the exercise of power where this is an irreducible feature of fundamental reality. However, closer examination of manipulability theories reveals that most are committed to relationalism.

A criticism levied against agency-based manipulability accounts is that they are problematically circular, because agency is a causal notion: *producing* and *bringing about* are causal concepts, hence agency-based theories purport to analyse causation in terms of causation. Von Wright responds to this objection by arguing that the relation between an action (e.g. cutting of the cake) and its result (the cake's coming to be cut) is *not* a causal relation; it is rather a logical one (if the cake does not come to be cut, then no-one cut it—the cutting-of-the-cake action did not take place):

> I am anxious to separate agency from causation. Causal relations exist between natural events, not between agents and events. When by doing *p* we bring about *q*, it is the happening of *p* which causes *q* to come. And *p* has this effect quite independently of whether it happens as a result of action or not. (1974: 49)

I think von Wright is right to sharply distinguish between agency on the one hand and causal relations on the other—he is correct that to demonstrate agency is not for an agent to stand in a causal relation to an event. However, I

do not think, as von Wright does, that this entails that agency is not a causal phenomenon. Von Wright does not recognise this because he subscribes to relationalism, the view that causation is always, everywhere a relation. Von Wright's view can be thought of as abiding by the following reasoning: causation is the relation between cause and effect; agency is not a relation between cause and effect; therefore, agency is not causation. This argument is sound only if relationalism is true. So, von Wright accepts relationalism.

The circularity objection can be directed against Menzies and Price's view as well. Menzies and Price respond to the circularity objection in the following way:

> The basic premise is that from an early age, we all have direct experience of acting as agents. That is, we have direct experience not merely of the Humean succession of events in the external world, but of a very special class of such successions: those in which the earlier event is an action of our own, performed in circumstances in which we both desire the later event, and believe that it is more probable given the act in question than it would be otherwise. To put it more simply, we all have direct personal experience of doing one thing and thence achieving another … It is this common and commonplace experience that licences what amounts to an ostensive definition of the notion of 'bringing about'. In other words, these cases provide direct non-linguistic acquaintance with the concept of bringing about an event; acquaintance which does not depend on prior acquisition of any causal notion. An agency theory thus escapes the threat of circularity. (1993: 194–195)

Unlike von Wright, Menzies and Price do not deny that agency is a causal phenomenon. What they deny is that acquiring the agency concept requires that one has already acquired the concept of causation. For Menzies and Price, even though agency itself is an essentially causal phenomenon, the *concept* of agency is one that can be understood and grasped independently of the *concept* of causation, and, because it can be independently understood, it can be used to analyse causation. As for whether Menzies and Price accept relationalism, it is not exactly clear. They describe the agency concept as 'a special class of successions' and as an action causing a result, which seems to suggest that they view agency in relational terms. However, ultimately I think it is unclear whether Menzies and Price's version of a manipulability account of causation accepts relationalism or not.

James Woodward (2003) argues that Menzies and Price's view is unacceptably anthropomorphic and subjectivist. Because Menzies and Price invoke a concept of agency that we grasp via direct experience of our own agency at work, their theory faces a difficult problem concerning causes that cannot be manipulated by human agents. To take an example from Menzies and Price (1993: 195), it seems to be true that movement of tectonic plates caused the 1989 San Francisco earthquake, but it is not true that movement of tectonic

plates was an event that could have been an effective means by which a human agent could have brought about the earthquake. Manipulating tectonic plates is just not within our power.

Woodward (2003), building on work by Judea Pearl (2000), offers his own manipulability theory of causation, which avoids this problem by using the concept of an intervention to analyse the causal relation, rather than manipulation by a human agent. Woodward contends that a variable *c* is causally related to a variable *e* if and only if intervention on *c* leaves the relationship between *c* and *e* invariant but changes the value of *e*. An intervention is any event that 'surgically' causes the value of *c* to change, that is, by blocking all causal influence over the value of *c* the usual causal antecedents of *c* have and without causally influencing the value of *e* except through *c*. An intervention is any event that has certain causal characteristics; an intervention need not involve human agency at all (although no doubt many interventions do involve human agency).

Woodward's theory is a kind of counterfactual theory of causation, since whether two variables are causally related to each other depends on how the relationship between those variables would change if certain interventions were made. However, there are key differences between Woodward and Lewis when it comes to the semantics of counterfactual conditionals. The most important difference is that in Lewis's account of how we should evaluate counterfactual conditionals in causal contexts it is never necessary to appeal to causal facts. By contrast, in Woodward's account of how we should evaluate counterfactual conditionals in causal contexts we are supposed to imagine that the antecedent of the counterfactual is made true by the occurrence of an intervention, which presupposes that certain causal facts obtain. To illustrate this point with an example, suppose event *c* caused e_1 and e_2, and e_1 and e_2 are not causally related to each other. Because counterfactual dependence is sufficient for causation, we would want the following counterfactual to come out false:

(a) If e_1 had not occurred, e_2 would not have occurred.

But, in a world where e_1 does not occur, we might suppose that this was because it was not caused by *c*, i.e. because *c* did not occur—but in that case e_2 would not have occurred either. This world—where e_1 does not occur because *c* does not occur—is therefore the wrong world to turn to when evaluating the truth of the counterfactual in a causal context. Lewis recommends that when we evaluate counterfactuals in a causal context we forbid 'backtracking'—i.e. we are forbidden from imagining that prior events and circumstances were also changed so as to cause the antecedent of our target counterfactual to be true. When we evaluate (a) we must imagine that a small miracle makes it the case that e_1 does not occur. So, the world we should use to evaluate the truth of (a) is a world where *c* still happens but then, miraculously, e_1 does not occur—in such a world e_2 would still occur (because *c* would still cause it), and therefore (a) comes out false.

Woodward achieves this same result using the notion of an intervention, rather than the notion of a 'small miracle'. For Woodward, when we evaluate (a) we are supposed to imagine that an intervention occurred to make it the case that e_1 did not occur—and such an intervention, by definition, leaves all causal relationships, except those which have e_1 as effect, unchanged. Evaluating the truth of (a) thus requires assuming certain other causal relations in the situation under discussion obtain. Even though Woodward's and Lewis's theories differ in this important way, it is not part of Woodward's theory that the truth of counterfactual conditionals depends on brute facts about powers, or a *sui generis* form of necessity. Thus, Woodward's theory is consistent with the view that counterfactual dependence can be understood without a primitive concept of power.

Does Woodward's theory embrace relationalism? The theory is intended to identify causal relationships between variables. On this theory, causation is something that exists between nodes in a network, and the concept of an intervention can tell us which relationships within this network are genuinely causal. On Woodward's theory, there is nothing extra in addition to the relationships between variables—such as the exercise of causal power or the bringing-about of events—which is essential to our understanding of causation. For this reason, I consider Woodward's theory a relationalist theory.

3.4 The relata of causation

Relationalism says that causation is always and everywhere a relation between distinct entities; however, it does not prescribe anything specific about what these entities must be. There is great disagreement on what the relata of causation are. There are many who hold that causation is a relation between *events* (Davidson 1967; Kim 1976; Lewis 1986). Some philosophers think that the relata of causation are *facts* (Bennett 1988; Mellor 1995). As mentioned above, Woodward (2003) holds that causation holds between *variables*. It has also been suggested that causation holds between *states of affairs* (Armstrong 1997), *conditions* (Mackie 1965) and *tropes* (Ehring 2011). I doubt this list is exhaustive. The situation is further complicated by the fact that there is very little agreement on the nature of entities like events, facts and states of affairs.

For example, among those who agree that causation is a relation between events, there is disagreement on what exactly events are. Davidson thinks that events are concrete particulars that can be redescribed and reidentified under different modes of presentation. This means that one and the same event can be referred to via different expressions, each of which identifies the event via a different intrinsic feature of it. For example, on Davidson's conception of events, Boudicca's death, Boudicca's suicide and Boudicca's poisoning are all one and the same event identified with different descriptions.

By contrast, Kim (1976) takes events to be 'exemplifications of properties at times'. Kim-events are located in space (they are where the objects exemplifying the properties are), and they are bound to a particular time (the times at which, or during which, the object exemplifies the properties) and they are contingent (they exist only if some object is a certain way). For Kim, the fact that his property exemplifications are bound to a particular time means that his events are particulars. However, Kim-events are also fact-like. Like facts, Kim-events indicate that an object is qualified. Also like facts, Kim-events have a propositional structure. The structure of Kim-events means that Kim-events are much more fine-grained than Davidsonian events. For example, Boudicca's exemplifying the property dying by suicide and Boudicca's exemplifying the property dying by poisoning would be distinct events as they involve distinct properties (dying by suicide and dying by poisoning).

Although relationalism is technically neutral with regard to what the relata of causation are, relational theories of causation that take causation to be a natural, extensional relation that holds between particulars (even fine-grained, fact-like particulars) are more supportive of other elements of the physicalist triad than theories of causation that allow causation to be an intentional relation. To see this, recall that the argumentative force of the causal argument is that, if mental and physical items are distinct, then they are in competition with each other for status as the cause of a physical effect. In order for there to be competition here, whether the mental item is cause of the physical effect cannot be something that depends on how the physical effect is described. The causal connection between the physical effect and its cause has to be a real relation. Furthermore, proponents of causal theories of intentional action state that their aim is to naturalise agency. As Bishop states, causal theories of action promise to 'make intelligible the possibility of agency within the natural order' (1989: 10). And Enç describes the causal theory of action as a 'treatment of action that confines itself just to events of the natural order of things, and to the causal relations among them' (2003: 3). Explaining agency in terms of events and causal relations could only be considered a project of naturalisation if causal relations are themselves natural relations that exist 'out there in the world'. As Giuseppina D'Oro explains:

> It is only if the term 'causation' is taken to be a category of revisionary metaphysics denoting a real relation, holding amongst events independently of how they are described, that the problem of causal rivalry between folk-psychological explanations of actions and naturalistic explanation of events can arise. The problem of explanatory exclusion simply does not arise within a descriptive conception of metaphysics precisely because, within such a conception of the role and character of philosophical analysis, causal relations are intentional relations that are not logically independent of the explanatory goals of a science. (2012: 219)

3.5 The importance of relationism

A central claim of this book is that the relational approach to causation is one of three mutually supporting views that form the physicalist triad. The relational approach to causation is, in some ways, the most fundamental of these three elements. The relational approach is appealing to both physicalists and those who endorse a causal theory of intentional action because of its associations with naturalism. The relational approach also lends support to both physicalism and causal theories of intentional action. If one adopts a relational approach to causation, then it seems inevitable that mental causation will be understood in relational terms, i.e. presented as a cause–effect relation between mental and physical entities. If all causation everywhere is the same, the only thing that can discriminate between different categories of causation is the nature of the relata involved. Furthermore, if one adopts a relational approach to causation, intentional action must be distinguished in terms of its aetiology. Alternatives to causal theories of intentional action, which purport to understand intentional action in terms of irreducible agent causation, are uncongenial to the relational approach to causation.

As I have already mentioned, I am not the first to suggest that there are intellectual connections between physicalism, philosophy of action and philosophy of causation (see for example Hornsby 2015; Lowe 2008). However, as Jennifer Hornsby (2015) notes, these connections have been underexplored. Some writers in philosophy of mind have suggested that the best way to respond to the causal argument for physicalism is to challenge the assumptions about causation implicit in the argument. For example, List and Menzies (2009) argue that construing causation as 'difference-making' allows one to argue that higher-level mental properties are not causally excluded by the physical properties that realise them. However, what these writers suggest is a fairly modest rethinking of the assumptions about causation at work in the causal argument, and the metaphysics of mind they eventually endorse is usually a kind of non-reductive physicalism. A number of writers in philosophy of action who are dissatisfied with causal theories of intentional action have suggested that Aristotelian views about causation are needed to properly understand agency. However, as we shall see in Chapter 5, although these neo-Aristotelian views of agency posit the existence of a special kind of causation (agent causation or substance causation), they do not explicitly challenge the idea that causation is always, everywhere a relation.

In the last three chapters I have tried to make salient the mutually supporting relationships between physicalism, causal theories of intentional action and the relational approach to causation. My next task is to explain why the best strategy for resisting the conclusion of the causal argument for physicalism is to use lessons from philosophy of action to challenge the relational understanding of mental causation.

References

Anscombe, G E M 1971 *Causality and determination: An inaugural lecture*. Cambridge: Cambridge University Press.

Armstrong, D M 1997 *A world of states of affairs*. Cambridge: Cambridge University Press.

Baumgartner, M 2008 Regularity theories reassessed. *Philosophia*, 36(3): 327–354. DOI: https://doi.org/10.1007/s11406-007-9114-4

Beebee, H 2006 Does anything hold the universe together? *Synthese*, 149(3): 509–533. DOI: https://doi.org/10.1007/s11229-005-0576-2

Beebee, H 2007 The two definitions and the doctrine of necessity. *Proceedings of the Aristotelian Society*, 107(1pt3): 413–431. DOI: https://doi.org/10.1111/j.1467-9264.2007.00231.x

Bennett, J 1988 *Events and their names*. Indianapolis, IN: Hackett.

Bishop, J 1989 *Natural agency: An essay on the causal theory of action*. Cambridge: Cambridge University Press.

D'Oro, G 2012 Reasons and causes: The philosophical battle and the metaphilosophical war. *Australasian Journal of Philosophy*, 90(2): 207–221. DOI: https://doi.org/10.1080/00048402.2011.583930

Davidson, D 1967 Causal relations. *Journal of Philosophy*, 64(21): 691–703. Reprinted in Davidson 2001 pp. 149–162.

Davidson, D 2001 *Essays on actions and events*. 2nd ed. Oxford: Clarendon Press.

Ehring, D 2011 *Tropes: Properties, objects, and mental causation*. Oxford: Oxford University Press.

Enç, B 2003 *How we act: Causes, reasons, and intentions*. New York: Oxford University Press.

Gallow, J 2022 The metaphysics of causation. The Stanford Encyclopedia of Philosophy, 14 April 2022. Available at https://plato.stanford.edu/archives/sum2022/entries/causation-metaphysics [Last accessed 2 September 2023].

Ganeri, J, Noordhof, P and Ramachandran, M 1996 Counterfactuals and preemptive causation. *Analysis*, 56(4): 219–225. DOI: https://doi.org/10.1093/analys/56.4.219

Hocutt, M 1974 Aristotle's four becauses. *Philosophy*, 49(190): 385–399. DOI: https://doi.org/10.1017/s0031819100063324

Hornsby, J 2015 Causality and 'the mental'. *HUMANA.MENTE Journal of Philosophical Studies*, 8(29): 125–140.

Hume, D 1964 *A treatise of human nature. Volume 1*. London: J M Dent/E P Dutten.

Hume, D 1975 *Enquiries concerning human understanding and concerning the principles of morals*. 3rd ed. Oxford: Clarendon Press.

Kim, J 1976 Events as property exemplifications. In: Brand, M and Walton, D *Action theory*. Dordrecht: D. Reidel. pp. 310–326.

Lewis, D K 1973a Causation. *The Journal of Philosophy*, 70(17): 556–567. DOI: https://doi.org/10.2307/2025310
Lewis, D K 1973b *Counterfactuals*. Oxford: Basil Blackwell.
Lewis, D K 1986 *Philosophical papers*. New York: Oxford University Press.
Lewis, D K 1994 Humean supervenience debugged. *Mind*, 103(412): 473–490. DOI: https://doi.org/10.1093/mind/103.412.473
Lewis, D K 2000 Causation as influence. *Journal of Philosophy*, 97(4): 182–197.
List, C and Menzies, P 2009 Nonreductive physicalism and the limits of the exclusion principle. *Journal of Philosophy*, 106(9): 475–502. DOI: https://doi.org/10.5840/jphil2009106936
Lowe, E J 2008 *Personal agency: The metaphysics of mind and action*. New York: Oxford University Press.
Mackie, J L 1965 Causes and conditions. *American Philosophical Quarterly*, 2(4): 245–264.
Mackie, J L 1974 *The cement of the universe*. Oxford: Clarendon Press.
McDermott, M 2002 Causation: Influence versus sufficiency. *The Journal of Philosophy*, 99(2): 84–101. DOI: https://doi.org/10.5840/JPHIL200299219
Mellor, D H 1995 *The facts of causation*. London: Routledge.
Menzies, P and Price, H 1993 Causation as a secondary quality. *British Journal for the Philosophy of Science*, 44(2): 187–203. DOI: https://doi.org/10.1093/bjps/44.2.187
Mill, J S 1843 *A system of logic, ratiocinative and inductive being a connected view of the principles of evidence and the methods of scientific investigation*. London: Longmans, Green, Reader, and Dyer.
Millican, P 2007 Against the new Hume. In: Read, R and Richman, K *The new Hume debate: Revised edition*. London: Routledge. pp. 211–252.
Ott, W 2009 *Causation and laws of nature in early modern philosophy*. Oxford: Oxford University Press.
Pearl, J 2000 *Causality: Models, reasoning and inference*. Cambridge: Cambridge University Press.
Psillos, S 2002 *Causation and explanation*. London: Routledge.
Sartorio, C 2005 Causes as difference-makers. *Philosophical Studies*, 123(1/2): 71–96. DOI: https://doi.org/10.1007/s11098-004-5217-y
Schaffer, J 2007 Causation and laws of nature: Reductionism. In: Sider, T, Hawthorn, J and Zimmerman, D W *Contemporary debates in metaphysics*. Malden: Blackwell, pp. 82–107.
Schaffer, J 2016 The metaphysics of causation. The Stanford Encyclopedia of Philosophy, 5 July 2016. Available at https://plato.stanford.edu/archives/spr2022/entries/causation-metaphysics [Last accessed 2 September 2023].
Steward, H 2012 *A metaphysics for freedom*. Oxford: Oxford University Press.
Strawson, G 1987 Realism and causation. *The Philosophical Quarterly*, 37(148): 253–277. DOI: https://doi.org/10.2307/2220397
Strawson, G 1989 *The secret connexion: Causation, realism, and David Hume*. Oxford: Oxford University Press.

Tooley, M 1990a Causation: Reductionism versus realism. *Philosophy and Phenomenological Research*, 50: 215–236. DOI: https://doi.org/10.2307/2108040

Tooley, M 1990b The nature of causation: A singularist account. *Canadian Journal of Philosophy*, 20(1): 271–322. DOI: https://doi.org/10.1080/00455091.1990.10717229

von Wright, G H 1971 *Explanation and understanding*. Ithaca, NY: Cornell University Press.

von Wright, G H 1974 *Causality and determinism*. New York: Columbia University Press.

Wilson, J 2010 What is Hume's dictum, and why believe it? *Philosophy and Phenomenological Research*, 80(3): 595–637. DOI: https://doi.org/10.1111/j.1933-1592.2010.00342.x

Woodward, J 2003 *Making things happen: A theory of causal explanation*. New York: Oxford University Press.