

UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

Rethinking Agent Causation: Essays on Freedom, Responsibility, and Action

Permalink

<https://escholarship.org/uc/item/9wd13529>

Author

Martinez, Joseph

Publication Date

2024

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA SAN DIEGO

**Rethinking Agent Causation:
Essays on Freedom, Responsibility, and Action**

A dissertation submitted in partial satisfaction of the
requirements for the degree

Doctor of Philosophy

in

Philosophy

by

Joseph Martinez

Committee in charge:

Professor Dana Nelkin, Co-Chair
Professor Manuel Vargas, Co-Chair
Professor Eddy Chen
Professor Gail Heyman
Professor Daniel Speak
Professor Eric Watkins

2024

Copyright

Joseph Martinez, 2024

All rights reserved.

The dissertation of Joseph Martinez is approved, and it is accepted in quality and form for publication on microfilm and electronically.

University of California San Diego

2024

Table of Contents

Dissertation Approval Page.....	iii
Table of Contents.....	iv
List of Figures.....	vi
Acknowledgements.....	vii
Vita.....	ix
Abstract of the Dissertation.....	x
Chapter 1: Introduction.....	1
1.1 Overview	1
1.2 Free Will and Moral Responsibility.....	2
1.3 Compatibilism and Libertarianism	3
1.4 Events and Agents.....	3
1.5 The State of Play.....	5
Chapter 2: Causal Pluralism: Agent Causation Without the Panicky Metaphysics.....	6
2.1 Introduction	6
2.2 The Dichotomy.....	7
2.2.1 Event-Causalism.....	8
2.2.2 Agent-Causalism	10
2.3 A False Dichotomy—And A Causal Pluralist Proposal.....	13
2.4 A Challenge For Agent-Causalist.....	17
2.5 Events Reconsidered.....	24
2.5.1 States Of Affairs	25
2.5.2 Spatiotemporal Regions.....	26
2.5.3 <i>Properties</i> Of Spatiotemporal Regions.....	27
2.6 Concluding Remarks	30
Chapter 3: Agents and Laws: Reconsidering Pereboom’s Wild Coincidence Objection.....	35
3.1 Introduction	35
3.2 Agent Causation and Indeterminism.....	37
3.3 The Objection	38
3.4 The Objection Reconsidered	45
3.4.1 The Second Horn Reconsidered	46
3.4.2 The Laws Reconsidered	47
3.4.3 Horizontal Reductionism	50

3.4.4	Vertical Reductionism	53
3.5	Three Replies.....	56
3.5.1	Baker’s Reply	56
3.5.2	Taggart’s Reply	59
3.5.3	Runyan’s Reply	60
3.6	Concluding Remarks	63
Chapter 4: Agent Causation and Motivating Reasons		65
4.1	Introduction	65
4.2	Two Theories of Action	66
4.2.1	Event-Causalism.....	69
4.2.2	Action and Free Will.....	70
4.3	Causalist Accounts of Motivating Reasons.....	71
4.3.1	Motivating Reasons as Co-Causes.....	73
4.3.2	Motivating Reasons as Prior Causes	76
4.3.3	Propensity-Based Account.....	80
4.4	Teleological Realism.....	84
4.5	A Causal Powers Account.....	88
4.6	Concluding Remarks	92
Chapter 5: Agent Causation, Responsibility, and Reasons-Responsiveness.....		94
5.1	Introduction	94
5.2	The Gap	95
5.3	Filling the Gap: Reasons-Responsiveness.....	98
5.4	Problems With Integration.....	104
5.4.1	Libertarian Commitments.....	104
5.4.2	Indeterminacy Itself.....	111
5.5	Probabilities to the Rescue?.....	115
5.6	Taking Stock.....	117
5.7	The Scope of the Problem.....	118
5.8	Where Do We Go From Here?.....	120
Chapter 6: Conclusion.....		126
6.1	Remaining Questions	126
6.2	Final Remarks	129
References.....		130

List of Figures

Figure 1.1: Free Will Diagram.....	4
Figure 3.1: Competing Explanations.....	44
Figure 4.1: Agent-Causal Theory of Action.....	67
Figure 4.2: Event-Causal Theory of Action.....	69
Figure 4.3: Motivating Reasons as Co-Causes.....	73
Figure 4.4: Motivating Reasons as Prior Causes.....	76
Figure 4.5: Modified Agent-Causal View.....	78

Acknowledgements

I would like to give special thanks to Dana Nelkin and Manuel Vargas for their extensive feedback and discussions on all parts of this dissertation. Their conversations greatly improved the ideas developed throughout, and their suggestions saved my project more times than I can count. I truly admire their example of philosophical excellence, and I hope to be able to approximate it one day.

I would also like to acknowledge and thank Eric Watkins, Daniel Speak, and Eddy Chen, for their feedback and discussions on various chapter drafts, all of which I greatly benefitted from, and which helped to push the dissertation forward in important ways. Thank you also to Gail Heyman for her thoughtful comments and questions about the dissertation project.

I am also grateful to Taylor Madigan, Mikayla Kelley, Ayoob Shahmoradi, Zack Brantz, Haigen Messerian, and Ryan Ravanpak for their comments and discussions on various chapter drafts. Taylor and Mikayla helped to greatly advance my understanding of the philosophy action, and Ayoob, Zack, Haigen, and Ryan all helped to hone various ideas developed throughout the dissertation.

I must also thank Nancy Cartwright for her comments on chapter 3 and extensive conversations about the dissertation project. Her work on the philosophy of science, and laws of nature in particular, helped me to see the empirical plausibility of agent causation, which in turn gave me the initial confidence to write a dissertation on the subject matter. And thank you to Christopher Shields for his helpful discussion on chapter 4, which helped me to improve my understanding of various claims I make there.

My deepest gratitude to Grace for her patience, love, and support throughout my time here at UCSD, without which I could not have made it through graduate school, as well as for her comments on a number of different chapter drafts. And, finally, thank you to Emmelia and Isidora, whose love provided me with the motivation to bring this project to completion.

Chapter 2, in part, has been accepted for publication (as “Causal Pluralism: Agent Causation Without the Panicky Metaphysics”) in *Synthese*, 2023, Martinez, Joseph. The dissertation author was the sole author of this paper.

Chapter 3, in part, is currently being prepared for submission for publication of the material (as “Agents, Laws, and Events: Reconsidering Pereboom’s Wild Coincidence Objection,” by Martinez, Joseph). The dissertation author was sole author of this paper.

Chapter 4, in part, is currently under commission for publication (as “Agent Causation and Motivating Reasons”) and may appear in *Philosophical Compass*, Martinez, Joseph. The dissertation author was the sole author of this paper.

Vita

- 2016 B.A. in Philosophy, University of California, Berkeley
- 2019 M.A. in Philosophy, University of California San Diego
- 2024 Ph.D. in Philosophy, University of California San Diego

Abstract of the Dissertation

Rethinking Agent Causation: Essays on Freedom, Responsibility, and Action

by

Joseph Martinez

Doctor of Philosophy in Philosophy

University of California San Diego, 2024

Professor Dana Nelkin, Co-Chair
Professor Manuel Vargas, Co-Chair

Taking things at face value, we are free and responsible agents. That is to say, sometimes we act in ways that merit praise and blame. I believe that the theory that best makes sense of this fact—supposing that it is one—is the agent-causal theory of free will. At the heart of the agent-causal theory is the idea that agents are directly causally involved in the production of their free actions; we cannot

wholly account for the agent's causal role in free will by referring to, as has been typically thought, things going on inside the agent's head. For much of the last century, most writers have neglected to take the theory seriously, often outright dismissing it. I argue that this neglect has not been justified. There is more to be said for the view than critics have previously realized. Furthermore, probing these criticisms and honing the agent-causal theory helps to shed light on a number of important issues for the free will dialectic.

The aim of this project is two-fold: first, it aims to address substantive objections traditionally leveraged against agent causation, and, second, it aims to begin to develop the agent-causal view in ways that make the view more plausible. Chapter 1 starts by introducing the background and framework that will be in place for the rest of the discussion. Chapter 2 addresses a long-standing concern about agent causation, which says that the commitment to substance causation is too ontologically demanding. I argue that agent causation does not necessarily require substance causation and is in fact compatible with certain event-causal frameworks. Chapter 3 addresses an empirical objection which says that agent causation conflicts with what we know about the laws of physics. I argue that the objection overstates the empirical evidence against agent causation. Chapters 4 and 5 turn to the positive project of elucidating under-discussed aspects of agent causation. Chapter 4 examines how best to account for acting on a reason within an agent-causalist framework. It is argued that the most plausible account involves taking the relation of acting on a reason as a primitive. And Chapter 5 then considers what a reasons-responsive account of agent causation might look like. It is argued that a straightforward integration between agent causation and the standard reasons-responsive account is implausible. Nonetheless, the failure of integration reveals a general problem for reasons-responsive accounts, suggesting that the most plausible account of reasons-responsiveness will be non-modal.

Chapter 1

Introduction

§1.1 Overview

I take it as a starting presumption that we are free and responsible agents. By this, I mean that we sometimes perform actions that merit praise and blame. It is my contention that the best overall theory that makes sense of the fact that we are free—supposing that it is a fact—is the agent-causal libertarian theory of free will. While I will not argue for the view directly here, I will do so indirectly. Agent causation's reputation has seen improvements in recent years. Still, many approach it with a great deal of skepticism. The view is often perceived of as, at best, relying on a framework that is empirically implausible and, at worst, resorting to an "obscure and panicky metaphysics" (Strawson 1962, 211). I aim to show that this skepticism is not warranted because these criticisms are not as well-founded as authors have made them out to be.

Another important aspect in motivating a view is developing it in such a way that it can do important explanatory work. This is one reason that I believe compatibilist views (especially reasons-responsive theories) have enjoyed such widespread support. Agent-causalists, it seems to me, have fared relatively poorly on this score. They have focused so much on developing a metaphysically coherent view that they have neglected to attend to issues that are central to our lives, especially matters related to our moral responsibility practices. With this in mind, I also aim to begin to develop two important and under-discussed aspects of agent causation: what it means to act on a reason and how to accommodate reasons-responsive elements into the view.

The motivation for attending to these two particular issues is two-fold. First, questions of moral responsibility, especially questions about when agents are excused or justified in what they do, are tightly connected to our ability to register and respond to reasons. Second, because the notion of reasons-responsiveness is parasitic on the notion of acting on a reason, it is important that we try to get

clear on the latter before trying to address the former. Addressing these issues will, I hope, help lay the groundwork for a more appealing picture of agent causation, one that is better able to address important matters, such as our inter-personal and social practices.

With that introduction out of the way, the rest of chapter will focus on introducing some core background assumptions that will be at play throughout the rest of this dissertation, as well as setting up the state of play in the free will literature vis-à-vis agent causation.

§1.2 Free Will and Moral Responsibility

I start with a few comments about what I take free will to be and the underlying reasoning for it. To start, here is the definition of free will that I will be working with:

Free Will: The control over one's actions that would render one morally responsible for those actions.

To say that some agent is morally responsible is just to say they are deserving of praise or blame—that praise or blame would be a fitting response to what they did. All things being equal, the agent would deserve praise for freely performing a morally good action and would deserve blame for performing a morally bad action. I leave the conceptions of praise and blame intentionally capacious. Praise and blame may encompass anything from inward judgements of approval and condemnation to external responses of state-sponsored awards or punishment.

As a simple illustration, consider an agent, Danny. On this definition, Danny would have acted freely in committing charity fraud just in case blaming Danny would be an appropriate response to his action. That is, it would at least be a fitting response to, at the very least, criticize Danny's action—or perhaps even to punish him with a fine or incarceration, depending on one's view about punishment.

Furthermore, an agent is morally responsible in virtue having a special kind of control over their actions. We can see this most clearly by considering cases where an agent is no longer considered

morally responsible. For example, imagine that we alter the above situation so that Danny performed charity fraud only because he was under the hypnotic influence of a third party. Or imagine that Danny spills a boiling cup of coffee onto someone's lap, but only because his arm had a muscle spasm. Our assessment of Danny is that he is not morally responsible. And what seems to explain why Danny is not morally responsible is that he lacks the relevant kind of control over his actions.

§1.3 Compatibilism and Libertarianism

There are two important divides in the literature that will be important to our discussion. First, one of the longest standing divides in the free will literature concerns the divide between **compatibilists** and **libertarians**. The divide revolves around whether free will is compatible with the truth of **determinism**. I'll have more to say about determinism in Chapter 4, but the basic idea for now is this: the world is deterministic if everything that occurs at a particular time had to occur given what came before it. For example, if Danny had to raise arm given the state his brain was in just a moment prior, then Danny's raising his arm was determined to occur by that brain state.

Compatibilism, then, is the view that we could still be free even if our actions were determined. Our having the control that would render us morally responsible is consistent with the fact that everything that we do is determined. Incompatibilism, in contrast, is the view that we would *not* be free if our actions are determined. Determinism would undermine the control needed for moral responsibility.

Incompatibilism can be further divided into two camps: free will skepticism and libertarianism. Free will skeptics are incompatibilists who believe that free will is incompatible with determinism, but that we are not free. And libertarians are incompatibilists who believe that we are free.

§1.4 Events and Agents

A second prominent divide in the literature concerns the divide between event-causalists and agent-causalists. More will be said about this in the following chapters (Chapters 2 and 4). For now, the divide roughly concern's the causal role that the agent plays in free will. Event-causalists maintain that, when an agent performs a free action, the agent's causal role is exhausted by the causal role of her *mental events*. Mental events, broadly construed, pick out events things like a person's feeling pain, believing something, desiring something, and intending to do something. In the context of free will, the relevant mental events are typically taken to be reasons, where a reasons mental event consists of a person's believing and desiring something. For example, if Danny desires to wake up and believes that drinking an espresso will wake him up, then Danny has a reason to drink espresso. Thus, event-causalists maintain that, when an agent performs a free action, it is the agent's reasons that cause the free action.

In contrast, agent-causalists maintain that actions are free only if the agent plays a direct causal role in the production of free action. Rather than saying that the agent's reasons cause a free action—as do event-causalists—the agent-causalist maintains that the *agent* is what causes her free actions. The agent's causal role is direct insofar there is no intermediate event that stands in for the agent.

With these distinctions, we can carve out five primary views about free will:

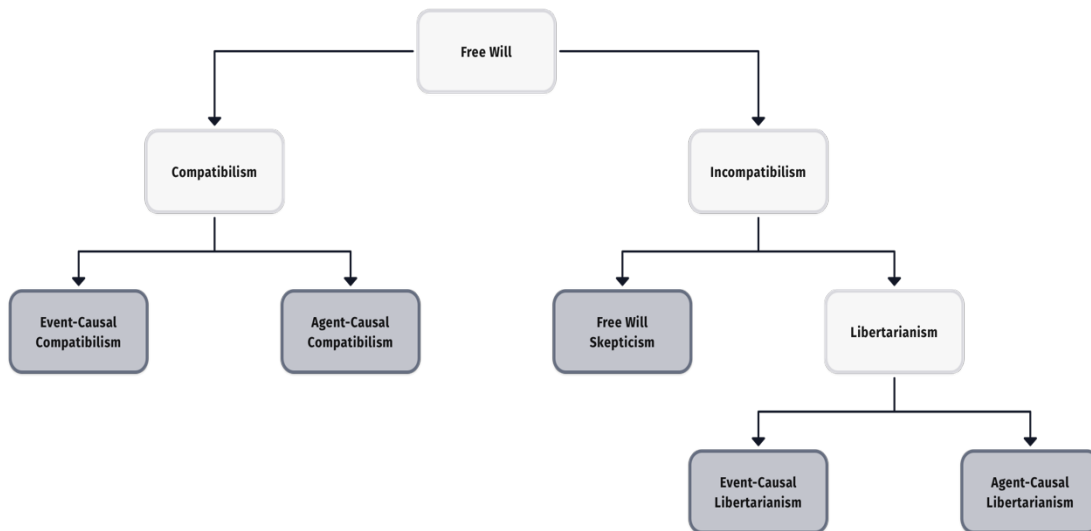


Figure 1.1: Free Will Diagram

§1.5 The State of Play

As things currently stand, agent-causal libertarianism (henceforth just agent-causalism) enjoys a fairly narrow degree of support. The primary reason for this, as stated at the outset, is that the notion of agent causation is looked at with a great deal of skepticism. The concerns can be divided into two kinds: metaphysical and empirical.

As per the metaphysical concern, many find the idea of agent to be metaphysically extravagant, if not mysterious. At the beginning of Roderick Chisholm's seminal (1964) paper where he reintroduced agent causation back into the free will literature, he makes the following remarks:

To solve the problem [of free will], I believe we must make somewhat far-reaching assumptions about the self or the agent—about the man who performs the act. (1964, 24)

This sentiment has continued. Most authors continue to believe that positing direct causation by the agent requires us to make too much of a "far-reaching assumption" in order to account for free will. What agent causation is and how it works are looked at with skepticism.

As per the empirical concern, a common sentiment is that agent causation is incongruent with our empirical picture of the world. This concern is perhaps best encapsulated in Derk Pereboom's Wild Coincidence Objection which he has developed over the recent years. As the objection goes, our best picture of physics leaves no room for anything like agent causation.

With that setup out of the way, I turn to addressing the first of our criticisms.

Chapter 2

Causal Pluralism:

Agent Causation Without the Panicky Metaphysics

§2.1 Introduction

As outlined in the previous chapter, there exists a deep divide between event- and agent-causalists. Simplifying things somewhat, event-causalists maintain that a free action is caused by appropriate (mental) events. In contrast, agent-causalists maintain that a free action is caused, not by any events, but by a substance—namely, the agent. So construed, agent-causalists maintain that substance causation is necessary for free will, while event-causalists maintain that event causation is necessary for free will. A central point of contention within the free will dialectic, then, surrounds whether substance or event causation is necessary to secure free will.

It is often assumed that the choice to include agents as fundamental causal relata marks a radical departure from the standard event-causal framework—so much so that, for the last few decades, the postulation of fundamental substance causes has often been taken as sufficient grounds to dismiss agent-causalism as a credible theory of free will. Because of this, event-causalist views have solidified themselves as the orthodox position within the free will literature.

In this chapter, I introduce and motivate a new view on free will, which I call *Causal Pluralism*. To foreshadow, Causal Pluralism is the view that free will is compatible with both (fundamental) event and substance causation. Indeed, as I develop the view here, free will is compatible with both kinds of causal relata *because* event causation can accommodate much of the metaphysics of substance causation. I show this by considering the nature of events more closely. In particular, I argue that, if the standard agent-causalist wants to reject the Causal Pluralist view I develop here, this requires

showing that there is a freedom-relevant difference between event causation (properly construed) and substance causation. In addition to this, I explain how Causal Pluralism has the advantage of capturing most of what extant agent-causalists have wanted to say about free will, all while making fewer ontological demands. By doing all of this, I hope to assuage the concern that agent causation resorts to an obscure and panicky metaphysics.

Still, even if one is not ultimately persuaded by the view, considering the Causal Pluralist proposal helps to draw out at least two important points. First, there is a lack of clarity regarding how the two types of causal relata—i.e., events and substances—relate. Second, getting clearer about how the two causal relata relate suggests that, in some important respects, the divide between event- and agent-causalists may be more superficial than it initially seems.

Here is the plan for the rest of the paper. §2 starts by providing a bit more detail regarding the event- and agent-causalist views, with a focus on how they differ. Following that, §3 explains why there is a false dichotomy when choosing between event and substance causation in a theory of free will. The Causal Pluralist view is introduced as a way of drawing out the false dichotomy. §4 turns, then, to providing further motivation for why extant agent-causalists should take Causal Pluralism seriously. Because the initial motivation for Causal Pluralism in §§3-4 relies heavily on a property exemplification model of events, §5 returns to examining the nature of events. More specifically, it is argued that Causal Pluralism does not stand or fall with a property exemplification model of events. §6 finishes with some reflection on Causal Pluralism's relation to agency reductionism and event-causalism. An important upshot of these remarks is that event causation and agency reductionism can be prized apart.

§2.2 The Dichotomy

Understanding the dispute between event- and agent-causalists will require a brief excursion through both campgrounds. While there is much to be said for and against each of these positions, my

aim in this section is not take to a stand on which position should be endorsed. Rather, it is merely to lay out the positions in order to draw out the event-substance divide. In the next section, I explain why the event-substance divide may be less important than initially thought.

§2.2.1 Event-Causalism

Before beginning, in order to preempt any confusion, it is important to note up front that the accounts which will be the focus of our discussion will be those that accept the control condition for free will, especially in its causal form. So, on this assumption, if an agent's action is to be free, the agent must stand in some suitable causal relation to that action. With that said, a view is a species of event-causalism when it seeks to explicate the control condition on free will wholly in terms of events causing one another. This would include both compatibilist and incompatibilist theories of free will.¹ While a number of different accounts fall under this heading (e.g., Franklin 2018; Fischer and Ravizza 1999), the various accounts appear to build upon a common core:

Event-Causalist Condition: An agent freely performed some action ϕ only if ϕ was caused by mental events involving the agent (e.g., the agent's reasons).

Though not essential to these views, it is commonly assumed that directly free actions are *mental actions*, such as the formation of an intention or the making of a decision. On this position, bodily movements—such as raising an arm or pulling a trigger—are derivatively free; that is, they are free in virtue of being caused by (directly) free mental actions. Though not much should hinge on this, for

¹ A closely related position is non-causalism. Non-causalists typically maintain that all causation consists in events causing one another, but also that free actions are uncaused—either by a (mental) event or an agent. Non-causalists, then, deny that the control required for free will is causal in nature. I take control to be a causal notion, so I put non-causalism to the side in the remainder of this discussion. See Widerker (2018) for a recent defense of the view.

ease of exposition, I will presume for simplicity that bodily movements (like the mental actions that cause them) are directly free.

With that said, consider a simple example. Suppose that Angelica makes an utterance. According to the event-causalist condition, Angelica's utterance was freely performed only if it was caused by mental events involving Angelica. As indicated in the description of the event-causalist condition, it is typically assumed that the appropriate mental events are the agent's reasons.

There are at least two ways to understand what the reasons mental state consists in. On the standard story, following Davidson, reasons are reduced to two more basic mental events: a belief-desire pair. So, for instance, Angelica's having a reason to make the utterance consists in her having both a desire to make the utterance and some kind of belief that her utterance will accomplish a particular goal. This, we might note, invokes an internalist account of reasons insofar as it posits that the appropriate reasons are *internal* to the agent.

An alternative account—one that is more sympathetic to externalist accounts—grants that reasons are, perhaps, something like considerations that count in favor of a particular action, but goes on to add that the appropriate mental event for the event-causalist condition is the *agent's taking something to be a reason*. If some reason *R* is a consideration in favor of making an utterance, then *Angelica's taking R to be a reason* can play the requisite role for the event-causalist (Franklin 2018).

According to event-causalism, (the control condition on) free will consists in events causing one another. On the one hand, we have mental events playing the role of causes, and, on the other hand, we have actions (whether mental or bodily [i.e., physical]) playing the role of effects. Whatever role the agent plays in the performance of free action, it will be mediated by their relation to these events. Furthermore, note that, on event-causalism, it is a *necessary* condition on free will that a free action be caused by a (mental) event.

The nature of events is generally left unanalyzed, even for event-causalists; and we will do so as well for the time being. Of course, later on, we will have more to say about the nature of events. It

should be sufficient for our purposes here to describe events at a high level of generality. Events, we may say, then, are simply “happenings” or “things that occur.” They may also be further picked out with common examples such as *the tree’s falling*, *Old Yeller running a mile*, and *Juan’s wanting a drink*.

Applying what we’ve said to a case of free action, the relevant events and their causal relation would look something like this: when Angelica has freely made an utterance, then *Angelica’s having a reason to make the utterance* has caused *Angelica’s making the utterance*.

§2.2.2 Agent-Causalism

Agent-causalism denies that the control condition on free will is explicable wholly in terms of events causing one another. Instead, they posit an irreducible agent as a cause. Just as with event-causalism, I am understanding agent-causalism to include both compatibilist and incompatibilist theories (e.g., Nelkin 2011; Markosian 1999; Jacobs and O’Connor 2013). Here is the core of such accounts:

Agent-Causalist Condition: An agent freely performed some action ϕ only if ϕ was directly caused by the agent (qua substance).

Agent-causalism requires that a free action not be fully causally explained by any event(s), but rather the explanation must make some essential appeal to the agent’s direct causal involvement. What it means for the agent to *directly* cause an action is for the agent to (literally) be the causal relatum that stands in causal relation to the action. So, instead of saying that Angelica’s freely making an utterance is explained in terms of an event causing *Angelica’s making an utterance*, the agent-causalist contends that Angelica herself caused the event of *Angelica’s making an utterance*. What we have here is a case of a substance standing in causal relation to an event.

Substances, for the agent-causalists, are said to correspond roughly to our conception of ordinary objects, and further described as “concrete particulars.” So, for example, cups, chairs, trees, tigers, and, of course, people, would all count as substances in this picture. Most proponents of agent-causalism do not provide a deeper analysis of what makes some complex entity a substance, as opposed to a mere aggregate, though a number of proposals are available (see, for example, van Inwagen 1990; Markosian 1998; Merricks 2001). Whatever criteria agent-causalists may come down on, they are at least committed to the position that agents are substances.

It is important to note that the agent-causalist need not posit any new kind of causation. Formally speaking, the causal relation that an agent stands in to an event is no different than the causal relation one event stands in to another event; it is the very same kind of causal relation that obtains when one event causes another event. The only difference is that the *causal relata* differ in each case (Clarke 2003: 186).

When the agent directly causes an action, this is said to occur at least partly in virtue of the agent’s possessing “a *causal power*, fundamentally as a substance, to cause a decision” (Pereboom 2014: 51, emphasis added). This causal power is typically referred to as an *agent-causal power*. Early agent-causalists seemingly thought that no deeper analysis of agent causation was available, remarking that agent causation was “primitive” (Bishop 1983: 74), “undefined” (Chisholm 1979: 70), and even “mysterious” (Taylor 1991: 53). More recent defenders of agent causation, however, have been more optimistic about giving such an account.

This optimism seems to have arisen in light of the resurgence of so-called neo-Aristotelian metaphysics, which often provide special priority to the notions of substance and causal power. Within neo-Aristotelian metaphysics, causal powers are taken to be a class of properties whose essences are linked to particular ends (e.g., Mumford 1998; Mumford and Anjum 2011; Bird 2007; Lowe 2003). That is, it is part of the nature of causal powers that they help to “bring about, or probabilify, certain effects” (Armstrong 2005: 312). An example of a causal power might be “charge.”

It is part of the nature of charge that particles that possess it have the capacity to repel other particles (cf. Kuykendall 2019: 340). It is typical, then, for powers theorists to explain all causation in terms of the exercise of causal powers like this.

While powers theorists are divided about what the fundamental causal relata are, most are quite friendly to the idea of substance causation, since substances are taken to be the bearers of causal powers. Given all of this, it is no surprise that agent-causalists have looked to neo-Aristotelian metaphysics for support. Jacobs and O'Connor (2013) and (O'Connor 2021) are perhaps the clearest cases where this commitment is explicitly made. On this kind of picture, the agent-causal power will only differ from other causal powers in terms of the kind of effect that it is directed at. Whereas, say, charge is directed at repelling other particles, the agent-causal power is directed at causing actions.²

It is sometimes thought that talk of agent (or substance) causation is merely elliptical for talk of a certain kind of event causation. For instance, Judith Thomson (2003) distinguishes between a non-fundamental “cause” and a fundamental “CAUSE.” She then maintains that only events can be CAUSES, but allows that non-events, such as agents, can be causes. She writes that, when x causes y, “[i]f neither x nor y is an event, then for x to cause y is for some event *appropriately related* to x to CAUSE some event *appropriately related* to y” (emphasis added). On such a construal, agents can play a kind of causal role, though this causal role is ultimately derivative of the more fundamental causal role of some events.³

Clarke (2017) similarly suggests that “[c]ausation by substances is non-fundamental if it can be reductively analyzed in terms of causation by things of one or more other categories.” More specifically, he suggests:

² A related issue is what role reasons play in the production of action. O'Connor's (2000) proposal is that reasons structure the agent's probabilities to act in different ways. Another proposal from Clarke (2003) is that reasons enter in as co-causes of actions. For the purposes of this chapter, I put this issue to the side and focus on the agent's causal relation to free action.

³ See also Kane (1998) for a helpful discussion on the matter.

Reductive Analysis: Substance S caused event e_2 just in case there was some event, e_1 , such that e_1 involved S and e_1 caused e_2 .

The relation of involvement can be interpreted in various ways, such as the “event’s being a change undergone by that substance,” or perhaps even the substance’s simply being the subject of the event.

Furthermore, the “just in case” idea expressed in Reductive Analysis is that the events standing in causal relation are both conceptually and ontologically prior to the substance causing an event, such that it is the event’s standing in causal relation that explains why the substance stands in a causal relation to the event (Clarke 2017). Put another way, the idea is that substance causation is non-fundamental if substance causation is grounded in—i.e., made true by—event causation. We might also add, then, that event causation is non-fundamental if event causation is grounded in substance causation.

The key takeaway of this discussion is that a pertinent distinction between event- and agent-causalism is the ontological and conceptual priority given to each causal relata (at least in the context of free actions). The event-causalist maintains that events have ontological and explanatory priority with respect to any substance causation—i.e., that events are the fundamental causes of free actions—while the agent-causalist maintains that substance causation has ontological and conceptual priority with respect to any event causation—i.e., that substances are the fundamental causes of free actions.⁴ (In what follows, all talk of causation refers to fundamental causation, unless stated otherwise.)

§2.3 A False Dichotomy—And A Causal Pluralist Proposal

⁴ The above discussion raises interesting questions about how best to understand the status of non-fundamental causes, as well as what their relation is to fundamental causes. Here, I leave it open how best to answer these questions. However, see fn. 6 for a sketch of some possible positions.

The boiler-plate dichotomy just laid out is almost as common in discussions of free will as the distinction between compatibilism and incompatibilism. Nevertheless, it is a false one. According to event-causalism, an action is free *only if* it is caused by certain mental events. And, according to agent-causalism, an action is free *only if* it is caused by an agent. Two positions have emerged out of this dialectic: either we go in for an event causation story of free will, or we go in for a substance causation story of free will. But, whichever we go in for, the two are mutually exclusive; we cannot allow for both in our theory of free will.⁵

The options, however, are not mutually exclusive. Free will theorists have unwittingly accepted that, assuming mental events (such as reasons) cannot be causes of free actions (as does the agent-causalist), no events of any kind could be causes of free actions. But, mental events do not seem to exhaust the relevant possibilities—there appears to be at least another option that deserves consideration. To see this, however, some setup is required. So, in the remainder of this section, I will explain why there is another kind of event that deserves consideration as being the cause of free actions. Furthermore, I will use this insight to begin to sketch out how this may open up the possibility of a view on free will that allows for both event causation and substance causation.

So far, we have been relying on a pre-theoretic conception of events—roughly understood as things that “happen” or “occur”—but, it will be helpful to home in on a particular conception of events. To be more specific, I will understand events in the sense explicated by Jaegwon Kim (1973). While Kim’s conception of events is by no means the only conception of events on offer, it seems to be the most widely assumed conception of events, especially in the free will literature. Below we will return to the question of how much of our discussion hinges on this particular conception of events.

⁵ One notable exception is Clarke’s integrated agent-causal account. On Clarke’s (2003) account, a free action is caused both by an agent (a substance) and the agent’s reasons (a mental event). With that said, the Causal Pluralist view advanced here diverges from it in significant ways. For one, Clarke requires *both* substances and events to be fundamental causes, whereas Causal Pluralism does not. Additionally, if it turned out that *all* fundamental causes were of the same kind—say, events—Clarke’s account would be false. In contrast, Causal Pluralism would not be falsified. This means that Causal Pluralism has the resources to avoid the uniformity problem (see Clarke 2003: 207-209) that Clarke’s account faces.

With that said, Kim defends a property exemplification model of events. According to this model, an event is the exemplification of a property, *P*, by substance, *x*, at a time, *t*. Schematically, then, all events have the structure of: *x*'s *exemplifying P* at *t*. So, for example, our previously mentioned event of *the tree's falling* would be short-hand for something like: *the tree's exemplifying the property of falling at midnight*.

We need only three ingredients to have an event on this model: a substance, a property, and a time. What we are looking for is a candidate event that can be the cause of a free action. Candidate substances and times are fairly straightforward. The substance will be the agent who is acting freely, and the time will be whenever the proposed event occurs. All that is left is a candidate property. One such property is the aforementioned agent-causal power. This property seems like a plausible candidate, since agent-causalists maintain that the exercise of such a property is a necessary condition for acting freely. The agent, then, will always exemplify this property when acting freely. Since we are also looking for a candidate event that will always be present in the occurrence of a free action, the agent-causal power seems like a good fit.

With that laid out, we have the following candidate event which might play the causal role of bringing about a free action: *an agent's exercising her agent-causal power at a time*.

The proposal is that events of this type deserve consideration as potential causes of free actions. Just as traditional event-causalists maintain that free actions are caused by certain mental events, such as a person's having certain reasons, the suggestion here is that free actions could be caused by a distinct sort of event—namely, *the agent's exercising her agent-causal power at a time*. However, in contrast to standard agent-causalist accounts, the agent is not directly causing the action, but is doing so by means of the event just described.

This insight can be used to generate a further proposal. Call the view that freedom is compatible with both event and substance causation *Causal Pluralism*. On the Causal Pluralist proposal, free will is wholly explicable in terms of event causation. But, the event which may be the cause of a free action

is not the typical mental event, such as the agent's reasons. Rather, it is an event directly involving the agent—the agent's exercising her agent-causal power at a time. However, the Causal Pluralist does not maintain that event causation is necessary for free will. For the Causal Pluralist, free will is equally compatible with substance causation. An action may be free if it is directly caused by the agent, in virtue of exercising her agent-causal power. How we might distinguish between these two types of causes of free actions is discussed in the next section:

A. *The agent's exercising her agent-causal power at a time.*

B. The agent, in virtue of exercising her agent-causal power at a time.

It is important to clarify that Causal Pluralism is not a commitment to something like a dual aspect view, where both substance and event causation are really one and the same, but looked at, or described, in two different ways. Rather, the idea is that our being free does not necessarily turn on whether the causal relata involved in free action turn out to involve substances or only events. The Causal Pluralist may remain neutral about such matters—that is, neutral about whatever the fundamental causal relata turn out to be. She need only maintain that substance and event causation be of the right kind if an action is to be free.

To sum up briefly, then, Causal Pluralism maintains that free will is compatible with both event and substance causation (given suitable characterizations). On the one hand, if the agent directly causes an action in virtue of exercising her agent-causal power at a time, then the action may be free. But, on the other hand, if an action is caused by an event of the sort: *an agent's exercising her agent-causal power at a time*, then the action may also be free. This is not to say that a free action must be caused

by both a substance and an event, but merely that, if either causal story turned out to obtain, our free will might still be secured.⁶

To be sure, none of this is proof that this distinctive form of Causal Pluralism is true. That it is possible, though, and even appealing, seems a particularly important thing to appreciate, given how worked-over and seemingly intractable debates have been about free will. There may yet be a way forward for convergence amongst event and agent-causalists.

§2.4 A Challenge For Agent-Causalists

⁶ One might worry about a potential overdetermination problem, like the kind developed in Kim's causal exclusion argument. The worry would presumably arise by adopting something like the following three commitments: (a) every effect has a sufficient fundamental cause, (b) sometimes non-fundamental causes are causally efficacious, and (c) no single effect can have more than one sufficient cause, unless it involves genuine overdetermination (like firing squad cases) (cf. Franklin 2018: 186). (a)-(c) are mutually inconsistent, so something has to give. While it is an open question to what extent a Causal Pluralist needs to accept all of these claims, it is nonetheless worth highlighting some of the potential ways of addressing the issue.

However, a full discussion of the matter is beyond the scope of this chapter.

At first glance, a plausible way forward would be to utilize the strategies developed in reply to Kim's own causal exclusion argument. Borrowing from Bernstein and Wilson (2016), and Wilson (2021), I sketch in very broad strokes four example strategies one could employ in this context. For a more in-depth discussion of potentially available options, the reader is encouraged to look at the two aforementioned sources. For ease of explication, I will assume that substances are fundamental causes, and that events are non-fundamental causes, though everything that I say could be applied *mutatis mutandis* to the reverse position.

(1) To start, one could simply go epiphenomenal, and deny any genuine kind of causal efficacy to events. (2) Another option might be to go in for a kind of identity view, where you reduce event causation to substance causation. Overdetermination would be avoided because there is really only one cause, not two. (3) Alternatively, if one wanted to "preserve the reality, distinctness, and efficacy" of event-causes, one could opt-in for a non-reductionist position—inspired, for example, by those nonreductive realization accounts that implement a proper subset strategy. The idea here would be that event-causes can be efficacious in virtue of having some "non-empty proper subset of the token powers of" substance-causes. That is, the idea would be that events can be efficacious by inheriting some proper subset of powers possessed by the relevant substances, and cause effects in virtue of those particular powers. Problematic overdetermination (like firing squad cases) would be avoided because there would be one set of powers manifested, not two. (4) Finally, another route could be to employ a "strong emergentist" strategy, where events can maintain their autonomy from the relevant substances, while also being causally efficacious. This would, in effect, involve denying claim (a) above—that every effect has a sufficient substance (fundamental) cause—thus avoiding overdetermination.

Two differences—between Kim's causal exclusion argument and the kind of overdetermination at hand—are worth mentioning. First, Kim's causal exclusion argument deals with overdetermination among the same kind of causal relata—typically, events—whereas the kind of overdetermination here deals with overdetermination between different kinds of causal relata—i.e., events and substances. Second, as I'm understanding the notion here, events are things that involve substances. In contrast, writers have not typically thought of mental events as involving physical events (or vice versa) in an analogous way. In light of these two differences, there is reason to think that at least some of the strategies sketched above may need important amendments, if developed more fully.

Ultimately, I leave it open to what extent any of the available options will work out, as well as which option would be the best route for the Causal Pluralist.

Bringing the underlying machinery of agent causation to the surface and contrasting it with a property-exemplification model of events forces us to think more critically about the differences between substance and event causation. In the previous section, I introduced the Causal Pluralist view and distinguished between two types of potential causes of free actions within the framework. While some have made similar suggestions (see, for example, Nelkin 2011: 75-79), here, I hope to develop these ideas in more detail and provide further motivation for them. In particular, I do this by examining more closely the nature and metaphysics of events. Most writers—especially agent-causalists—seem to assume an overly simplistic account of events when contrasting event causation with substance causation—which, I think, has made it easier for them to be dismissive of event causation (in the context of free actions). In this section, then, I provide further motivation for thinking that both kinds of causes may be consistent with our being free. I start by providing a way of distinguishing between these two types of causes. After that, I consider a challenge for why agent-causalists in particular should take seriously the possibility that both causes are compatible with free will (as the Causal Pluralist maintains).

If all agent causation involves the exercise of an agent-causal power at a time, one might initially wonder why agent causation is not simply event causation by another name. Indeed, Clarke recognizes this concern and writes:

If a substance causes an event, it does so in virtue of having some causally relevant property. . . .The accommodation, however, comes perilously close to acknowledging that it is *the substance's having the property at the time* in question that is the cause.
(Clarke 2003: 201-202)

While the concern about whether substance causation collapses into event causation has not garnered much attention, some authors have briefly weighed in on the matter.⁷ For instance, in response to Clarke, Pereboom (2014: 54-58) contends that it is at least *conceivable* that an agent cause an effect independently of any causally relevant property. Since we can conceive of this kind of substance causation, then it really must be distinct from event causation. Though helpful, Pereboom's point does not provide us with a concrete way of demarcating substance causation from event causation, and so it may be worthwhile to consider alternative proposals.

Ann Whittle has also responded to Clarke's concern, but has gone further than Pereboom in providing a concrete means of demarcation. She writes:

[W]hile the views sound similar, they are not equivalent. To say that an effect occurred 'in virtue of' the laws does not thereby make the laws causes of the effect. Similarly, saying that an effect occurred 'in virtue of' a substance having a property at a time does not render the having of that property the cause. . . we may insist that the 'in virtue of' relation holding between substances and their powers need not be understood in terms of what the substance's properties cause. Rather, properties are responsible for the fact that substances cause things. (Whittle 2016: 16)

According to Whittle, substance causation differs from event causation with respect to the role that the causally relevant property (in our case the agent-causal power) plays. The idea is that the 'in virtue of' relation—in the claim *an effect occurred 'in virtue of' a substance having a property at a time*—can take on at least two meanings, one causal and one non-causal. On the causal reading, a substance's exemplifying a power is the cause of an effect. On the non-causal reading, a substance is

⁷ A closely related (and highly valuable) discussion can be found in Buckareff (2011) and Kuykendall (2019, 2021). The concern there is whether substance causation collapses into *powers causation*. I will have a little more to say about this dispute below, but it is worth flagging here that it is an interesting question whether Causal Pluralism could be extended to other kinds of causal relata. As noted earlier, the standard dichotomy in the free will literature is between event and substance causation, so I keep that as my focus here. I leave it for another time to address whether Causal Pluralism should be extended to other kinds of causal relata.

the cause of an effect, with the exemplification of the power playing a non-causal role in explaining how the substance caused the effect.

Whittle stops short of explicating how to understand this non-causal ‘in virtue of’ relation that she invokes. While she suggests that there is an analogy with the laws of nature, it is not entirely clear what the details are supposed to look like. Part of the problem is that there are a number of different ways to understand what laws of nature are—some of which threaten circularity in the present circumstances.⁸

One natural interpretation of Whittle’s non-causal ‘in virtue of’ relation is that of metaphysical dependence, or grounding. It is typically thought that one thing (fact, etc.) *F* grounds another thing (fact, etc.) *G* when (*inter alia*) *G*’s obtaining (non-causally) depends on *F*’s obtaining, and *F*’s obtaining (non-causally) explains *G*’s obtaining (Clark and Liggins 2012). To take an example, it is commonly thought that the singleton set {*Socrates*} is grounded in the person Socrates, since the existence of the singleton set depends on, and is explained by, the existence of the person Socrates.

Although the grounding interpretation has some plausibility, it is unclear if Whittle would be willing (or would have the need) to take on board all of the commitments that come along with metaphysical grounding. For that reason, I will leave it open what exactly the non-causal ‘in virtue of’ relation amounts to—although, it seems to me that metaphysical grounding still provides a useful approximation for trying to understand what the relation is.

To return to the main thread of our discussion, Whittle’s proposal provides us with a means of demarcating substance causation from event causation. In particular, the distinction looks something like this: *event causation* occurs when **the substance’s exemplifying (or exercising) the causally relevant property at a time** stands in causal relation to an effect. And, in contrast, *substance causation* occurs when **the substance** stands in causal relation to the effect, but the substance does so in virtue

⁸ For instance, if laws are summaries of how causal powers are exercised (Demarest 2017), then the analogy is not illuminating.

of (in some non-causal sense) the substance's exemplifying (or exercising) some causally relevant property (at a time).

Whittle's suggestion provides us with a helpful way of bringing out the potential differences between event and substance causation. The suggestion also seems to be in line with what current agent-causalists maintain about the metaphysics of agent causation (cf. Jacobs and O'Connor 2013). So, there is good reason to think that the proposal captures a plausible way in which agent-causalists might try to explicate the difference between event and substance causation.

Let us grant that Whittle's foregoing distinction provides a way of demarcating event causation from substance causation. While we might dispute which story provides us with the best metaphysics of causation, I believe a more important issue deserves attention.

The issue is this: granting the distinction between substance causation and event causation, is the difference between the two a *freedom-relevant* difference? Suppose Angelica makes an utterance. Could her acting freely in this case turn on which of these is the cause of her action?

A*. *Angelica's exercising her agent-causal power at a time.*

B*. Angelica, in virtue of (in some non-causal sense) her exercising her agent-causal power at a time.

It is not entirely clear that it would, as it is not clear that there is any freedom-relevant difference between the two causes.

In light of this, if the agent-causalist is correct that substance causation is compatible with our being free, then we need some principled reason for maintaining that *only* substance causation is compatible with our being free. That is, we need some principled reason for ruling out event causation as being compatible with free will. But, it is not clear there is any such reason for thinking this—that

there is some freedom-relevant difference. Because of this, we should take seriously the proposal that *both* substance and event causation are compatible with our being free.

Freedom-relevant differences come in many different shapes and sizes. Here are some examples that have been proposed throughout the years: the causal history of an action, the presence of indeterminism, the ability to do otherwise, and the reduction of agency to mental events. In the present circumstances, the onus is on the agent-causalist to show that there is more than a mere difference between substance and event causation. They must show that the difference between the two captures some *freedom-relevant* feature. And, as the list just provided indicates, there are several ways to do that. The important point here, though, is simply that arguing about potential freedom-relevant differences is a mainstay of the free will literature, and so it should not be seen as unusual to ask for one in this context.

It may be worth reminding the reader at this point that the foregoing challenge is meant to apply to both compatibilist and incompatibilist (i.e., libertarian) agent-causalists. Thus, if one is to propose a freedom-relevant difference between substance and event causation, it is important to be clear who the interlocutor is. For instance, suppose one thinks that the problem of luck shows that indeterminism (of the sort postulated by libertarians) undermines free will.⁹ If the Luck Objection is presented to a libertarian agent-causalist, this might show that they ought to give up being a *libertarian*, but not necessarily that they ought to give up being a compatibilist agent-causalist. Perhaps there are freedom-relevant differences that apply equally well to both compatibilists and incompatibilists. That is all well and good. The important point here is that we need to be careful who is being addressed when considering potential freedom-relevant differences.¹⁰

⁹ The problem of luck states, roughly, that indeterminism introduces luck in such a way that it undermines the control required for free will. See Mele (2006) and Levy (2011) for further discussion of this problem.

¹⁰ Some agent-causalists (cf. O'Connor 2000; Steward 2012) distinguish between a production view of action and a component view of action, and maintain that the former is crucial to a theory of free will. On the component view, actions are not events that are caused, but are rather agent-causings of events. One might wonder, then, whether adopting a component view of action provides a freedom-relevant difference that cannot be captured if all causes are events. I'm not persuaded that this provides a freedom-relevant difference which cannot be accounted for in terms of event causation. On the component

While I have stressed the significance of freedom-relevant differences, one might also wonder what role non-freedom-relevant differences should play in this discussion. As noted earlier, there are a variety of options when picking one's preferred causal relata—even within a neo-Aristotelian framework. Buckareff (2011, 2017) and Mumford and Anjum (2011) maintain that powers are causes. Kuykendall (2019, 2021), Whittle (2016), and Skow (2018) all maintain that substances are causes. And most of analytic philosophy over roughly the past century has worked under the assumption that events are causes. Each of the foregoing camps will have considerations and arguments in favor of their choice of causal relata. So, what should we say about these divides?

I think that these considerations will be important in picking one's preferred causal relata, but I also believe that they don't make a difference to free will. Potential considerations are likely to include things like: being the most metaphysically plausible, having the best fit with our best scientific theories, cohering with one's preferred theory of causation (e.g., a counterfactual account), and so on. While I take these all to be legitimate considerations, it is important to keep in mind that they don't show that *free will* specifically requires one or another causal relata. For example, suppose one prefers a counterfactual analysis of causation, and suppose further that one believes that this account of causation requires that events be causes. While this would give one reason to believe that all causes are events, it would not show that free will requires that events be causes. I take the same point to apply to all considerations that don't qualify as freedom-relevant differences.

I believe a two-fold challenge for agent-causalists can be drawn from our discussion. The first challenge is to show that substance causation does not collapse into event causation. The second challenge is to show that the *reason why* substance causation does not collapse into event causation is due to some freedom-relevant difference. While I have suggested one plausible way of addressing the

view, an action can be described as a complex event which has as its constituents a substance standing in causal relation to an event (such as the formation of an intention). The Causal Pluralist can maintain much of the same story. However, instead of the first constituent being a substance, it would be the event of an agent's exercising the agent-causal power at a time (cf. Clarke 2017: 10).

first challenge, I have also explained why I think this is not enough to address the second challenge. Since Causal Pluralism avoids the foregoing concerns, the agent-causalist has special reason for taking the proposal seriously.

Before moving on to the next section, it is worth pausing to highlight an important advantage of the Causal Pluralist view. In particular, I would like to suggest that Causal Pluralism provides a way of securing central aspects of agent-causalism, without the same ontological stringency.

Causal Pluralism is able to capture much of what agent-causalism offers in part by accommodating much of its metaphysics. Causal Pluralism takes as its starting point the agent-causal story, with irreducible agents and their agent-causal powers, and builds it into its own proposal. However, motivated by the fact that there is not a clear freedom-relevant difference between substance and event causation (given suitable characterizations), the Causal Pluralist also allows for event causation. But, it is important to notice that even the event allowed for in Causal Pluralism is not far off, ontologically speaking, from the substance-cause postulated by agent-causalism. In both cases, we have irreducible agents exercising their agent-causal powers. By allowing for such features no matter what the causal relata are, Causal Pluralism captures a core and guiding tenet of agent-causalism. Furthermore, it does this without restricting itself to fundamental substance causation.¹¹

§2.5 Events Reconsidered

One might naturally wonder whether the plausibility of Causal Pluralism, especially the challenge raised in the preceding section, turns on a specific conception of events—namely, the property-exemplification model. And, while I think it would be an interesting and important point if free will

¹¹ It is worth saying a few words about what distinguishes the account proposed here from others, such as Buckareff (2011). While there is clearly overlap in our discussions, I believe the aims are quite different. First, Buckareff argues that the fundamental causes are powers, whereas I argue that we should stay neutral between events and substances. Second, and relatedly, my concern is whether substance causation collapses into event causation, whereas Buckareff is concerned with whether substance causation collapses into powers causation. Finally, I argue that these points have important implications for the free will dialectic between event- and agent-causalists.

turned out to be compatible with both substance causation and only one form of event causation, I think there are ways of extending our remarks to other accounts of events. So, the goal of this section is to motivate the point that Causal Pluralism does not stand or fall with the property-exemplification model of events.

With that said, a few brief remarks are in order before moving on to the proceeding discussion. The first is that I will assume that substances (such as agents) and causal powers (such as an agent-causal power) are taken to be fundamental constituents in our ontology. I will then briefly canvas three different accounts of events and motivate the claim that agents and their agent-causal powers may be so intimately involved with these various construals of events that it is not clear that, if any of them should turn out to be correct, our being free would be undermined. My suspicion is that intuitions may vary from account to account for various readers, so that Causal Pluralism may seem to be more or less plausible depending on the account of events under consideration. It is important to state upfront, then, that it is not my aim to show that Causal Pluralism is equally plausible on every construal of events, but simply that there exists at least *some* flexibility when it comes to what theory of events a Causal Pluralist might endorse.

§2.5.1 States Of Affairs

Start first with a conception of events that diverges the least from Kim's. Roderick Chisholm (1990, 1992, 1994) defends an account where events are construed as states of affairs.¹² Chisholm construes states of affairs as the exemplification of properties by substances. Chisholm, then, departs from Kim only insofar as he thinks the *times* are not constituents of events. To provide an example for concreteness, a fully specified Kimian event would be something like: *Angelica's waving at 12am*,

¹² Early Chisholm (1970) had a much different conception of events. On his earlier account, events were a subspecies of propositions and thus were abstract entities. Chisholm explicitly gave up this view in later years for the one discussed above.

whereas a full Chisholmian specification of the same event would simply be: *Angelica's waving* (which has the property of occurring at 12am).

If a Kimian event is compatible with Causal Pluralism, then so is a Chisholmian event. The Chisholmian event-cause of our free actions would simply be *the agent's exercising her agent-causal power* (which occurs at a particular time). The only modification that would be needed to make room for this kind of event would be to drop the requirement that times are constituents of the events, and there's no obvious reason to think that this move would carry any complications that would threaten our acting freely.

§2.5.2 Spatiotemporal Regions

Now consider an account of events that departs further from Kim's. This account construes events as spatiotemporal regions (Davidson 1985; Quine 1985). We can get an intuitive grasp on spatiotemporal regions by first thinking of the universe in four-dimensionalist (or eternalist) terms, where the totality of the universe—past, present, and future—exist “all at once” in a four-dimensional block. Events can then be picked out by taking “slices” or subregions of the universe. For instance, an event such as *Angelica's waving at 12am* will occupy some subregion within the block, and that subregion will be the event.

Note that, while it is easier to get a grasp on the idea of spatiotemporal regions within four-dimensionalism, the Davidsonian need not be committed to such a view. To give just one example, some presentists—i.e., those who say that only the present exists—allow that the present consists of a (much) smaller, and constantly changing, four-dimensional block. This would allow a presentist to account for the existence of spatiotemporal regions in much the same way as the eternalist.¹³

¹³ See Dainton (2010) for a helpful summary of presentist options.

At face value, such an account might seem highly antithetical to the Causal Pluralist project, given how far the account departs from our initial conception of events. However, things are less problematic than they initially seem. As long as we are content with saying that agents are physical substances that are located in time and space, we can say that the relevant events (i.e., spatiotemporal regions) involving free actions are simply *delimited* by the location of the agent (cf. Bennett 1989). When an agent acts freely, the agent exercises her agent-causal power at a time and place. The event of the agent exercising her agent-causal power would then be the exact—no more, no less—spatiotemporal region where the agent exercises her power. Conceiving of events as delimited by the substance involved would seemingly “blur” the distinction between substance and event causation (cf. Simons 2003).

§2.5.3 *Properties Of Spatiotemporal Regions*

Consider one last account of events—one that departs even further from Kim’s account than the one just considered. On this alternative conception, events are a special class of properties—they are properties of spatiotemporal regions (Lewis 1986; cf. Bennett 1988). This account builds on the previous one by incorporating spatiotemporal regions into its analysis of events, but does not go so far as to reduce events to those spatiotemporal regions.

Conceiving of an event as a property is *prima facie* puzzling. What *kind* of property is this? Lewis himself admits to only giving constraints for when a property is “formally eligible” for being an event, though the constraints provided are not particularly helpful in this context.

To get a better fix on how we might understand Lewis’ claim that events are a certain class of properties, it may be helpful to consider some of what Richard Montague has said on the matter. Montague has defended a nearby version of Lewis’ account—one according to which events are properties of times, rather than properties of spatiotemporal regions. My interest is not in Montague’s own account, but in the following comments he makes:

A third possibility, and one that seems to be the only reasonable suggestion, is to take as the event corresponding to a formula the property expressed by that formula. Thus the event of *the sun's rising* will be *the property of being a moment at which the sun rises*, and events in general will form a certain class of properties of moments of time. (Montague 1969: 160, emphasis added)

The general strategy proposed here appears to be to, first, take what seems (pre-theoretically) to be an event—such as *the sun's rising*—and then, second, make a property out of it—such as *the property of being a moment at which the sun rises*.

Since Lewis admits to making a similar kind of proposal, it seems we could apply Montague's remarks to Lewis' own account. This would mean that events correspond to properties such as *the property of the sun's rising*; but, instead of attributing such properties to times, we would simply attribute them to spatiotemporal regions. While this does not give us a principled way of individuating events, it at least should give us an intuitive grasp on the proposal that Lewis is making.

It may be worth addressing that Lewis takes properties to be sets, and so ultimately reduces events to sets of regions spread across different concrete possible worlds. We may not want to adopt all of Lewis' machinery, so it is worthwhile to flag that options are available. For instance, Bennett (1988) proposes an account of events that takes them to be *tropes* that are had by spatiotemporal regions, where tropes are said to be “abstract particulars” or “property instances.” Since Bennett takes events to be properties of sorts (which are had by spatiotemporal regions), he can be understood as proposing a variation of Lewis' account. I see no reason why, then, one could not also adopt a different account of properties on this matter.

We have said that, according to this account, events *are* properties of spatiotemporal regions. For example, *Angelica's waving* is an event and a property. It is an event because it is a specific sort of property. And this property is instantiated at a region—namely, the region where Angelica waves. So, for Lewis, there is an event at that region because this property is instantiated at that region.

The same can be applied to cases where an agent acts freely. If an agent acts freely, she thereby exercises her agent-causal power to freely cause some action. The exercising of her agent-causal power may be described as a property: *the property of exercising her agent-causal power*. We may thereby also characterize that property as an event. *The agent's exercising her agent-causal power* is a property that is instantiated at the very region where the agent exercises her agent-causal power, and it is this kind of event that may be thought to cause our free actions.

Before concluding, it is worth taking a moment to make a few comments. First, the foregoing accounts by no means exhaust all of the available accounts of events and the possible permutations.¹⁴ Such a task would take us far beyond the purposes of our present one. (I leave the task to the reader to plug in their preferred theory of events and check the results.) Our present task has been to show that Causal Pluralism does not stand or fall with a property exemplification model of events. There exist other accounts of events one may endorse while also endorsing Causal Pluralism. To lay my cards on the table, Chisholmian events and events conceived of as spatiotemporal regions seem just as consistent with Causal Pluralism as Kimian events. I am less confident about events conceived of as properties of spatiotemporal regions, though this may be due to my uneasiness about talk of properties as causes more generally. Of course, intuitions may vary.

It seems to me that the spatiotemporal regions theory is more in line with how philosophers of science tend to think of events, so it may be of special interest how well it meshes with Causal Pluralism. And, on this score, I think Causal Pluralism does well enough. So long as we grant that substances and powers are fundamental constituents in our ontology, we can construct events using these materials in much the same way that a Kimian constructs events—that is, out of substances, properties, and times.

¹⁴ See, for example, Lombard's (1986) account that conceives of events as change. See also Simons (2003) and Casati and Varzi (2020) for helpful surveys of different accounts of events.

All in all, I hope to have motivated the point that the plausibility of Causal Pluralism does not stand or fall with the Kimian conception of events.

§2.6 Concluding Remarks

Causal Pluralism gives us a new perspective on a number of important issues pertaining to free will. While I do not have the space to fully address all such issues here, I would like to stop and take note of one issue in particular: agency reductionism. Space constraints preclude a fuller exploration of agency reductionism; but, before concluding, I would like to (i) sketch out how I think Causal Pluralism could address the matter, and (ii) draw out an interesting upshot of having the Causal Pluralist view on the table.

Agency reductionism concerns whether “the causal role of the agent in all agential activities (from nonintentional to autonomous action) is reducible to the nondeviant causal activity of appropriate mental states and events involving the agent” (Franklin 2018: 15). It is standard, then, to characterize traditional event-causalist views as adhering to agency reductionism, since they maintain that all actions (including free actions) are caused by mental events. And, it is also standard to characterize traditional agent-causalist views as adhering to agency non-reductionism, since they maintain that at least *some* actions are caused by substances.

A natural question is whether Causal Pluralism is best characterized as adhering to agency reductionism or non-reductionism. Under the Causal Pluralist view, the question divides itself: (a) if the (fundamental) causes of free actions turn out to be substances, is Causal Pluralism committed to agency reductionism? And (b), if the (fundamental) causes of free actions turn out to be events (of the sort specified earlier), is Causal Pluralism committed to agency reductionism?

I think that, if we start by considering (a), Causal Pluralism will straightforwardly classify as agency *non*-reductionist. This is because the view—under the assumption that the fundamental causes

of free actions are substances—simply adopts the same metaphysical story as traditional agent-causalist views. The more interesting question, it seems to me, is what we should say about (b).

I believe that Causal Pluralism is also best classified as adhering to agency non-reductionism *even if* the (fundamental) causal relata of free actions turn out to be events. The reason for thinking this is that the kind of events allowed for by Causal Pluralism are crucially and relevantly different from the kind of events postulated by traditional event-causalist views. The guiding idea behind agency reductionism, I take it, is that the agent (or agential activity) can, in some sense, be wholly decomposed into the workings of the subcomponents or subsystems of the agent (see, for example, Velleman 1992; cf. Kane 1996, 193-194). This is what makes traditional event-causalist views reductionistic about agency. Mental events postulated by traditional event-causalism, such as beliefs and desires, are events that occur within, or inside, the agent. Traditional event-causalism, then, reduces the agent (or agential activity) to the workings of the mental subsystems of the agent.

The kinds of events that Causal Pluralism would allow for are not events that take place within the agent. Rather, the events postulated are pitched at the level of the agent (*qua* substance). This means that there is no attempt to reduce the agent (or agential activity) to the workings of the subcomponents of the agent. What this indicates is that Causal Pluralism is not committed to agency reductionism, whether the first causal relata are agents or events (of the sort previously specified).¹⁵

Though not stated explicitly, authors sometimes seem to talk as if agent causation is the only way to avoid agency reductionism (cf. Velleman 1992; Franklin 2018). But, an important upshot of the preceding discussion is that the two seem to come apart: one can be an agency non-reductionist while still allowing for event causation. This is because it is not the adoption of event causation *per se* that makes an account agency reductionist, but rather, the adoption of certain kinds of events.

¹⁵ An agent-causalist (or Causal Pluralist) need not deny that such subcomponents and subsystems are explanatorily relevant, only that the agent (or agential activity) is wholly reducible to such things.

Given the foregoing, Causal Pluralism looks to have at its disposal the resources to develop a promising line of reply to formidable objections to standard event-causalist accounts, such as Pereboom's (2014a, 2014b) *Disappearing Agent Objection* and Franklin's (2016, 2018) *It Ain't Me Argument*. Arguments like those developed by Pereboom and Franklin aim to undercut traditional event-causalist accounts by showing that the mental events that are purportedly the cause of free actions fail to supply what is needed for free will. In the case of the Disappearing Agent Objection, it is contended these mental events cannot account for an agent's settling a free action; and, in the case of the It Ain't Me Argument, it is contended that they cannot account for an agent's self-determining a free action. Both authors maintain that agent causation presents the most promising solution to their respective arguments.¹⁶

We do not need to get into all of the details of the arguments here. What is important for our purposes is concentrating on what lies at the heart of these objections. It seems to me that these arguments get their traction by leveraging the distinction between an agent (conceived of as a substance) and the agent's mental events. In doing so, it is easy to get an intuitive grip on why one might think that the agent's mental events cannot account for their settling a free action, or for their self-determining a free action. These arguments, it seems to me, provide lucid ways of articulating the concern raised by Richard Taylor (1992, 51): "If I believe that something not identical to myself was the cause of my behavior—some event wholly external to myself, for instance, or even one internal to myself, such as a nerve impulse, volition, or whatnot—then I cannot regard the behavior as being an act of mine, unless I further believe that I was the cause of that external or internal event."

Now, with all that laid out, it's not clear to me that these arguments are going to have the same kind of pull when applied to Causal Pluralism (*qua* its event-causal specification). The kind of event

¹⁶ While Pereboom believes that agent causation is sufficient to account for free will, he also believes (for empirical reasons) that we are *not* agent-causes in the actual world (Pereboom 2014, 65-69). And, although Franklin was on the fence about agent causation in Franklin (2018), he seems to have fully embraced it in Franklin (2019).

allowed for by Causal Pluralism is pitched at the level of the agent. There is no effort to reduce the agent's causal contribution in free action to the causal contribution of the agent's mental events (or mental subcomponents). The concerns that the agent disappears from the picture, or that it really "ain't" the agent determining the action, consequently don't have the same kind of force when targeted at Causal Pluralism.

Furthermore, recall that traditional agent causation is taken as a solution to both authors' arguments. Since Causal Pluralism adopts much of the same underlying metaphysics in its event-causal specification, it seems to me that both arguments are going to have a difficult time consistently maintaining that traditional agent causation is a solution to their respective arguments, but that Causal Pluralism in its event-causal specification is not.¹⁷

Let us now step back and take stock of some of the central ideas that have been put forth in this paper. First and foremost, I have tried to introduce a novel view on the issue of free will—namely, Causal Pluralism. Causal Pluralism maintains that free will is compatible with both fundamental event and substance causation. Thus, the view allows one to stay neutral about what the fundamental causal relata are, so long as the relata are of the right sort.

Second, while I think Causal Pluralism deserves wider consideration, I have also tried to show that the demand to consider Causal Pluralism is all the more pressing for extant agent-causalists. Insofar as the move to adopt Causal Pluralism is more natural for the agent-causalist, agent-causalists ought to reconsider the fact that they have unnecessarily wedded themselves to substance causation, thereby making their account more ontologically demanding than it needs to be. Causal Pluralism may then be a way of securing the goods of agent-causalism without all the ontological demands.

Third, I hope to have shown that, independent of whether one is ultimately attracted to the view, having Causal Pluralism on the table helps us to see a number of issues in a new light, such as: (a)

¹⁷ It is an open question, of course, whether these arguments can be reformulated to raise problems for Causal Pluralism.

there are overlooked ways to address the dichotomy when choosing between event and substance causation in a theory of free will, and (b) event causation may come apart from agency reductionism.

No doubt further issues remain that deserve consideration. For instance, a fuller and more systematic treatment of how well Causal Pluralism fares against its competitors is in order, as is a closer look at the nature of agency reductionism. Furthermore, although the focus has been on the event-substance divide in the context of free will, there remains the possibility of extending the considerations made here to the nature of action more broadly. Regardless, such treatments will have to come at a later time. For now, agent-causalists who wish to resist Causal Pluralism would do well to reconsider the relation between substance and event causation, as well as what substance causation secures that event causation cannot.

Chapter 2, in part, has been accepted for publication (as “Causal Pluralism: Agent Causation Without the Panicky Metaphysics”) in *Synthese*, 2023, Martinez, Joseph. The dissertation author was the sole author of this paper.

Chapter 3

Agents and Laws:

Reconsidering Pereboom's Wild Coincidence Objection

§3.1 Introduction

In the previous chapter, we addressed the concern that agent causation relies on a metaphysical picture that is too ontologically demanding. In this chapter, I turn considering an empirical criticism of agent-causal libertarianism. While empirical critiques to agent causation abound, one of the most forceful empirical objections in recent years comes from Derk Pereboom (2001, 2014), known as the Wild Coincidence Objection. The objection's aim is to show that our current empirical evidence regarding the laws of nature conflicts with agent causation. Roughly, Pereboom's Wild Coincidence Objection argues along the following lines:

On this view, when an agent-cause makes a free decision, she causes it without being causally determined to do so. But she at some point would affect the physical world distinct from the agent-cause. On our best physical theories the physical world is law-governed....On the path to action that results from this undetermined decision, alterations in the physical world, for example in her brain or some other part of her body, are produced. But it would seem that we would at this point encounter divergences from the deterministic laws....One might object that it is possible that the physical alterations that result from free decisions just happen to dovetail with what could in principle be predicted on the basis of the deterministic laws, so nothing

actually occurs that diverges from these laws. But this proposal would, at least prima facie, involve coincidences too wild to be credible. (Pereboom 2014, 65-66)

He continues:

At this point, the libertarian might venture that there are indeed departures from the probabilities that we would expect on the basis of the physical laws, and they are likely to be found in the brain....The problem for this proposal, however, is that we have no evidence that such divergences occur. (Pereboom 2014, 69)

In short, the agent-causalist faces a dilemma. She must either claim there are no deviations from the laws of nature or that there are deviations. If she claims that there are no deviations, this would be a coincidence “too wild to be credible.” And, if she claims that there are deviations, then she is making an empirical claim for which we lack evidence. Either way, the agent-causalist has sufficient reason to give up her position.

Recently, however, a number of authors (Baker 2017; Runyan 2018; Taggart 2020) have called into question the argument’s plausibility. Jordan Baker (2017, 3091), for instance, contends that Pereboom’s objection falls short because the “objection fails to engage the target theory.” And Christopher Taggart (2020, 15) claims that the “Problem of Wild Coincidences does not show that current empirical evidence makes agent-causal libertarianism unlikely...even if all events are governed by physical laws...and agents can overrule those laws, even though they never do.” Still yet, Jason Runyan (2018, 4564) argues that the objection fails because “current findings from neuroscience would suggest that the antecedent probability that certain behavior will occur fluctuates from one occasion to the next, and is subject to change depending on what one does.”

The goal of this chapter is two-fold. First, I aim to show that the Wild Coincidence Objection is more forceful than previously realized. I do this by showing that a proponent of the Wild Coincidence Objection has the resources to address the foregoing responses. Second, in light of this, I consider an

alternative line of reply to the Wild Coincidence Objection that involves calling into question whether the world is wholly governed by the laws of physics. I explain what this commitment would involve and how it provides a more promising way out of the Wild Coincidence Objection.

Here is an overview of the rest of the chapter. In §3.2, I provide a few brief remarks about the commitments of agent-causal libertarianism that will be relevant to the discussion of the Wild Coincidence Objection. In §3.3, I lay out the objection. Along the way, I expand the Objection by developing an Explanatory Exclusion Problem. Following that, in §3.4, I explain how this modified version of the Wild Coincidence Objection has the resources to address the responses of Taggart, Baker, and Runyan. Finally, in §3.5, I outline an alternative way of addressing the Wild Coincidence Objection—namely, by calling into question whether the world is wholly governed by physical laws. I conclude by considering the tenability of this proposal and how it fares with respect to alternative responses to the Wild Coincidence Objection.

§3.2 Agent Causation and Indeterminism

There are two core commitments that will be relevant to our discussion of Pereboom's objection. The first relates to agency non-reductionism and the second to libertarianism. While I believe that agency non-reductionism is the real target of the Wild Coincidence Objection, because libertarianism is relevant at certain junctures in the argument, it is worth making both explicit.

1. Agency Reductionism: An action is free only if the agent is the direct cause of that action.
2. Libertarianism: An action is free only if the agent is not causally necessitated to perform that action.

In the previous chapter, we examined claim (1) in detail. The proposal there was that there are at least different, but permissible, ways to interpret this claim. The first was that the agent *qua* substance must stand in causal relation to the action. The second was that the agent *qua* substance need not stand in

direct causal relation to the action; instead, it is also possible that some agent-level event be the cause of the action—an event such as *the agent's exercising her causal power at a time*. Although Pereboom's objection targets the former interpretation, I believe that the objection can also be formulated to target the latter interpretation as well. The reason is that, on the picture sketched in the previous chapter, this event is still a non-reductive event in the sense that an agent's causal contribution cannot be wholly explicated in terms of lower-level events.

As for claim (2), this stipulates an indeterministic condition for agent causation. It says that an agent must not be causally determined to perform an action if that action is to be free. Or, to put it slightly differently: at the moment just prior to acting, it must be open to the agent to perform that action or not perform that action. If Ellie freely raises her arm, then it must be within Ellie's power to not raise her arm (and vice versa).

Granting—as does Pereboom—that this picture of agent causation is coherent, it is a further question whether the world is hospitable to these requirements. That is, even if it's metaphysically possible that agent causation is instantiated, it's still an open question whether *our* world is set up in such a way to allow for agent causation. One way in which this question might be assessed is by examining whether agent causation is consistent with the general order that the world exhibits; or, that is, with the laws of nature. In the next section, we turn to considering Pereboom's objection that agent causation conflicts with the empirical data regarding the laws of nature.

§3.3 The Objection

The Wild Coincidence Objection takes the form of a dilemma. Its general structure looks like this:

1. The world is wholly governed by physical laws.

2. So, if agent causation is true, then agents will either (a) (freely) conform to the predictions of the physical laws or (b) (freely) deviate from the predictions of the laws.
3. (a) should be rejected. (Horn 1)
4. (b) should be rejected. (Horn 2)
5. So, we should reject agent causation.

Let's start from the beginning and work our way through the argument.

Premise 1. I take premise (1) to constitute the foundation of the argument. At various points, Pereboom makes the following sorts of claims: “on our *best physical theories* the physical world is law-governed” (Pereboom 2014, 66, emphasis added); and, if agent causation were to hold, “such agent-causes would be embedded in a world that, by the evidence that supports our current theories in physics, is nevertheless wholly governed by *the laws of physics*” (Pereboom 2001, 79, emphasis added).¹⁸ Interestingly, however, despite the foundational role the premise plays, Pereboom does not provide much commentary on what he means when he says that the “world is wholly governed by the laws of physics.”

Nonetheless I believe we can reasonably infer what the claim roughly consists in given what Pereboom goes on to say later in the argument. I take the main idea to be that the sciences currently strongly suggest some form of physics reductionism. We can break the idea of physics reductionism down into two parts. First, everything that exists is composed of the same fundamental particles posited by quantum mechanics, such as the fermions and bosons. Second, we appear to have physical laws that govern and explain the behavior of this fundamental stuff, laws such as Schrödinger's equation and

¹⁸ Often times, the phrase “law-governed” is reserved for describing a particular view of laws of nature where laws are said to govern (Armstrong 1983) rather than describe (Lewis 1973). I take it that Pereboom's argument is not actually intended to turn on a governing conception of laws. The primary reason for this is that it is highly contentious whether our best physical theories actually require an Armstrongian conception of laws. For this reason, I try to stay neutral about the exact metaphysics of laws.

the Von Neumann equation. Putting these two claims together, it follows that everything that happens is governed and explained by these fundamental physical laws. For instance, consider a soccer ball rolling down a hill. It, just like everything else, is wholly composed of fundamental particles. To explain how the ball moves down the hill, then, just is to explain the movement of the particles that constitute the ball. Because the laws can explain the movement of those particles, the laws can explain the movement of the ball down the hill.

The same goes for (purportedly free) human actions. A person is a physical entity composed of fundamental physical stuff like the soccer. Suppose Ellie raises her arm. To explain this action would be to explain the behavior of the particles that constitute her (arm). Because we have laws that can explain the behavior of these particles, just as in the case of the soccer ball, we can explain Ellie's action.

But how do mental events fit into this picture? Presumably when an agent acts, various mental events—such as beliefs, desires, intentions—are causally relevant. Pereboom provides a few ways to square this idea with physics reductionism (Pereboom 2001, 63–79). The first is to go in for some kind of reduction of the mental to the physical. For instance, perhaps beliefs, desires, intentions, etc., *just are* certain physical events, such as brain events.

An alternative option is to embrace (what Jessica Wilson 2021 has recently described as) *weak emergence*. The basic idea behind weak emergence is to grant that there is a real distinction between the mental and the physical, but cash out the distinction between the two in terms of differing causal powers. For instance, suppose that mental event m_1 supervenes on physical event p_1 . The idea is that p_1 will be associated some set of causal powers, say $\{A, B, C\}$, and m_1 will inherit some *subset* of those causal powers, say $\{A, C\}$. Still, what's important on this view is that mental is entirely constrained by the physical. A mental event always instantiates a causal power that is likewise instantiated by the physical event on which it supervenes. This allows us to accommodate some minimally robust form of

mental causation while also retaining the idea that everything that occurs is wholly governed by fundamental physical laws.

Premise 2. Physical laws allow us to make predictions about the world. So, it follows that, if the physical laws wholly govern the world, they will make predictions about everything that happens (or could happen).¹⁹

As Pereboom's notes, it seems to be a matter of dispute as to whether our physical laws are deterministic or probabilistic. To say that the laws are deterministic is to say that, holding fixed the laws of nature and the state of the world at any moment, then there is only one way for the world proceed—there's only one possible future state of the world. Whatever the future state of the world is, it has a probability of 1 of occurring, and all other possible future states of the world have a probability of 0 occurring.

It is commonly thought that the Newtonian mechanical laws that govern the movement of planets are deterministic. If so, then holding fixed the state of a planet, say Mars, and the Newtonian laws, then there is only one possible future movement of Mars. That particular movement has a probability of 1 of occurring. All other movement of Mars have a probability of 0 of occurring.

In contrast, probabilistic laws, don't always specify one unique future, but may specify a range of possibilities all with a certain probability of occurring. So, if the laws are probabilistic, the possible future states of the world will be assigned a probability *ranging from* 0 to 1 of occurring.

The rate at which radioactive atoms decay is generally taken to be probabilistic and governed by exponential decay laws. Let's suppose that at some particular moment a radioactive atom has 70% chance of decaying and 30% chance of not decaying. Holding fixed the laws and the state of the atom at that time, the atom's decaying is assigned a probability of .7 of occurring and the atom's not decaying is assigned a probability of .3 of occurring.

¹⁹ Or at least the behavior of the physical components can be wholly explained in these terms.

The relevant upshot is that, whether the fundamental laws turn out to be deterministic or probabilistic, every event will be assigned *some probability* of occurring. However, because the agent-causalist denies that determinism is true, let's grant for the sake of argument that the fundamental physical laws are probabilistic.

With all that in place, let's focus on a possible concrete prediction that the laws might make. Suppose that there is a .7 probability that Ellie raises her arm each time a question is asked in a class and that there are 10 questions that are asked over the course of the class. The laws tell us to expect that Ellie will raise her arm (close to) 7 times. Now further suppose that Ellie has free will (in the agent-causal sense). What should the *agent-causalist* expect to happen? Well, conceptually, there are only two possibilities: (a) Ellie freely raises her arm (close to) 7 times, or (b) Ellie freely does not raise her arm (close to) 7 times. Option (a) amounts to saying that Ellie will freely *conform* to the expected frequencies predicted by our laws and option (b) amounts to saying that Ellie will freely deviate from the expected frequencies predicted by our laws. I take the relevant kind of deviation here to be a contrastive notion, where the *actual* frequency of cases doesn't match up with the *expected* frequency ascribed by our laws. For instance, suppose the following is the case:

7. The laws *L* and the state of the world *H* entail that *Ellie raises her hand at time t*, will occur with a probability of 1.
8. *L* are true physical laws.
9. *H* obtains.
10. Ellie does not raise her arm at time *t*.

In this case, a deviation from the laws would have occurred, because the laws tell us to expect Ellie to raise her arm, but she does not raise her arm.

Though the details differ, the same point applies to probabilistic laws. Although probabilistic laws cannot be falsified by individual cases, they can be falsified by sets of cases. Deviations from probabilistic laws occur when the actual distribution of cases falls outside of the expected frequency predicted by those laws.²⁰ For example, if the laws tell us to expect Ellie to raise her hand 7 times over the next 10 occasions, and Ellie only raises her hand once, then a deviation would have occurred.

It is here the dilemma is meant to kick in. Pereboom wants to argue that, whether the agent-causalist embraces option (a) or option (b), problems arise. And because the options are exhaustive, we should reject agent causation. Let's consider the options in turn.

Premise 3. Now consider the first horn of the dilemma. Here, the agent-causalist anticipates that Ellie will raise her arm (close to) 7 times. I believe that there are roughly two reasons for being skeptical of this option. The first reason is that there is *no principled reason* to expect conformity; and the second reason is that, even if there were *some* principled reason to expect conformity, the better explanation would be the one involving fundamental physical laws.

To bring out the force of these concerns, let's characterize the problem in terms of competing explanations.²¹ The issue is explaining why a given phenomenon occurred, in this case the fact Ellie raises her arm (close to) 7 times. We have two competing explanations. Explanation 1 involves some complicated story about the fundamental physical laws governing the fundamental physical parts of Ellie in such a way that she raises her arm (close to) 7 times. This first explanation is a reductionist explanation that makes no mention of anything like agent-causal free will. Explanation 2, in contrast, involves making reference to *Ellie* freely causing the raising of her arm (close to) 7 times. This second

²⁰ While you cannot logically falsify probabilistic laws, you can empirically falsify them. The greater the deviation from the expected distribution, the stronger reason we have to think the probabilistic laws are false. (cf. Popper 2002)

²¹ Doing this also helps to side-step certain concerns that would otherwise seem to arise. In particular, one might worry that the Wild Coincidence Objection fails to get off the ground if one endorses a regularity view of laws. For, if the laws merely describe regularities, then surely there can be no wild coincidences about agents conforming to the laws. By making it clear that there are competing explanations for any purportedly free action, one that excludes agent causation and one that includes agent causation, then it seems to me irrelevant which theory of laws one endorses. The question is merely pushed back to which laws are the correct laws—that is, what is the *content* of the laws that give us the best explanation for what occurs.

explanation is non-reductionist. It denies that you can wholly explain why Ellie (freely) raised her arm by only making reference to fundamental physical laws governing her fundamental physical parts.

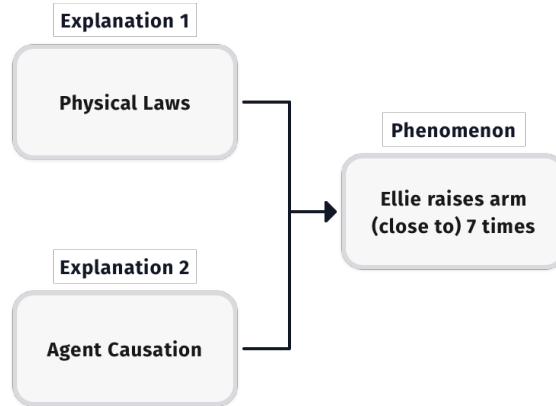


Figure 3.1: Competing Explanations

Merely granting that Ellie happens to conform to the predictions of the physical laws in this case might seem innocuous—and indeed it might be if it were merely an isolated incident—but it’s important to keep in mind that this is only one example, and that the laws will make countless other predictions about what agents will do. Going in for option (a), then, is tantamount to saying that agents will *always* freely conform to expected frequencies predicted by our laws. And this, Pereboom objects, would be too surprising “to be credible”—it would be a *Wild Coincidence*. It just seems *prima facie* true, the thought goes, that agents would deviate from the laws at some point, *given that they can* (Pereboom 2014, 66–67).

Consider an analogy. Imagine that I had a large jar filled with letters on my table. Now, imagine that every time I spilled the jar of letters onto my table, some sentence from *The Brothers Karamazov* was formed. Let’s suppose, now that we have two competing explanations. On the one hand, for the first explanation, imagine that I explained to you that the jar and the table were some kind of toy that I had purchased online, and that the letters were designed to form random sentences from *The Brothers Karamazov* when spilled onto this table. On the other hand, suppose someone insisted that the letters

just happen to fall this way. There is no principled explanation. That's just how it is. Given that we have both explanations in hand, which explanation should we prefer? Presumably the first explanation.

The agent-causalist seems to be in much the same situation when it comes to explaining (purportedly) free actions, such as Ellie raising her arm (close to) 7 times. If so, the agent-causalist should deny that agents freely conform to the predictions of the physical laws.

However, suppose that the agent-causalist did have some principled explanation for why we should expect conformity to the physical laws (we will canvass some one such option in §3.5.1). This would still not be sufficient for addressing this first horn of the dilemma. In addition to providing a principled explanation, the agent-causalist must also show that this explanation is empirically superior to (or at least as empirically plausible as) the explanation involving the physical laws. Short of this, the agent-causalist should reject this first horn of the dilemma.

Premise 4. Now consider the second horn of the dilemma, where the agent-causalist instead maintains that we should expect deviations from the laws (Pereboom 2014: 69). While this is conceivable, the problem with this line of argument is that there appears to be no empirical evidence to support the claim that there are such deviations from the physical laws. In essence, the agent-causalist would be making an empirical prediction that we lack empirical evidence for. And “without evidence for the departures from natural law that this view predicts, we have insufficient reason to accept it” (Pereboom 2014, 69).

§3.4 The Objection Reconsidered

Before examining some extant replies and I think they're insufficient on their own, I want to outline what I take to be the most plausible line(s) of reply to Pereboom's objection. The first, and perhaps less substantive, line of reply involves putting pressure on the reasons given for abandoning the second horn of the dilemma. The second, more substantive, line of reply puts pressure on the idea that the world is wholly governed by fundamental physical laws.

§3.4.1 The Second Horn Reconsidered

Let's start by reexamining the second horn of the dilemma and Pereboom's reasons for rejecting it. While I take my points here to be less definitive than the line of reply I want to advance in the next section, I believe it is still worth considering for at least two reasons. The first is that, as far as I can tell, no one has explicitly tried to defend the second horn of the dilemma. And the second is that it helps to build my cumulative case against the Wild Coincidence Objection that I will advance in the next section.

Recall that the second horn of the dilemma is the one wherein the agent-causalist says that we should expect Ellie (and free agents in general) to deviate from the predictions of the fundamental physical laws. Pereboom argued that this position should be rejected because it involved making an empirical prediction for which we lack empirical evidence. And without the empirical evidence to back up the prediction, the view, along with its prediction, should be rejected. The reasoning seems to rely on some more general principle with the following form:

Principle: If a view v makes an empirical prediction p , and we have no empirical evidence for p , then we should reject v .

This principle seems to me much too strong. Theories often make predictions that are neither verified nor disconfirmed for various reasons. For example, Einstein's relativity theory made various predictions which took many years to be verified, and physicists still accepted the theory prior to many of those verifications. In fact, relativity theory still makes predictions that have not yet been fully verified, but that does not mean we ought to reject.

Putting to the side cases in the sciences, it seems me that Pereboom's own argument fails to meet the principle. As previously noted, Pereboom's claim that the world is wholly governed by fundamental

physical laws entails that every event has some specified objective probability of occurring. However, we currently lack knowledge of what such probabilities could be. The claim that the world is wholly governed by fundamental physical laws makes an empirical prediction that lacks empirical evidence. It would follow, according to this proposed principle, that we should reject that the world is wholly governed by physical laws.

Perhaps there is some way to finesse the principle so that the objection goes through as intended. Though I'm not sure, I'm suspicious that it could be done. Regardless, the discussion helps to bring out an important point which extant authors have gestured at various places in their replies, which is that we currently lack the ability to test whether agents conform to or deviate from some predictions made by our fundamental physical laws. Without any way of testing such predictions, the argument as a whole appears dialectically weak.

This points to what I think is really the crux of the objection: namely, whether physics reductionism is true. The claim that the world is wholly governed by physical laws just is a specific form of physics reductionism. So, if we have strong empirical evidence that physics reductionism is true, we have strong prima facie reason to believe that agent causation is true. That's what I want to examine in the next section.

§3.4.2 The Laws Reconsidered

Within most circles in analytic philosophy, some form of physics reductionism is typically taken for granted. I think that this partly explains why Pereboom doesn't provide any specific evidence for the claim that the world is wholly governed by the laws of physics. In this section, I want to pressure on this idea. More specifically, drawing on recent work in the philosophy of science, I want to argue

that we currently *lack sufficient reason* to believe that the world is wholly governed by our fundamental physical laws.²²

Before beginning, a few brief clarifications are in order. I want to put pressure on the idea that the world is wholly governed by physical laws by putting pressure on both of these claims. We should note, however, that merely showing these claims to be false in their unqualified sense may not be enough to show that Wild Coincidence Objection does not go through. For example, suppose that there is one faraway patch of the universe that is not wholly governed by physical laws, thus falsifying horizontal and vertical reductionism. At face value this would not undermine the Wild Coincidence Objection insofar as it is compatible with *everything about persons* being wholly explainable by the physical laws. And it is this latter fact that we are primarily concerned with. What this shows, then, is that we don't merely need to put pressure on these claims in any old way, but we need to pressure on them in a way that allows us to doubt that human actions are wholly governed by physical laws.

My strategy for doing this is to put pressure on the positive case for thinking that human actions are wholly governed by physical laws. And I take this positive case for thinking this not to be a direct one, but rather an indirect and cumulative one. We don't simply apply our laws of physics to predict human actions, but we infer that the laws wholly govern human action given how much of the rest of the world we can explain with them. So, I want to say that this general inference is undermotivated because horizontal and vertical reductionism likely fail in importantly substantive ways. Given the way they fail, we are not licensed to infer that human actions are wholly governed by physical laws.

First, it is important to be clear here that I am not arguing that we *know* it to be false that the world is wholly governed by the laws of physics. While I would be happy that if that is the reader's takeaway, I do not need that conclusion for my current purposes. Here, I just want to motivate the

²² There are some authors that have also suggested similar doubts about Pereboom's objection. O'Connor (2003, 309) raises some concerns, but doesn't discuss it in much detail. Vargas' (2013, 68-69) comments also seem to be in the spirit of rejecting this claim, although Vargas ultimately does not endorse agent causation.

thought that the objection fails to undermine agent causation, and to do that, I only need the weaker epistemic claim: that we should withhold judgement about whether the current empirical evidence shows that the world is wholly governed by the laws of physics.

Second, insofar as I am raising doubts about the idea that the world is wholly governed by physical laws one may be inclined to read this as an attack on science writ large, that we should not trust science. Importantly, this is not what I am advocating. Instead, my point is that the science is much more complicated than one is initially inclined to think, and these complications justify a moderate form of skepticism about a *particular* scientific picture: that everything is explained by physics.

With those clarifications out of the way, I would like to fix our discussion on two different claims that are connected to the idea that the world is wholly governed by our physical laws:

Vertical reductionism. All higher-level laws are reducible to our fundamental laws in physics.²³

Horizontal reductionism. Our fundamental laws of physics are universal in scope across all the types of events in their domains.²⁴

Together, vertical and horizontal reductionism form a set of interlocking principles that ensure that everything that happens is dictated by fundamental laws of physics. Horizontal reductionism allows us to say that the fundamental physical laws reach to all areas of the universe in virtue of the fact that everything is composed of the same fundamental stuff. However, we also recognize that there are different “levels” to nature, not only are the physical facts (in the physics sense), but there are also

²³ I take it that a kind of ontological reduction generally goes hand-in-hand with vertical law reduction. When some higher-level law, in say chemistry, is reducible to a lower-level law in physics, it is partly because the chemical phenomena are reducible physical phenomena.

²⁴ Vertical and horizontal reductionism are universal generalizations. Vertical and horizontal reduction are individual instances. The distinction between vertical and horizontal reductionism is adapted from Cartwright 2016.

(e.g.,) sociological, biological, and chemical facts. Vertical reductionism allows to in turn explain all the higher-level phenomena by reducing what goes on at the higher level to what goes on at the fundamental level.

By putting pressure on both of these claims we thereby put pressure on the claim that the world is wholly governed by the physical laws. In light of this, one might wonder whether casting doubt on either of these claims is sufficient, necessary, or jointly necessary and sufficient for the purposes of this argument.

The answer to this question seems to me a complicated one. What's relevant for the success of the Wild Coincidence Objection is whether we're licensed to say that the laws of physics wholly govern human actions. So, to block the objection, what we need is sufficient doubt about whether the laws of physics wholly govern human actions. And this could potentially be blocked by undermining either vertical or horizontal reductionism, so long as the domains in which they failed led to a breakdown in being able to wholly explain human actions using the fundamental physical laws.

What makes the matter more complicated is that the best motivation for thinking that human actions are wholly governed by physical laws is not a direct argument. It is not an argument which involves trying to show that we can use fundamental physical laws to predict specific human actions. Rather, it is something like an indirect and cumulative case that reasons along the following lines: given the success of physics explanations across some relevantly robust set of cases, we can infer that human actions are also likely explained by the laws of physics. Because of the structure of this type of reasoning, the strategy here for blocking this generalization is to present a corresponding cumulative case to the effect that these physics explanations do not rise to the requisite level to infer that everything is explained by physics.

§3.4.3 Horizontal Reductionism

Horizontal reductionism is the claim that all physical laws are universal in scope. Roughly, the idea is that all laws have universal quantifiers attached to them. For instance, to say that Newton's second law ($F = ma$) is universal in scope is to say that Newton's second law holds for all forces, masses, and accelerations.

To deny horizontal reductionism with respect to some particular laws is not to claim that those laws are false. Instead, it is to claim that their scope is limited. Rather than reading laws (like $F = ma$) with a universal quantifier in front of them, we should read the laws as having a *ceteris paribus* clause—an “all things being equal” clause—in front of them. The significance of having such a *ceteris paribus* clause in front of, for instance, Newton's second law, is that you simply cannot assume that $F = ma$ holds for any phenomena that have force, mass, and acceleration. To make the generalization that $F = ma$ explains the behavior of that specific phenomena, you need specific evidence that all things are indeed equal in that particular domain where the law is being applied.

But why doubt horizontal reductionism? Although various authors have provided reasons to doubt this claim (e.g., Cartwright 1999, Giere 1999, Horst 2011), I want to focus on a concern that stems from the predictive accuracy of our laws. Generally speaking, we take ourselves to have good evidence that laws govern phenomena in a given domain when they can be used to accurately model and predict the phenomena in that domain. Predictions in lab-like settings present the clearest case of this where all factors are perfectly controlled for. In these cases, we can make accurate models and predictions about how, for example, a falling object will behave using Newtonian laws. Sometimes we can even do this outside the lab. An especially clear example of this is the motion of planetary bodies. Using Newtonian laws, we can also accurately model and predict how planets will move through space. All of this gives us good reason to believe that Newtonian laws govern those particular phenomena.

A concern arises, however, when we realize that this kind of modelling and prediction fail in cases that appear to be ubiquitous. Consider an example presented by Cartwright. She writes:

In some cases a physicist is a worse prophet than a [behaviourist psychologist], as when he is supposed to specify where in Saint Stephen's Square a thousand dollar bill swept away by the wind will land, whereas a [behaviourist] can specify the result of a conditioning experiment rather accurately. (Cartwright 1999, 27)

Cartwright's contention is that, if $F = ma$ is universal, then we should be able to use it to determine the acceleration of the thousand dollar bill, given the relevant forces and masses, and thus predict its inevitable placement in Saint Stephen's Square. But this cannot be done. $F = ma$ allows us to predict the acceleration, given the forces, and we know one of the forces well—the force of gravity. But this does not allow us to make the correct predictions. And we know why: the wind pushes the bill, too. The problem is that we don't know how, using our knowledge from physics, to write a force for the wind. Neither does physics provide us with a way to model the “inner structure” of the wind, so as to arrive at an accurate value for the force of the wind. Since our physical laws cannot be used to predict what will occur here, we should be reconsider whether we have good empirical evidence that it's our physical laws that explain what is going on. And as Cartwright notes, the example is not simply an anomaly, but rather can be generalized to any number of different laws and scenarios.

Some want to say that the wind exerts a proper force, and that, when that is added, $F = ma$ will allow us to make an accurate prediction. But that is merely what we *say*. It is supposed that we can do it because sometimes we are able to accurately describe all the forces. But, supposing we always (or usually) can is well beyond the empirical evidence, at least so Cartwright argues.

The points generalize to our fundamental physical laws that are said to govern the fundamental constituents of this world. While our fundamental physical laws can make accurate predictions in specific scenarios, this does not necessarily warrant saying that they generalize to explain the behavior of all the fundamental constituents at all times and at all places, especially when we cannot use those laws to predict the behavior of a given phenomenon. Many want to say that our inability to use our laws to predict a given phenomenon is merely an epistemic limitation on our part—and this could very

well be true. But the point is that this shouldn't be assumed, but rather arguments need to be given to substantiate this claim.

§3.4.4 Vertical Reductionism

I turn now to vertical reductionism, or the claim that all higher-level laws are reducible to our fundamental laws in physics. If vertical reductionism is true, then everything that an agent does is wholly explainable in terms of the goings on at the fundamental level. Vertical reductionism, then, puts direct pressure on agent causation insofar as it posits a non-reductionist picture of agency. (For the remainder of this section, I refer to vertical reductionism simply as reductionism.)

It might be thought that reductionism is obviously true on the grounds canvassed earlier: everything as far as we know is composed of the same fundamental physical stuff and this fundamental physical stuff is governed by a set of fundamental laws. So, if an agent is wholly composed of, and governed by, those same physical laws, then what an agent does must be wholly explainable by those same laws.

While I think that this might provide some initial inclination towards reductionism, it's nonetheless a question that ought to be informed by what the empirical data suggest—especially in this context where we're dealing with an empirical objection to agent causation. What I want to suggest is that, when we look at the empirical evidence for reductionism, the evidence does not require that we endorse such a picture, and in many cases seems to incline against it. Of course, this is not to say that nothing is ever reducible to something more basic or fundamental. Far from it. What this does say is that reductionism cannot simply be assumed, and so we need to specific reason to think that it obtains.

Indeed, in recent years there has been a growing number of philosophers of science who have started to doubt reductionism. A large of number of such philosophers can be found in the growing

movements of scientific pluralism and scientific disunity.²⁵ Though their reasons for adopting these views can be diverse, scientific pluralism and scientific disunity are often motivated by the failure of, and opposition to, reductionism (Ludwig and Ruphy 2021).

If one simply takes a bird's eye view of how the sciences operate, this does not obviously suggest a picture where each domain of science is wholly reducible to a more fundamental science, with everything bottoming out at our most fundamental science, quantum mechanics. Psychologists do not generally take themselves to be merely doing complicated biology, biologists complicated chemistry, and chemists complicated physics. The way that sciences operate is typically much more complicated. Scientists are constantly borrowing from, and relying on, data across a wide range of sciences, but it's not typically done in a way that straightforwardly suggests a reductionist picture.

The case of chemistry is, I think, illustrative of how the sciences work as a whole. Chemistry suggests itself as the domain of science most ripe for reduction. Chemistry is the study of molecules, molecules are composed of particles, and quantum mechanics studies the behavior of these particles.

In practice, there doesn't appear to be a neat reduction of chemistry to quantum mechanics. There are various reasons for skepticism about reduction here. Lombardi suggests, for example, that there is a strong discontinuity between various core concepts and terms in chemistry and those in physics (Lombardi 2015). Another reason, one that I'm inclined to find more compelling, is that the reduction appears is not only incomplete, but not a principled reduction to quantum mechanics.

The attempt to reduce chemistry to quantum mechanics revolves around the ability to predict the structure of the various molecules in chemistry using special versions of the Schrödinger equation. Here, there are a number of limitations that arise which others have noted (Chang 2015, Hendry 2010,

²⁵ Scientific pluralism is “an umbrella concept for philosophical engagement with conceptual, methodological, theoretical, and social diversity in science” (Ludwig and Ruphy 2021). Scientific disunity is a related view that “argues for the replacement of the emphasis on global unity—including unity of method—by emphasizing disunity and epistemological and ontological pluralism” (Cat 2017).

and Lombardi 2013) First, the reduction only works for simple molecules like hydrogen. And second, even in the case of simple molecules, the reduction is not a pure reduction.

The reduction is not pure for various reasons. Here are a just a few. One reason is that the Schrödinger equation has assumptions that conflict with quantum mechanics. In order to predict or explain the location of a nucleus in a molecule, the nucleus must be treated as a classical particle, with a fixed location and momenta. This is said to conflict with the Heisenberg uncertainty principle. In addition to this, the location of the nucleus is already taken from “pre-quantum” chemistry, which means that the location of the nucleus is not even independently predicted by quantum mechanics. The purported reduction of chemistry to quantum mechanics is entirely infused with non-quantum theory.

A second point to highlight is that the domain of physics itself does not exemplify a unified kind of reductionism. To be sure, there are plausible cases of reduction, such as the reduction of thermodynamics to statistical mechanics, but this seems to be exception rather than the norm. Most branches of physics do not reduce to more fundamental branches of physics, and sometimes even conflict, such as the case of relativity theory and quantum mechanics. Chang (2015, 200) puts the point the point nicely here:

In the here-and-now of physics, it is important to note that many important and successful uses of physics draw from various parts and levels of physics, which have not been unified with each other in general. A good example illustrating that point is the global positioning system (GPS), which draws from Newtonian mechanics (for satellites), quantum mechanics (for atomic clocks), and special and general relativity (for the correction of atomic clocks), without any attempt to unify those theories on the whole.

Reductionism is not something can just be assumed about the higher-level sciences, but specific empirical evidence is required to substantiate it.

What does all of this show? Two things. First, Pereboom's claim that the world is wholly governed by physical laws is a much more substantive empirical claim than one might initially think. Second, and while this empirical evidence is by no means comprehensive, it suggests that the claim that the world is wholly governed by physical laws is not obviously forced upon us by the empirical evidence. It is reasonable for one to doubt the claim in an empirically respectful manner.

Once this assumption is no longer granted, the Wild Coincidence simply fails to get off the ground. The objection requires that we be able to offer explanations from fundamental physics about everything, especially human actions. This in turn generates a dilemma for agent-causalists. Without the assumption, no dilemma can be generated. And with no dilemma generated, no pressure is put on the agent-causalist.

§3.5 Three Replies

In recent years, a number of different replies have been put forth to Pereboom's empirical challenge. I turn, then, to examining three extant replies to Pereboom's Wild Coincidence Objection. One feature that these replies have in common is that they grant (or at least do not pressure on) the assumption that the world is wholly governed by physical laws. Instead, they appear to offer ways to different ways of undermining the objection without calling into question Pereboom's empirical claim. Addressing these replies will allow us to draw out the why the objection is forceful once Pereboom's assumption is granted.

§3.5.1 Baker's Reply

Let's start first with a line of reply initially proposed by Clarke (2003), but further developed and defended by Baker (2017).²⁶ The response relies on a modified conception of agent causation. On Clarke's agent-causal *integrationist* account, a free action is caused both by the agent *and* a corresponding event involving the agent's reasons. Furthermore, it is said to be a *law of nature* that, when an agent causes an action, a corresponding event also causes the action. Agents are "integrated" into the laws of nature with events.

As Baker notes,

The integrated approach circumvents Pereboom's argument by showing that agent-causes are already *built into* the microphysical laws. Thus, there is no space to consider how the action of an agent-cause *just so happens* to follow the physical probabilities because the agent-cause is *already* described when considering the probabilities.
(Baker 2017, 3093)

If the integrationist line of reply is to succeed, then recall that it must do two things: (i) address the Wild Coincidence Objection by providing a principled explanation for why we should expect conformity to the physical laws and (ii) show that this principled explanation is at least as empirically plausible as the explanation relying solely on our physical laws.

With respect to addressing the Wild Coincidence Objection, integrationism seems to fare quite well. Integrationism has a straightforward explanation for why purportedly free agent-caused actions would conform to our physical laws—the integrationist simply builds into the physical laws agent-causes. Suppose that our physical laws tell us to expect, more or less, that Ellie will raise her hand three times over the course of five occasions and that Ellie in fact raises her hand three times over the course of five occasions. Since Ellie's agent-causing her hand to raise will be part of the physical laws

²⁶ There may be at least one potential difference. While Baker suggests that he is defending a strong-emergentist version of agent causation, Clarke (2003, 177–181) does not seem to posit this.

themselves, conformity to the physical laws is expected—indeed, conformity is guaranteed on the integrationist account.

Agent-causal integrationism provides an attractive way of addressing the Wild Coincidence Objection. However, the response seems to lose plausibility when trying to address the second challenge. In order to address this challenge, the agent-causalist must show that her explanation for conformity to the physical laws is just as empirically plausible as an explanation relying solely on our physical laws; but, the integrationist explanation seems to be *less* empirically plausible than the explanation relying solely on the physical laws.

Call the integrationist explanation for why there is conformity to the laws *E1*. According *E1*, physical laws are in fact explaining what is occurring, so conformity to the physical laws obviously holds. However, for the integrationist, the physical laws have two components: (a) events and (b) agent-causes. When a purportedly free action is caused, it is caused both by an event and an agent. Now, call the explanation that relies on non-integrationist physical laws *E2*. These physical laws only have physical events as their components—they do not contain agent-causes. The question arises: which is more empirically plausible? *E1* which contains agent-causes in contents of the physical laws or *E2* which does not contain agent-causes as contents in the physical laws?

Consider once again what is supposed to give rise to the more general dilemma within which our discussion is embedded. Earlier we saw Pereboom's claims that "on our *best physical theories* the physical world is law-governed" (Pereboom 2014: 66, emphasis added); and that, if agent causation were to hold, "such agent-causes would be embedded in a world that, *by the evidence that supports our current theories in physics*, is nevertheless wholly governed by the laws of physics" (Pereboom 2001: 79, emphasis added). While Pereboom doesn't explicitly mention which physical laws (theories) he has in mind, examples aren't hard to come by. It's plausible that what Pereboom has in mind are cases such as Schrödinger's equation. And so, if it is physical laws like these that are thought to wholly

govern the world, it is rather clear that our current physical laws don't countenance anything like agent-causes.

Since the integrationist adds agent-causes into the contents of our physical laws, they are making an addition which is not currently supported by empirical evidence. *E1*, then, seems to be empirically suspect. In contrast, *E2* simply relies on our physical laws in their current form and thus doesn't raise any empirical concerns. It seems clear that *E2* is more empirically plausible than the agent-causal integrationist's *E1*.

§3.5.2 Taggart's Reply

Taggart (2020) has recently put forth a line of reply wherein the Wild Coincidence Objection is said to fail even by its own lights. He believes that the Wild Coincidence Objection fails *even if* we grant (a) that we have very good empirical evidence that all events are governed by probabilistic physical laws, (b) that an agent's directly causing an action is not governed by (probabilistic) physical laws, and (c) that agents can, but never do, deviate from the probabilistic physical laws (Taggart 2020: 8).

Taggart's criticism seems to hinge on what he takes to be the crucial assumption of the Wild Coincidence Objection:

11. Epistemically, the probability that both all events are governed by physical laws and that agent-causal libertarianism is true is extremely small.

The problem with this claim, Taggart argues, is there is insufficient reason to believe that it is true. This is because the claim "is not an empirical premise. Any reason we have to believe [this claim] is independent of any empirical evidence" for thinking that all events are governed by physical laws (Taggart 2020, 8). The idea appears to be that, simply knowing that we have empirical evidence that all events are governed by physical laws does not tell us the relative likelihood of it both being the case

that all events are governed by physical laws and that agent causation is true. In short, Taggart thinks that an agent-causalist need not provide a principled explanation for why we should expect free agent-caused actions to conform to our physical laws. The Wild Coincidence doesn't get off the ground to begin with.

It's not entirely clear to me why Taggart thinks that (11) is implausible. Taggart here seems to be assuming that the world can *both* be wholly governed by physical laws and that agents can freely agent-cause actions. But this seems to be false. If the world is wholly governed by physical laws—such as Schrödinger's equation—then it is these laws that explain the actual frequencies that we see. And since these laws exclude agent-causes, an agent-causal explanation for any purportedly free actions occurs will be inconsistent with the thesis that our physical laws wholly govern the world.

§3.5.3 Runyan's Reply

Now let's turn to our third and final reply. Runyan (2018) has advanced a response that critiques the Wild Coincidence Objection on empirical grounds by providing neuroscientific evidence that the neural laws involved in human behavior have changing probabilities. More specifically, Runyan uses this empirical evidence to advance a two-part reply for why conformity to the laws would not be a coincidence:

First, inclining factors may wax and wane depending on what [an agent such as] Ellie does. This, at times, sets parameters on how frequently she will exercise her agency in certain ways across a set of occasions. Second, Ellie either exercises her agency in a way that shifts expectations going forward (when these expectations change), or in a way that conforms with expectations and doesn't change them (when they don't change)...It is this combination of restrictions and influence which might very well

explain why, *in every case*, her behavior conforms with expectation. (Runyan 2018, 4573)

In short, it is both inclining factors and changing probabilities that are meant to address the Wild Coincidence Objection.

Given that agents can influence the probability that later actions will occur, we can assume there will usually be no long-term frequencies for which we should expect human actions. Still, we can assume that there will usually be short-term (and constantly changing) frequencies for which we should expect human actions. How, then, do the two factors—changing probabilities and inclining factors—explain why it is reasonable to expect conformity to the short-term frequencies? It is because, on the one hand, inclining factors (e.g., mental events such as desires and urges) “[set] parameters on how frequently [an agent] will exercise her agency,” while, on the other hand, the agent can still exercise her agency to shift (or not shift) the probabilities going forward.

To build on a case from Runyan (2018, 4571-4572), suppose there is, at the outset, a .32 probability that Ellie will raise her hand across 100 occasions. Further suppose that Ellie conforms to expectations for the first five occasions, but the way in which she exercises her agency in those five occasions shifts the probability for the rest of the 95 cases to, say, .7. Ellie then conforms to the new probability of .7 for the next five cases, but then once more shifts the probabilities going forward (for the remaining 90 occasions), and so on until the 100 occasions have passed. Ellie meets expectations in each of the short-term cases because of the restrictions placed on her by the inclining factors, but nonetheless has an influence (qua agent-cause) on the probabilities along the way.

While Pereboom initially frames the Wild Coincidence Objection in terms of long-run frequencies, it's not entirely clear to me that the objection needs to be framed in this way. What seems to be doing the real work is that there is conformity to the physical laws and that this conformity is need of an explanation. Indeed, Runyan throughout his discussion reaffirms that his response is meant to explain why it might be that agent-caused “behavior always conforms with expectations.” It is just

the expectations involve fluctuating short-term frequencies. With that said, recall that, if Runyan's response is to succeed, two things need to be met: the response must provide (i) a principled explanation for why conformity to the *constantly changing short-term* frequencies predicted by our laws is reasonable to expect and (ii) this principled explanation must be at least as empirically plausible as an explanation relying solely on our physical laws.

Let's grant that Runyan's two-fold response provides a principled explanation for why agents would conform to the short-term and changing frequencies predicted by our laws. Even if this is granted, it's not entirely clear that this explanation is at least as empirically plausible as an explanation relying solely on our physical laws.

Call Runyan's two-fold explanation *E1*. *E1* is meant to explain why conformity to the laws is reasonable to expect. If there is conformity to the short-term and fluctuating frequencies predicted by our physical laws, how might a proponent of the Wild Coincidence Objection explain this? Again, presumably one could argue that it's our physical laws that are explaining the conformity—laws that don't make reference to anything like agent-causes. Call this explanation *E2*. Given everything we've discussed so far, *E2* appears to be more empirically plausible than *E1*. Since we are already conceding at this point in the argument that we have good empirical evidence that the world is wholly governed by physical laws, and that these laws don't include agent-causes, this gives us good reason to prefer *E2* to *E1*. Is there anything that can be said on behalf of Runyan's reply?

The most obvious response, it seems to me, would be to show that agent causation is needed to explain why the probabilities of future actions change, or at least that we have good empirical evidence that the changing probabilities are attributable to agent causation. But it's not clear that Runyan has shown this to be the case, he's only shown that it's in principle possible. Indeed, much of the evidence Runyan provides to substantiate the claim that there are fluctuating frequencies seems to suggest that shifting probabilities, from one decision to the next, is a general phenomenon regarding action more broadly and is not limited to human action in particular (e.g., Klaes et al. 2012; Barker et al. 2014;

Friedman et al. 2015; Bellay et al. 2015). If there are fluctuating short-term frequencies even for non-human animals, this strongly suggests that the fluctuating frequencies can be explained in terms of physical laws without recourse to agent-causes. If so, we have good empirical reason to think that we don't need agent causation to explain updating frequencies, and conformity to these frequencies, even in the case of human action.

§3.6 Concluding Remarks

What should an agent-causalist make of this proposed line of reply? At first glance, calling into question that the physical laws wholly govern the world may seem to be a large price to pay to save agent causation. However, it's not so clear to me that things are so bad. In order to avoid Pereboom's empirical objection, one need not necessarily reject the claim the world is wholly governed by physical laws. What agent causation is incompatible with is that our world is wholly governed by fundamental physical laws like Schrödinger's equation, laws that rely on a reductionist picture of the world. This is in principle compatible with the idea that there are physical laws that somehow accommodate the agent-causal picture of free will, similar to the integrationist proposal.

Another possibility—one that is consistent with the integrationist proposal, but still distinct—is to draw a distinction between two senses of laws of nature. A neo-Aristotelian theory of laws says that they are descriptions, or generalizations, of how causal powers are manifested in the world. In a weak sense, everything is law-governed because everything that happens is the outcome of interacting causal powers. But we may want to reserve a more technical and robust notion of laws of nature for those network of interacting causal powers that are describable in a suitably rigorous way (cf. Cartwright 1999 and Demarest 2017). Furthermore, among some of those causal powers may be those had and exercised by agents. These, of course, are just sketches, but they nonetheless present themselves as live possibilities.

Of course, other lines of response may be open to the agent-causalist to develop. For instance, perhaps the agent-causalist could defend the second horn of the Wild Coincidence Objection—that we should at some point expect deviations from our physical laws—on *non-empirical* grounds, such as on the basis of philosophical arguments or non-evidential grounds (cf. Speak 2004). While Pereboom suggests that it would be infelicitous to form scientific beliefs about the empirical world on these bases, this does not seem to me to be beyond dispute.

Of course, there may yet be other ways of addressing Pereboom’s empirical dilemma. While I’m optimistic about the empirical case against Pereboom’s claim that the world is wholly governed by fundamental physical laws, I hope to have also shown that Pereboom’s argument is more forceful than previously realized, and that the objection highlights why it is important for the agent-causalist to get clearer about agent causation’s place in nature.

Chapter 3, in part, is currently being prepared for submission for publication of the material (as “Agents, Laws, and Events: Reconsidering Pereboom’s Wild Coincidence Objection,” by Martinez, Joseph). The dissertation author was sole author of this paper.

Chapter 4

Agent Causation and Motivating Reasons

§4.1 Introduction

Let us take stock. The previous two chapters have focused on defending agent causation from two different objections. I have tried to make the case that agent causation is neither too metaphysically demanding nor in conflict with our best picture of science. Having addressed those concerns, I want to shift our focus to a different matter, namely, trying to develop the view so that it can account for central aspects of our agency and agency-related practices. In particular, I want to examine two closely related issues: acting on reasons and being morally responsible. While neither of these topics have been entirely neglected by agent-causalists, they have not occupied much of their attention either. Important aspects of these issues remain unaddressed and so are in need of further clarification. If agent causation is going to make any more headway, it's imperative that the matters receive further treatment.

This chapter is an attempt to take up one such task of furthering the agent-causalist project. In particular, I focus on the problem of integrating motivating reasons²⁷ with agent causation.²⁸ When we act, we often do so *for* certain reasons. For instance, when Malia goes to the store, she does so *for the purpose of* getting milk. And, when Pablo drinks coffee, he does so *because* he needs to wake up. In each of these cases, we are citing reasons that motivated, or explained, the action in question. That we take ourselves to be able to provide these kinds of rationalizing explanations of our behavior is, I take it, a very familiar and uncontroversial idea. That agent causation should be able to accommodate this into its picture, then, is an obvious desideratum.

²⁷ Sometimes motivating reasons are also referred to as explanatory reasons. More on this below.

²⁸ Going forward, "reasons" will refer to motivating reasons, unless stated otherwise.

Here is how the rest of the chapter will proceed. In the next section, I begin by outlining the agent-causal theory of action and contrasting it with the event-causalist picture of action. I then consider, each in turn, five different ways of articulating the relationship between agent causation and reasons for action—three causal accounts, one teleological account, and a less discussed account, which I call the *causal powers* view. Along the way, I lay out what I think the prospects for each are. I argue that none are entirely satisfactory, and thus conclude that the most plausible option for the agent-causal theorist is to take the relation of acting on a reason to be a primitive and basic feature in her overall account.

§4.2 Two Theories of Action

This section takes up the following question: what *makes* some event an action? As we shall see later on (see §2.3), not all agent-causalists will agree on how to answer this question. This is because some agent-causalists posit agent causation only in order to account for *free* action, while others posit agent causation to account for action *as such*. While I remain motivated to develop an agent-causal theory of free will, I believe that the most plausible form of an agent-causal theory of free will is one that embraces the wider scope application of the view—namely, one that aims to account for action as such. What we should want to know, then, is what a unified picture of agent causation looks like. The best way to do this, it seems to me, is to start by developing an agent-causal theory of action, and subsequently extending it to an agent-causal theory of free will. So, in the rest of this section, I outline the core the agent-causalist account of action, highlighting some of the core commitments along the way. I then contrast the agent-causal account of action with the orthodox view of action, i.e., the event-causalist view. The next section takes up the central question of the chapter: how can an agent-causal theory of action make sense of acting *for a reason*?

Let's start by considering some examples of actions: *Talia's thinking through a math problem*, *Diego's reaching for a craft beverage*, and *Isabel's raising her arm*. With such examples in mind,

what kinds of things can we conclude about the nature of action? Here are two things. First, actions seem to be *events*. That is to say, they are things that *happen* or *occur*. Second, they're not just any sorts of events, but they're events that crucially involve the agent. For ease of reference, let's call such events *movements*.

Not all movements are actions. Consider, for instance, a case where the movement of *Isabel's arm going up* is brought about by a sudden muscle spasm. While clearly a movement of Isabel, it is not an action that she performed.

It was not an action because Isabel did not exhibit the right kind of *control* over the movement. Evidently, we exhibit control over things (at least partly) in virtue of being able to cause various effects. For instance, I can control the direction in which my car moves *by* causing the steering wheel to move in various ways; and, I can control the car's speed *by* causing the gas pedal to go down. The same, it would seem, for our actions:

Causal Condition: A movement *e* is an action if only if an agent *S* stands in a suitable causal relation to *e*.

The agent-causalist maintains that agents exhibit control over their movements by *directly* causing those movements. More specifically, the account looks like this:

Agent-Causal Theory of Action: A movement *e* is an action if and only if *S* (qua substance) directly caused *e*.

Consider a simple example. Suppose that Isabel performs the action of raising her arm. In this case, we have: the agent (Isabel), a movement (*the raising of her arm*), and the two standing in causal relation to one another. More specifically, on the agent-causalist picture, Isabel directly (literally) causes the

movement of *the raising of her arm*. It is in virtue of this causal relation that Isabel exhibits control over this movement, and, it is in virtue of this causal relation that the movement is an action.^{29, 30}

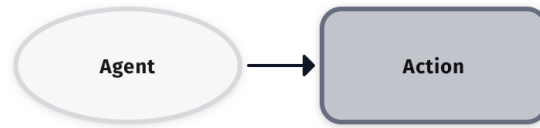


Figure 4.1: Agent-Causal Theory of Action

Philosophers of action typically distinguish between two kinds of actions: basic and non-basic. Basic actions are actions that are performed without *doing* anything else. In contrast, non-basic actions are actions that are performed by performing basic actions. For example, raising my arm is considered by many to be a basic action. I don't raise my arm by doing anything else; I simply raise my arm.³¹ In contrast, climbing a ladder is a non-basic action; I climb the ladder by moving my arms and my legs. Here, we are primarily focused on basic actions.

While most philosophers of action agree that there is a genuine divide between basic and non-basic actions, they disagree about where basic actions are to be located. On the one hand, we have those who maintain that basic actions are *mental* actions, such as the formation of an intention. On the other hand, we have those who maintain that basic actions are *bodily* movements, such as moving my leg. While I have a preference for mental actions, for the purposes of this chapter, I leave it open.

²⁹ Here I work with a product view of actions. Actions are events that are distinct events that are caused by the agent. Some agent-causalists espouse a component view of action, where actions are complex events that have as their constituents the agent and some further event. So, while the product view says that *the raising of her arm* is an event (in virtue of being caused by Isabel), the component view would say that the action is *Isabel's causing the raising of her arm*. Actions, on this latter view, are agent-causings of events.

³⁰ For clarification: circular nodes represent substances (more on this below) and rectangular nodes represent events. Additionally, shaded nodes pick out actions and arrows represent causal relations.

³¹ Of course, this is not to say that nothing else happens. When I raise my arm, certain neurons fire, for example. However, I don't raise my arm *by* firing any neurons. Indeed, I don't raise my arm *by* doing anything else. I simply raise my arm. For a helpful overview on basic action, see Amaya (2017). For a critical discussion on how we should think of non-basic actions, see Kelley (2024).

However, for the ease of explication, I will continue to use bodily movements as examples of basic actions, except in cases where an account explicitly endorses a mentalist account of basic actions.

§4.2.1 Event-Causalism

The event-causal theories of action enjoy the status of orthodoxy within the literature. These theories deny that actions are directly caused by the agent. Rather, the agent's causal contribution to an action is exhausted by the causal role of the agent's mental events—where the mental events in question are typically taken to be the agent's reasons. Reasons, on the standard view, are reducible to a belief-desire pair.³² For instance, if Isabel has a desire to cast a vote, and Isabel also believes that by raising her arm she will cast a vote, then Isabel would have a reason to raise her arm on this view.

What makes some movement an action on the event-causalist view, then, is that it is caused by the agent's reasons. If Isabel's reasons caused *the raising of her arm*, then it is in virtue of this that that movement is an action:

Event-Causal Theory of Action: A movement e_2 is an action just in case it is (non-deviantly) caused by an appropriate mental event e_1 involving S (e.g., the agent's reasons). (cf. Franklin 2018: 12)

³² Of course, not all will be inclined to accept this psychological reductionist picture of reasons. Some think that reasons are not mental events, but are facts. Still others think that reasons should be taken as a primitive—they're simply considerations that count in favor of a course of action. The event-causal theorist of action can accommodate these alternative accounts of reasons by stipulating that there is a mental event of *an agent's taking something to be a reason*, and that this event can play the same role as a belief-desire pair in this account of action. See Franklin 2018: 12, fn 3.

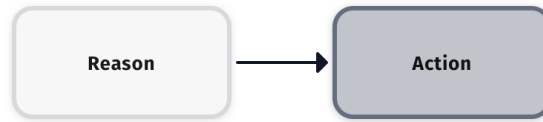


Figure 4.2: Event-Causal Theory of Action

§4.2.2 Action and Free Will

By now, readers may have noticed that agent causation can be put to use for two different, though overlapping, purposes in action theory: (a) in order to provide an account of action *as such* (cf., Bishop 1983), and (b) in order to provide an account of *free* action (cf., O'Connor 2000 and Clarke 2003). While I remain motivated by concerns for free will, I believe that the most plausible version of an agent-causal theory of free will embrace both (a) and (b). This seems to me preferable insofar as it makes the agent-causal position more parsimonious. Hence, going forward, in weighing the different available options, I will keep in mind both kinds of considerations. With that said, some clarificatory remarks are in order.

Scope. The agent-causal theory of action and agent-causal theory of free will differ in scope. The action theorist is committed to saying that all actions must involve agent causation, whereas the free will theorist is only committed to saying that free actions involve agent causation. For instance, Clarke (2003) explicitly maintains an event-causal theory of action alongside an agent-causal theory of free will. So, on this combination of views, all actions are caused by reasons, but free actions are also co-caused by the agent. If one starts with an agent-causal theory of action, one is committed to saying that both free and unfree actions involve agent causation. So, while the agent-causal theory of action entails an agent-causal theory of free will (assuming one believe that free will exists), the agent-causal theory of free will does not entail an agent-causal theory of action.

(In)Compatibilism. As mentioned in chapter 2 , there are both libertarian and compatibilist versions of agent causation. By extension, I'm inclined to think that an agent-causal theory of action is also in principle consistent with our being causally determined. While I remain committed to agent causation in its libertarian form, I plan to stay neutral about the matter in our discussion here, though I make it clear in places where I think this commitment to libertarianism is relevant. Apart from the fact this broadens the reach of the discussion to agent-causalists who are compatibilists, there are also relevant implications for the agent-causal libertarians like myself.

The thought is that, if we are going to say that all actions are agent-caused, we should want to allow for the possibility that some of these actions are unfree. And among these unfree actions, it is likely that some of these actions may be the result of being causally necessitated, such as in the case of an irresistible urge or desire. Plausibly an agent whose action results from an irresistible urge is still just that, an action. If so, this will only possible we allow that, sometimes, an agent can be the direct cause of her action while being causally necessitated to do so.

§4.3 Causalist Accounts of Motivating Reasons

In the previous section, we saw how the agent-causalist accounts for action in terms of the direct causal involvement of the agent. The rest of the chapter takes up the following question: given an agent-causal theory of action, how can we accommodate motivating reasons into the picture? We will focus our discussion on three different families of views of motivating reasons: (a) causalist accounts, (b) teleological accounts, and (c) powers accounts. The remainder of this section focuses on causalist accounts.

Agents typically act for a specific goal or purpose. This is to say that, ordinarily, they *act on* (or *for*) *reasons* that they have. For instance, Diego might drink a craft beverage *in order to* satisfy his thirst, and Isabel might raise her arm *for the purpose of* casting a vote. When an agent acts on (or for) a particular reason, call that reason a *motivating reason*.

Philosophers sometimes distinguish between two other kinds of reasons: normative and explanatory. Normative reasons are reasons for an agent to act in a particular way. They are, as Scanlon puts it, considerations that count in favor of a course of action. For instance, Isidora may recognize a reason to not raise her arm, but nonetheless not act on that reason. It is a normative reason, but not a motivating one.

Explanatory reasons are said to be reasons that explain an action, without necessarily being reasons that motivate the action. To adapt an example from Alvarez (2017), suppose that Isabel knows that a candidate has committed some infelicitous act. *The fact that* the candidate has committed an infelicitous act may be a reason that explains why Isabel voted against him, but this is not what motivates Isabel. What motivates Isabel to vote against the candidate are her mental states (or events), e.g., her belief that voting against him will get him out of office and her desire to vote him out because he committed an infelicitous act. The view that explanatory reasons are distinct from motivating reasons is controversial. While I don't distinguish between the two, readers should feel free to substitute motivating reasons in wherever I make reference to explanatory reasons.

The causalist view of motivating reasons can be put as follows:

Event-Causal Theory of Motivating Reasons: An action A is done for a reason R just in case A is (non-deviantly) caused by R .

Readers may notice that the event-causalist view of action and the event-causalist account of motivating reasons supply the same answer to two different questions. The former posits a causal connection between the action and the agent's reasons to give an account of action, while the latter posits the same causal connection to give an account of motivating reasons. Thus, most theorists who have adopted one have adopted the other. Nonetheless, the two can in principle be prized apart—and that is what we shall do here.

It is difficult to overstate the influence of the foregoing picture of action. It would not be an exaggeration to say that, within both the free will and action theory literatures, the vast majority of philosophers have simply assumed the view; and, for those developing alternative views, they have had to recognize that the burden is on them to show why one should reject the event-causal theory.

But why has the view (especially with respect to the account of motivating reasons) had a hold on the literature? The answer is relatively straightforward: it's difficult to see how else a reason could explain an action, especially in light of the fact that we typically take causal explanations to be the paradigm of explanatory relations in nature. Why does the moon affect the tides? Because the moon exerts a gravitational pull on the earth. Why did Malia's blood pressure drop? Because she was pregnant. So too, it would seem, for motivating reasons: a reason explains an action by standing in causal relation to that action. What other relation could a reason stand in to an action such that it would explain that action? It is difficult to see what other options there could be.

The following three proposals, then, take this challenge seriously. That is, they offer three different ways of maintaining that a reason stands in a causal relation to an action, while doing so within an agent-causal picture of action.

§4.3.1 Motivating Reasons as Co-Causes

The agent-causal account of action maintains that a movement is an action when it is caused by the agent. The causalist account of motivating reasons maintains that an action is done for a reason when it is caused by that reason. So, one way to integrate the two is to simply postulate two causes of an action: the agent and her reason. The agent's causing the action explains why that movement is an action, and the reason's causing the action explains what motivated that action.

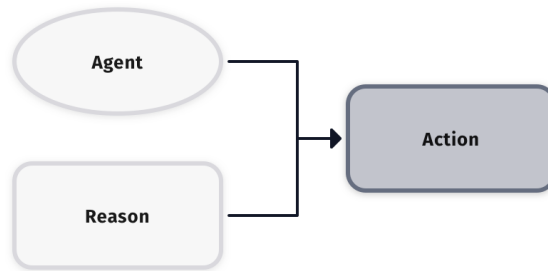


Figure 4.3: Motivating Reasons as Co-Causes

This proposal takes its cue from the agent-causal account of free will developed by Randolph Clarke (1993, 2003)—and more recently endorsed by Franklin (2019)—though it’s important to highlight that it differs in some important ways. Perhaps the most important difference is that Clarke espouses an event-causal view of action (cf. Rice). So, for Clarke, a movement is an action in virtue of its being caused by the agent’s reasons, and an action is motivated by whatever reason it is caused by. Unlike the proposal made here, the agent’s causal contribution does *not* make some movement an action—it is only meant to account for an enhanced degree of control. This enhanced degree of control is in turn meant to explain how an action could be free.

Putting Clarke’s view to the side, what should we say about our current proposal? I think the view faces two primary concerns.

Suppose Isabel causes the raising of her arm, and further suppose that it is also caused by one of Isabel’s reasons. On our current proposal, this would suffice to explain that event’s being an action, and its being done for a reason. If we ask why Isabel raised her arm, we can cite the relevant reason that causally contributed to it.

Adapting an objection originally directed at Clarke (Rice 2011), one might argue that this answer will not do—the picture leaves an important question unanswered. Suppose, instead of asking why Isabel raised her hand, we asked: why did Isabel agent-cause the raising of her hand? Whereas we can cite the reason as a cause of the action, we can’t cite it as a cause of the agent’s agent-causing, as there

is no causal relation between the two. Only by answering both questions can we have a satisfactory account of motivating reasons, or so the thought seems to go.

I'm uncertain to what extent this objecting undermines the view. On the one hand, there is something unsatisfying about not being able to cite a motivating reason as an explanation for why that particular agent-causing occurred. On the other hand, explanations must bottom out somewhere. And, at any rate, the current proposal *does* provide an explanation for why Isabel performed the action that she did (i.e., why she raised her arm). It was because of the reason that caused it. Since the action can be given this kind of rational explanation, it remains to be seen what is insufficient about the view.

Nonetheless, I believe the concern is getting at something important. The deeper concern can be drawn out, I suggest, with the following case:

Donation. Juan and Isabel are out on a date. On the way out, they both see someone soliciting donations for UNICEF. Upon seeing this, Juan finds himself with two reasons to make a donation: (a) donating in order to impress his date, and (b) donating in order to alleviate some suffering. In the end, Juan makes a donation on the basis of reason (b).

Suppose Juan's acting on this motive constitutes a morally good action. Now consider whether we would hold Juan morally responsible for his action (i.e., whether we would praise him for his action). Intuitively, Juan should be praised, since he did something good. But, with our current proposal in mind, I think this intuition begins to fade when we consider the fact that Juan had no *control* over which reason he would act on. That is to say, he had no control over *which* reason would causally contribute to his making the donation. To be sure, Juan had control over whether he would make a donation. It is just that he had no control over which reason would be efficacious. By extension, Juan would seem to have no (or at least a greatly diminished sense of) control over whether he would be praiseworthy or blameworthy for his action.

Putting to the side the matter of free actions, the view seems to fare well enough. At face value, it is able to account for both why some movement is an action, as well as why that action is done for a reason.

Turning now to whether the view can be extended to free actions, I believe that agent-causalists are likely to be of two minds about this conclusion, depending on their view on freedom. On the one hand, if one is libertarian (like I myself am), I believe the view will be troubling. Libertarians place a great deal of emphasis on the agent's being the causal source of their free actions. If it turns out that, in at least some circumstances, it's out of the agent's control whether they are blameworthy or praiseworthy, this cuts against the spirit of libertarianism. On the other hand, the concern about control seems to me less weighty if one is compatibilist. Compatibilists, insofar as they allow for us to be fully causally determined, have comparatively weaker criteria for sourcehood. The concern that some aspect of moral responsibility turns out to be outside of the agent's control would, it seems to me, not have the same kind of pull—though, of course, intuitions may vary.

In sum, the proposal of this section—that motivating reasons are co-causes with the agent—seems to capture non-free actions well enough. However, when it comes to free and responsible action, the view has different results. Compatibilists may rest content with the view, but libertarians are likely to be unsatisfied.

§4.3.2 Motivating Reasons as Prior Causes

Consider, now, an alternative proposal that places the motivating reason in a different location. Perhaps, instead of saying that reasons act as co-causes, we can say that reasons act as prior causes of the agent's bringing about their action. When, say, Isabel raises her arm for a reason *R*, this is to be understood in terms of *R*'s causing Isabel's agent-causing of *the raising of her arm*.

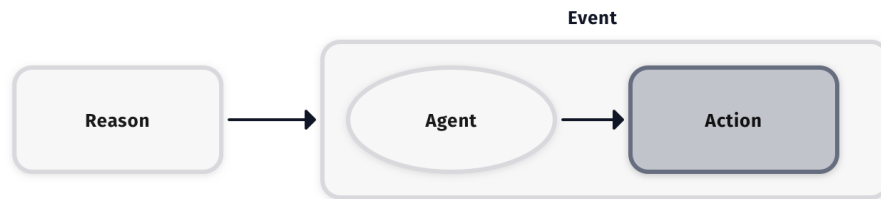


Figure 4.4: Motivating Reasons as Prior Causes

The formulation might seem needlessly convoluted. Why say that a reason causes *an agent's agent-causing*, as opposed to simply saying that the reason caused the agent? The formulation, puzzling as it sounds, is deliberate. The reason is that most agent-causalists object to the idea of substances being effects (cf. Clarke 2003). Most find it infelicitous to say, for example, that a reason caused Isabel full-stop. A reason might cause Isabel *to do something*, or *to be in a certain state*, but those would be events, not substances. Substances, if they stand in a causal relation to anything, can only play the role of causes—or at least so it is widely thought.

The formulation outlined above is meant to get around this concern. The agent's agent-causing of an action is a complex event. It is an event that has as its constituents an agent and an event, with the former causing the latter. Because the reason is said to cause the complex event, and not the agent directly, the substance is never directly caused.

This proposal can get off the ground only if you think at least one of the following is true: (i) that substances can be effects, or (ii) that something can cause a complex event *without* causing the first causal relata in that complex event (cf. O'Connor 2000).

Claim (ii) is doubtful. Suppose that a strike of lightning causes a bush to catch fire, which in turn causes a nearby house to burn down. Can we say that *the lightning's striking* caused *the bush's catching fire causing the burning of the house*? I believe we can, but only if this taken to mean that *the lightning's striking* caused *the bush's catching fire*, which in turn caused *the burning of the house* (cf. Clarke 2000). Short of that, the claim seems implausible. The same, we should say, about the agent-

causal event. We can say that a reason caused the agent's agent-causing an action only if the agent herself can be caused.

Christopher Franklin (2016) has objected that this response fails because the burden of proof lies with the opponent to show that it's *metaphysically impossible* for an agent-causal event to be caused. Franklin maintains that, at most, it's only been shown that substance causation cannot neatly fit neatly into our event models of causation. In fact, Franklin suggests that we should *expect* that some of our assumptions about causation will need to revision since we're allowing in a new kind of causal relata (i.e., substances).

Though helpful, it's unclear to me how satisfactory Franklin's reply is. While it's certainly true that it has not been shown to be metaphysically impossible for a substance to be a causal effect, it's also true that none of our current causal frameworks can make sense of an agent-causal event having a cause. Perhaps there is some to-be-discovered framework that makes sense of this, but it's not clear how much weight this should be given. What matters is whether, with the evidence that we now have, we can make sense of an agent-causal event having a cause.

A more promising way to cash out the idea of an agent-causal event having a cause is the proposal discussed in chapter 2. There, I argued that agent causation can be translated into an event-causal framework without losing anything significant with respect to free will. On this view, an agent-causal event can be wholly formulated in terms of event causation, where the first event is roughly: *the agent's exercising her agent-causal power at a time*. Assuming that agent causation can be so translated in this context, the present concern is dissolved. *The agent's exercising her agent-causal power at a time* is now an event. And because there's no problem with events having causes, there's no reason to think that this complex event can't be caused by a reason.

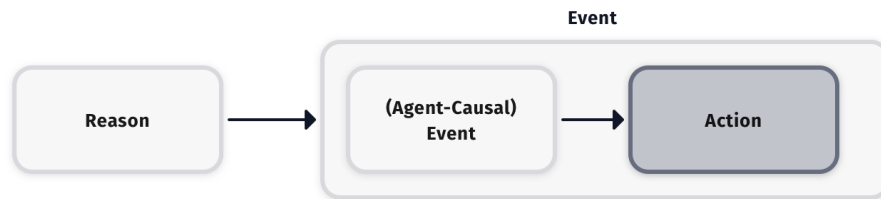


Figure 4.5: Modified Agent-Causal View

Like the previous proposal, the assessment of the prior causes view can be divided between unfree and free action. Starting with the former, I believe that the view fares at least, if not better, than the previous proposal. The current proposal evidently explains why some movement is an action by citing the agent as a direct cause. It also explains why that action was done for a reason by citing a reason as a prior cause of the agent’s agent-causing. Furthermore, since the agent’s agent-causing has a prior reason as a cause, it answers the question raised by Rice in previous section—i.e., why does an agent agent-cause an action?

Further assessments of the plausibility once more turn on whether one is compatibilist or libertarian. Libertarians are likely be wary of positing some antecedent causal source of the agent. In fact, the prior causes account seems to cut against sourcehood intuitions motivating libertarianism even more than the prior co-causes account. Compatibilists are, once again, less likely to be moved by this concern, since everything may have an antecedent causal explanation. So, if an agent has a reason as a prior cause, there is nothing particularly troublesome about this.

In sum, the proposal that reasons are prior causes of agent-causings has its own costs and benefits. On the one hand, the view would appear to get off the ground only by granting that the agent’s exercising her agent-causal power is an event. On the other hand, the view answers why agents agent-cause as they do. Furthermore, just like the proposal that reasons are co-causes, the current view is likely to align better with compatibilism than it is with libertarianism.

§4.3.3 Propensity-Based Account

Turn now to our to our third and final variation of a causal account of motivating reasons. O'Connor (2004, 2009)—and also in (O'Connor and Churchill 2005)—has defended (what I take to be) a variation of the prior causes account of motivating reasons. However, instead of reasons causing the agent, the reasons cause changes in the *propensities* (or probabilities) of the agent to act. Strictly speaking—at least according to O'Connor—the reasons don't cause the agent, they only exert *causal influence* on the agent's propensities to act.

To understand O'Connor's account, we need a brief clarification regarding O'Connor's account of action. While we've so far been assuming that basic actions are bodily movements, O'Connor takes them to be mental events. So, on O'Connor's account, the agent performs a basic action by directly causing an intention. (Of course, assuming nothing funny is going on, this will usually result in the agent's performing some bodily movement.)³³

O'Connor's full account involves distinguishing between two different relations an agent can have to reasons: (i) the relation of an agent acting *on* a reason, and (ii) the relation of acting *for* a reason. So, where we have been taking these claims to be synonymous, we will distinguish between them while dealing with O'Connor's account.

Start, first, with O'Connor's analysis of acting *on* a reason. His proposal is as follows:

Acting on a Reason. Agent *S* acts on a reason *R* just in case: (i) *R* increases the objective probability that *S* performs action *A* and (ii) *S* performs *A*.

Suppose that Isabel is considering casting a vote. She's undecided whether to vote in favor or against the candidate. A reason *R* presents itself to Isabel, increasing the probability that she raises her arm in order to cast a vote. She decides to raise her arm. In this case, she has acted *on R*. *R* increased the

³³ We might also note that O'Connor also adheres to a component theory of action, though nothing should turn on this.

probably of her raising her arm and she in fact raised her arm. (Note that, while Isabel is conscious of the reason in this case, one need not be to act on a reason.)

O'Connor then uses the acting *on* relation to further develop his relation of acting *for* a reason. The full analysis looks like this:

Acting *for* a Reason. An agent *S* acts for a reason *R* just in case: (i) *S* acts on *R* when performing action *A*, (ii) *S* is conscious of *R*, (iii) *S* causes the intention to *A* for the sake of a goal *G* (where *R* promotes *G*) and (iv) no other reason promotes *G* (O'Connor 2009: 121-122).

Conditions (i), (ii), and (iv), I take it, are straightforward enough. So, let me say a few words about condition (iii).

Recall the idea that reasons promote certain ends. For instance, suppose Isabel has a belief that raising her arm will cast a vote, and that she also has a desire to cast a vote. She would thus have a reason *R* to raise her arm. *R* promotes the action of raising her arm. But, it also promotes a further end—namely, that she raise her arm *for the sake of* casting a vote.

Now, also recall that O'Connor maintains that agents directly cause intentions to act. What does this look like? Suppose Isabel finds herself with reason *R*. While aware of *R*, Isabel causes an intention to act with the following content: *raise arm for the sake of voting*. Both *R*, and the intention caused by Isabel, have overlapping content—namely, the goal of casting a vote. Supposing *R* increased the probability that Isabel would raise her arm, she has not merely acted *on* *R*, she has also acted *for* *R*. According to O'Connor, it is in virtue of these additional features that “a further explanatory connection between that reason and the [action] is forged” (O'Connor 2009: 121).

Because O'Connor posits two distinct relations an action can have to a reason, we must decide which relation is best fit to play the role of motivating reasons. At least at one point, O'Connor (together with Churchill) suggests that “either type will help to explain the occurrence of a decision to

A'' (2005: 252). Given that the acting *on* relation is shared by both relations, acting *on* a reason appears to be both necessary and sufficient for being motivated by a reason. So, let's start by considering the relation of acting *on* a reason. The idea here would be that an agent is motivated by a reason *R* just in case the agent performed some action, and *R* increased the probability that the agent would perform that action.

The proposal strikes me as *prima facie* implausible. It entails that, whenever an agent performs an action, she is motivated by *every* reason that increased the probability that she would perform that action. But there seem to be clear cases where we would want to deny this.

To take one case of this, consider our previous example *Donation*. Suppose that the two reasons that Juan has to make a donation influence the probability that he will do so to varying degrees. Suppose, for instance, that the altruistic reason R_1 increases the probability from 0 to .8, and the self-motivated reason R_2 increases the probability from .8 to .8001. Now suppose that there is this increase in probability despite the fact that Juan completely disavows R_2 . If Juan were to make the donation, we should want to allow for the possibility that Juan was *not* motivated by R_2 (as well, of course, as the possibility that he may have been motivated by both reasons). The problem is that the proposal at hand can't allow for this, since every reason that increased the probability (of an action that an agent performed) is to be counted as a motivating reason.

So, while O'Connor seems to suggest that acting *on* a reason is what constitutes being motivated by a reason, perhaps it the relation of acting *for* a reason that is better fit to play this role. On this alternative proposal, an agent is motivated by a reason *R* just in case the agent acts *for* that reason (in the sense explicated above).

Some have objected that the account acting of *for* a reason is problematic because it invokes a *non-causal* relation between the reason and the action (Clarke 2003; Rice 2011). More specifically, the objection is that the reason's partially entering into the content of the agent's intention does nothing to forge a greater explanatory connection between the reason and the action.

In my view, the objections are compelling. Consider an example borrowed from Clarke (2003: 142): suppose a flash of lightning causes a brush fire. In such a case, can we cite some feature *of the fire* to explain why the flash of lightning caused the brush fire to occur? No. Whatever features the brush fire has are consequent to the lightning flash's causing it. In other words, since the brush fire occurs *after* the lightning strike, it's unclear how some feature of the brush fire could explain why the brush fire occurred. The same would appear to be the case when acting for a reason. The fact that part of the content from the reason enters into the content of the agent's intention cannot even partly explain why that intention occurred. Acting *for* a reason, then, carries with it no obvious advantage over acting *on* a reason.

In light of foregoing concerns, it would appear that the propensity-based proposal is unlikely to account for motivating reasons, whether it comes to unfree or free actions. Agent-causalists, it would seem, would do better by looking elsewhere for an account of what it means for an agent to be motivated by a reason.

In this section, we considered three variations of the causalist account of motivating reasons. Both the co-causes proposal and the prior causes proposal seem to do well enough in accounting for motivating reasons in the context of unfree actions—although the prior causes proposal may require taking on board a modified form of agent causation. In return, however, the agent-causalist is able answer why an agent agent-causes as they. On the matter free actions, neither account is entirely friendly to libertarianism, since invoking a reason as an added cause cuts against sourcehood intuitions that generally motivate libertarianism. Compatibilists are unlikely to have the same concerns, so I see no reason why a compatibilist couldn't, in principle, adopt either. As for the propensity-based proposal, it seems unable to account for motivating reasons across the board.

Having canvassed the available options for causalist accounts motivating reasons, I now turn to considering some alternative options.

§4.4 Teleological Realism

I turn now to considering teleological realist accounts of motivating reasons. While these accounts have had relatively few adherents in contemporary debates, George Wilson (1989) and Scott Sehon (2005, 2016) are two notable authors that have provided extended defenses. Here, I focus on the account developed by Sehon, since it has received the most extensive development of the two.

The core of teleological realism is the idea that there is a *sui generis* and primitive teleological relation between actions and reasons. As such, the teleological relation cannot be reduced to some further relation, such as a causal connection between the action and the reasons. This is not to say that the account is *anti-causalist*—i.e., that it is incompatible with there being a causal connection between the reasons and the action—rather, it treats any such causal relation as orthogonal to teleology.

Like event-causal theorists, teleological realists have typically presented their view as a way of explaining both action and motivating reasons. As we turn our attention to the teleological account, it is worth noting that the idea of a motivating reason will shift from a mental event to a potential state of affairs. For extant teleological realists, some movement is an action in virtue of its being directed at a reason, and that action is motivated by whatever reason it is directed at. Here, once more, I intend to prize these two commitments apart. I focus on whether the teleological realist account of motivating reasons can be integrated with an agent-causal theory of action:

Teleological Account of Motivating Reasons: An action *A* is motivated by a reason *R* just in case *A* is *teleologically directed* at *R*.

For the teleological realist, teleological directedness is expressed by a number of different locutions: acting for the sake of *R*, acting in order to accomplish *R*, and acting because of *R*, are all some examples. Again, when a reason is cited as an explanation of the action, this is to be understood as irreducibly teleological. If Isabel raises her arm in order to cast a vote, then the raising of her arm is teleologically

directed at casting a vote. That Isabel performed this action for that particular reason is neither “reducible [nor] analyzable into some further explanatory relation” (Sehon 2016: 66).

Teleological accounts provide a potentially attractive alternative for agent-causalists, especially those of the libertarian variety. We saw in the previous section that introducing a causal relation—between the reason and an action—ended up coming into conflict with agent causation in various ways. By introducing a non-causal, teleological relation, those concerns are side-stepped. The integration, then, seems to be a promising candidate for integration. On this proposed integration, some movement is an action in virtue of its being agent-caused, and that action is motivated by whatever reason it is teleologically directed at.

Positing an irreducible teleological relation might seem to make action explanation elusive. How can we be sure that an agent acted on a particular reason? Where the causalist account of motivating reasons allows us to test for this by looking for a causal relation, the teleologist might seem to offer no such alternative.

While Sehon maintains that reasons explanations are irreducibly teleological, he also believes that there are a number of criteria we can use to guide us in this process. He writes: “teleological explanations [are] part of an overall attempt to construct a theory of an agent, and part of [the] aim is to produce a theory according to which the agent is as rational as possible (Sehon 2016: 27).”

Sehon sees this interpretive endeavor as guided by two core principles:

Principle₁: Agents act in ways that are appropriate for achieving their goals, given the agent’s circumstances, epistemic situation, and intentional states.

Principle₂: Agents have goals that are of value, given the agent’s circumstances, epistemic situation, and intentional states.

Here’s an example adapted from Sehon that is meant to illustrate the interpretative process. Imagine that we see Isabel turn on the coffee maker. In trying to decide what explains Isabel’s behavior, we might

propose different explanations (or theories). For instance, we might posit that Isabel turned on the coffee maker in order to make herself a cup of coffee (*Explanation 1*). However, we could also consider more unlikely explanations, such as the possibility that Isabel turned on the coffee maker in order to send a message to an alien race on Mars (*Explanation 2*). Within the teleological framework, how are we to settle which theory is best?

According to Sehon, the principles of rationality tell us to make the agent as rational as possible by looking at all of her (past or present) beliefs, desires, values, and so on. So, if we look at Isabel's beliefs and desires, we see that she believes that the coffee maker is a reliable way of making coffee, that she has a project due soon, that she is tired, that coffee is a reliable way of alleviating her tiredness, and so on. In contrast, Isabel neither believes that aliens exist, nor that the coffee maker has the ability to transmit any messages. In light of these facts, we can conclude that *Explanation 1* wins out. The same sort of interpretive process can then be applied to any case of action explanation.

As I see it, Sehon's enhanced account of teleology adds too much. On the one hand, no explanation is given for what the relation is between the rationality principles and teleology. On the other hand, the two seem to come apart.

First, the relation between teleology and the rationality principles is in need of explanation. Sehon talks as if the rationality principles somehow constrain what teleological relations obtain, but no justification is given for this. As far as I can tell, there isn't any kind of conceptual or dependence relation between teleology and the rationality principles. Indeed, some comments of Sehon indicate that there is no such dependence relation between the rationality principles and teleology. On at least one occasion, he maintains that the rationality principles are brute, and so aren't explained in terms of any other facts (Sehon 2016:72). However, if they're brute in this sense, they can't be grounded in

teleology. It is thus hard to see what relation there is between teleology and the principles of rationality, such that latter constrains the former.³⁴

Second, I see no reason why the rationality principles and teleology couldn't come apart. Surely there are some possible worlds where agents are motivated by reasons, but those reasons never make the agents even *prima facie* rational.

Perhaps one could object that these principles are simply part of the *concept* of teleology. But, this raises its own difficulties. It now appears that teleology is no longer primitive, but is rather reduced to these rationality principles. Suppose, however, that one is willing to accept this reduction—being teleologically directed at a reason *just means* that that reason makes the most sense of the agent's behavior. In this case, the account fully collapses into a *rational interpretation* view of motivating reasons (Ryle 1949; cf. Paul 2021). In the spirit of charity, I take it a teleological realist would reject the conclusion that their view collapses into a competing account.

In light of the foregoing, one may opt in for a weaker reading of Sehon, according to which the principles of rationality are merely useful heuristics for determining what motivates an agent. However, it is unlikely that Sehon holds this position. Sehon (2016: 38) explicitly maintains, in cases where all the available considerations weigh equally for competing explanations, that it is simply *indeterminate* what reason an agent acted on. If the principles of rationality were only heuristics, then Sehon would only be committed to saying that we have no reliable way of knowing what reason the agent acted on.

Suppose we parted ways with Sehon, instead maintaining that the rationality principles are simply useful heuristics for trying to understand the behavior of agents. How much better does the view fare? As far as I can tell, the view has more plausibility. Furthermore, combining a teleological account of motivating reasons with an agent-causal account of action seems to raise no obvious concerns the way

³⁴ See Sehon (2016: Ch. 3) for an attempt to deal with irrational actions.

an integrated causal account did. The plausibility of the integration stands or falls with how sympathetic one is to irreducible teleology.

Speaking for myself, I have an insufficient grasp on irreducible teleology to make any definitive assessment. Perhaps this is to be expected when dealing with a *sui generis* and primitive notion. However, I don't find myself with the same feelings when considering other primitive notions. For instance, I find myself convinced (and perhaps you do too) that causation is an irreducible and primitive relation. When one thing causes another, there are no further facts into which this can be reductively analyzed. Despite all that, I find myself with a strong sense that causation is a real, irreplaceable part of the world's furniture in a way I simply do not when I consider teleology.

§4.5 A Causal Powers Account

I turn now to our final view of motivating reasons, which I shall refer to as a *powers account of motivating reasons*. The view outlined here is an attempt to reconstruct an account suggested by a handful of theorists (e.g., Hyman 2015, O'Connor and Jacobs 2013, O'Connor 2021). So, while the account is broadly in the spirit of these authors' proposals, I leave it open to what extent it diverges from them.

The core of this view is that reasons are causal powers. Hyman, for example, has suggested a dispositional account of motivating reasons, wherein reasons just are dispositions. Dispositions are tendencies, potentialities, capacities, etc., of objects to behave in certain ways across different circumstances. For example, salt has the disposition of *solubility*. It has the potential, tendency, etc., to dissolve in the appropriate circumstances (e.g., when placed in water). Hyman takes reasons to be of the same species. If Isabel has a reason to raise her arm, then she has the potential, tendency, etc., to raise her hand in the relevant circumstances.

Similarly, O'Connor has suggested on different occasions that reasons may be causal powers. Causal powers are, as we mentioned earlier, properties whose essences are linked to certain effects.

We mentioned *charge* as a causal power of particles to repel other particles, and the *agent-causal power* as a causal power to bring about actions. On this proposal, then, reasons are causal powers that are inherently linked to *particular kinds* of actions. If *R* is a reason to *A* for the sake of *G*, then *R* is partly directed at *A*-ing.

Dispositions, as I use the term, are a *species* of causal powers—they are *passive* powers. Passive powers are powers that need an external stimuli for their manifestation. Solubility is a passive power, since it needs an external stimuli (e.g., water) in order for that power to be manifested.

By contrast, active powers do not need an external stimuli for their manifestation. Mass is one such potential example. A body with mass is at all times exerting a gravitational force on all other objects; it does not need an external stimuli in order for it to exert its gravitational force. It is constantly and actively exercising its power.

In between these two classes of properties lies *spontaneous* powers. Like active powers, spontaneous powers do not need an external stimuli for their manifestation. However, unlike active powers, they are not constantly active. Rather, they are manifested spontaneously. For instance, *decay* may be an example of a spontaneous power. A radioactive atom is something that has the power to decay. The atom does not need a stimulus for the decayed state to be manifested—generally, at any moment in time, the atom may simply exercise its power. Many authors, then, take an agent-causal power to be a spontaneous power that can be exercised at will by the agent.

Thus understood, Hyman and O'Connor agree that reasons are causal powers—they just disagree over what kind of causal powers they are. Hyman evidently takes them to be passive powers, while O'Connor is likely construing them as spontaneous powers. In what follows, I focus on the idea that reasons are spontaneous powers.

Postulating a 'reasons powers' involves invoking two kinds of powers in action: an agent-causal power and a reasons power. So, how do they relate? O'Connor seems to make the following suggestions: (i) an agent-causal power is "an enduring, more general power of choice;" (ii) reasons are

“particular ‘all-in’ powers one variably has on any given occasion to choose this or that;” and (iii) reasons powers are “partly grounded in” the agent-causal power. Distilling the core idea, one way to understand these claims is that an agent-causal power is a sort of higher-order power to take on reasons. If this is correct, an agent-causal power needs a reason for its manifestation—without a reason, an action cannot be performed. This is because, according to O’Connor, reasons are needed to make a decision, to “choose this or that.” By analogy, we might think of an agent-causal power as a firearm, and reasons as the ammunition to use the firearm on a particular occasion.

One upside of a powers-based account would be that it incorporates aspects from both the causalist and teleological accounts of motivating reasons (Hyman 2015). It incorporates causalist aspects insofar as the reasons are causally relevant in explaining the action, in much the way that the disposition of solubility is causally relevant in explaining the salt’s dissolving. It likewise incorporates aspects of teleology insofar as the powers are *directed at* certain manifestations (i.e., actions).

One concern for the current proposal is related to how reasons are to be individuated. Suppose that Juan takes R_1 to be a reason to make a donation, where R_1 promotes self-interest. Suppose also that Juan takes R_2 to be a reason to make a donation, where R_2 promotes altruism. Because powers are typically thought to be individuated by the effects that they’re directed at, our current proposal would seem unable to individuate these two powers at hand, since both R_1 and R_2 are directed at the same action—namely, making a donation.

The way out of this, it seems to me, is to take on board O’Connor’s suggestion that reasons are actually directed at causing intentions with specified content. So, what R_1 is really directed at is causing an intention *to make a donation for the sake of making oneself look good*, and what R_2 is really directed at is causing an intention *to make a donation for the sake of alleviating suffering*.

Earlier, it was objected that causalist views can’t account for an agent’s having control over which reason they acted on. Does a powers account fare any better? At face value, the answer would seem to be *no*, since agents can’t cause which power they act on. However, upon further inspection, that answer

is less obvious. Consider whether agents have control over whether they exercise their agent-causal power. Plausibly, if agent causation is possible, then agents don't exhibit control over the exercise of this power by causing anything—what they do is exhibit control over their actions by simply causing those actions. They exert control of their agent-causal power by choosing what to cause. The same might be said of an agent with respect to her reasons. The agent exhibits control over which reasons they act on by choosing which intentions to cause.

One might think that a version of Clarke's previous objection is unaddressed. By appealing to the intentional content that is caused by agent, aren't we getting the explanation backwards? How can something that comes after the agent-causing help to explain the agent-causing itself? To this, the powers theorist may object that the reason is actually operative throughout the whole causal process. Indeed, that causal process was only made possible by that reason's partly grounding it.

The discussion, however, invites a potentially troubling question for the agent-causalist: does the powers-based account accurately capture what we want in an account of motivating reasons? More specifically, in what sense is an agent *motivated* by a reason simply in virtue of exercising the relevant reason's causal power? On the one hand, the reason (construed as a causal power) does help to explain the action—the reason allowed the agent to cause the action they did. On the other hand, this analysis seems to leave something out.

Reasons, I take it, exert some kind of “pull” on the agent, and it is partly in virtue this pull that the agent is supposed to be motivated by it. When an agent is motivated by a reason, they act in light of that reason. If Isabel has a reason *R* to raise her arm, she has, according to this proposal, a causal power that enables her to raise her arm. If she in fact does raise her arm, she has exercised her power *R* to do so. But has she therefore acted in light of that reason? Was she *motivated* by that reason? Whatever this ultimately amounts to, it seems to be absent from our current picture, since all we have

is the bare exercise of a causal power. And whether anything can be supplied to fill this gap, I'm at present unsure.³⁵

§4.6 Concluding Remarks

The central task of this chapter has been to get clearer on the available options for developing an account of motivating reasons that can be integrated with an agent-causal theory of action. In this chapter, we canvassed five different proposals. The key takeaways were that the first two causalist proposals—the co-causal and prior causes views—fared well enough in accounting for action as such. When extended to free action, the conclusion was that they were more friendly to compatibilism than libertarianism. The third causalist proposal—the propensity-based view—was found wanting when accounting for action in general. The final two proposals—teleological realism and the causal powers views—were likewise found to come up short in accounting for action in general. As a libertarian, if I were forced to bet, my money would be put on the powers account—but I would not bet much. In my view, libertarians have more work to do if they want to provide a unified account of motivating reasons.

I will close with a more speculate comment. In seeking an account of motivating reasons, we have been assuming that there *is* such an account to be found. Elsewhere, however, others have argued that any such account is bound to fail (cf. Boghossian 2014). Assuming, as I do, that something like agent-causal libertarianism is true; and, assuming the conclusions of the chapter are not entirely off-base, perhaps the agent-causalist is lead to a similar conclusion. Perhaps, that is, the agent-causalist must be content with treating motivating reasons as an unanalyzable primitive in her overall theory of action. To some, this may be no better than postulating irreducible teleology. But note that, unlike the teleologist, this would not involve postulating a new *sui generis* category to account for action

³⁵ One might think that O'Connor's claim that reasons alter the probabilities of an action can help supply what is needed. However, in my view, the mere alteration of probabilities seems inadequate to fill this gap.

explanation. Rather, it would simply take action explanation as something that resists reduction. At any rate, for those who believe that a reduction is in order, more work needs to be done.

Chapter 4, in part, is currently under commission for publication (as “Agent Causation and Motivating Reasons”) and may appear in *Philosophical Compass*, Martinez, Joseph. The dissertation author was the sole author of this paper.

Chapter 5

Agent Causation, Responsibility, and Reasons-Responsiveness

§5.1 Introduction

According to agent causation, we are free and morally responsible partly in virtue of being agents who are directly causally involved in bringing about their actions. While much has been said to work out the metaphysical underpinnings of this sort of view, less has been said about how the view handles cases of moral responsibility. It might be thought that this will more or less involve applying the conditions for free will to various cases. If the agents meet the conditions, they are morally responsible; and if they don't, they are not morally responsible. I think that this is a good starting point for the agent-causalist. However, as I will argue here, applying the standard agent-causalist conditions for free will are insufficient to account for important aspects of our moral responsibility assessments.

Here, I think there are important insights to glean from the recent explosion of so-called reasons-responsive accounts of moral responsibility and the special attention they have paid to addressing wide-ranging issues in moral responsibility, such as addiction, psychopathy, implicit bias, and more. Arguably, the traction these views have gained is partly explained by their ability to provide compelling treatments of these core questions. Integral to the agent-causalist project is that the view likewise offer, or at least take seriously, these questions. The aim of this chapter is to help begin to fill in this lacuna. I hope to accomplish this by drawing out shortcomings of extant agent-causal accounts on matters of moral responsibility, and then pointing the way towards developing a more plausible form of agent causation—one that embraces a robust form of reasons-responsiveness.

While the catalyst for this discussion is agent-causal libertarianism, the conclusions here have wider-ranging consequences for those interested in moral responsibility. As just noted, reasons-

responsive accounts of moral responsibility have gained widespread support in the literature. Chief among these views are those that ground reasons-responsiveness in *counterfactual* conditionals about how an agent would act in various circumstances. This chapter considers the possibility of adopting these resources for the purposes of enhancing agent-causal accounts. However, if the conclusions here are correct, agent-causal libertarianism is incompatible with this form of reasons-responsiveness. Nonetheless, as I shall argue, the features that make the two incompatible can be generalized in an important way to extend to anyone that grounds reasons-responsiveness in counterfactual conditionals. Not only should the agent-causal libertarian seek out a form of reasons-responsiveness that is not tethered to counterfactuals, so should (almost)³⁶ every other reasons-responsive theorist.

The chapter will proceed as follows. In the next section (§5.2), I start by briefly recapitulating the core elements of agent causation and explaining why they are unable to account for important responsibility-relevant features. I then consider how reasons-responsive accounts appear to fare much better in accounting for these features and so suggest that the agent-causalist could benefit from accommodating those elements from the reasons-responsive accounts that allow them to capture those relevant features. After laying this out, I examine what I take to be the central problem for this integration—whether the relevant counterfactual conditions can be satisfied within the agent-causal framework. My verdict is negative. We have good reason to be skeptical about whether the indeterministic counterfactual conditions can be satisfied. I explain how the point generalizes out and conclude by examining potential ways forward. I suggest that the agent-causalist should develop a more robust, though non-modal, form of reasons-responsiveness.

§5.2 The Gap

³⁶ As we shall see, those compatibilists who think that free will and moral responsibility *require* determinism will be insulated from this concern.

In their current forms, agent-causal views cannot fully account for central aspects of our moral responsibility practices. To see this, we need to lay out the conditions that an agent-causalist posits for moral responsibility and show that they leave some important aspects of moral responsibility assessments unaccounted for.

With a few caveats, the following three criteria are generally taken by agent-causal theorists to be necessary and sufficient for settling an agent's moral responsibility status.³⁷ The first condition was discussed at length in chapter 2, the second condition in chapter 3, and the third condition in chapter 4:

1. **Substance Causation:** An agent *S* is morally responsible for performing some action *A* only if *A* was directly caused by the agent.
2. **Indeterminism:** An agent *S* is morally responsible for performing some action *A* only if *S*'s causing *A* was not itself casually necessitated.
3. **Acting on a Reason:** An agent *S* is responsible for performing some action *A* only if *S* did *A* for a reason.^{38,39}

There is room for greater precision. Nonetheless I believe these three conditions serve as a useful approximation for how agent-causalists have typically wanted to cash out the conditions for moral responsibility. If this is correct, we should be able to use the foregoing three conditions to adjudicate the various cases of moral responsibility.

³⁷ Nearly all agent-causalists have explicitly endorsed the first two conditions (e.g. Chisholm 1964; Taylor 1992; O'Connor 2000; Clarke 2003; and Franklin 2019). Furthermore, most recent treatment of agent causation appear to endorse the third condition as well (e.g., O'Connor 2000; Clarke 2003; and Franklin 2019).

³⁸ Condition 3 is likely too weak for our purposes here. Plausibly, we want the person to have been able to do the right thing for the right reason in that context. I don't have the space to address this point here, so feel free to read this third condition as making this provision.

³⁹ We may also need an epistemic condition for moral responsibility. To expedite the discussion, however, I will simply assume that any such condition is satisfied going forward.

Start with a case where I think these conditions fare well, such as Jasmine stealing her brother's phone. Let's suppose that Jasmine, and not merely some events going on inside Jasmine's head, brought about her action of stealing the phone. Furthermore, suppose that Jasmine was not causally necessitated to steal the phone, and Jasmine acted on intelligible, though objectional reason, such as getting back at her brother. Jasmine performed a morally bad action and she meets all of the relevant conditions for moral responsibility, so Jasmine morally blameworthy for what she did. This seems to be the correct verdict.

Now consider the following two cases⁴⁰:

- A. Bella has clinical depression, which often produces systematic weakness of will in the form of listlessness and apathy. One day, while experiencing an episode of depression, she fails to keep an important promise that she made.

- B. Gael has acquired damage to the prefrontal cortex of the brain. A side effect of this is that he often struggles to conform to his own judgments about what he ought to do. One day, when his friend asks him for help in an urgent situation, he fails to extend a helping hand, though his considered judgment was to help his friend.

(A) and (B) are meant to be representative of a large range of cases where agents face various kinds of difficulties in acting. I submit that cases like (A) and (B) are not easily handled by the above three conditions. They cannot be used to fully adjudicate an agent's responsibility status.

Let's focus our attention on Bella. According to our three conditions, if we want to know whether Bella is morally responsible for failing to keep her promise, we need only consider (a) whether she directly caused her failing to keep her promise, (b) whether her doing so was causally necessitated, and

⁴⁰ Adapted from Brink (2013).

(c) whether she caused the action that she did for a reason. For our purposes, let's suppose that all the conditions are satisfied. That is, suppose that (a) Bella directly caused her failing to keep her promise, (b) her doing so was not causally necessitated, and (c) she acted on a morally bad reason (say, she wanted to watch a movie instead). Given what we have said, it would follow that Bella is morally blameworthy for failing to keep her promise.

This strikes me as too fast. Intuitively, even if Bella satisfies all three of the above conditions, I can still imagine her not being (fully) morally responsible. Part of the concern is that Bella has a certain cognitive impairment that is affecting her control in a way that is not being accounted for. Our account of moral responsibility ought to make room for such features.

I believe that there are accounts that are present in the literature that fare much better on this matter. Reasons-responsive accounts are, perhaps, the strongest example. As others have noted, one reason that many have found reasons-responsive accounts appealing is precisely because "they offer elegant explanations of the conditions in which various excuses and exemptions apply" (McKenna 2013). I am inclined to agree with McKenna's assessment. Reasons-responsive accounts appear to provide richer resources when it comes to dealing with cases like those involving Bella and Gael. Supposing that this is correct, it would be potentially fruitful for the agent-causalist to examine why this is so, and whether she too can take can partake in these resources. In the next section, I examine this possibility in more detail.

§5.3 Filling the Gap: Reasons-Responsiveness

In the previous section, I identified a gap for agent-causalists and suggested that reasons-responsive accounts are in better shape when it comes to this problem. In this section, I explain how reasons-responsive accounts address cases like (A) and (B). After doing this, I briefly sketch out what it would look like for an agent-causalist to integrate with the kind of reasons-responsive account sketched here, and why it would potentially be beneficial to do so.

Reasons-responsive accounts encompass a rather large, and in some ways diverse, number of views. While, at their core, they all attempt to capture freedom and responsibility by reference to the agent's appropriate sensitivity to reasons, they do so in different ways. Here is a sampling of authors who may reasonably be considered reasons-responsive theorists, despite their divergence in the way they cash this out: Fischer and Ravizza (1998), McKenna (2013), Nelkin (2011), Sartorio (2016), Wallace (1994), and Wolf (1990).

Given the number of distinct views listed, it would certainly take us too far afield to consider these views individually. Here, then, I plan to focus on one particularly influential reasons-responsive account—namely, the one developed Fischer and Ravizza. Since Fischer and Ravizza's initial defense of the view, various writers have developed similar views, or have built off their initial account (e.g., Haji (1998), Brink and Nelkin (2013), and Vargas (2013)). For many in the literature, Fischer and Ravizza represent the *de facto* reasons-responsive view. What follows is a brief and partial reconstruction of the core aspects of their account. It is worth noting upfront that, while my goal is to stay faithful to the original account, the following reconstruction takes on board a few amendments that subsequent authors have suggested over the years. Noteworthy amendments will be noted.

The guiding idea behind reasons-responsiveness is that moral responsibility is intimately connected to our ability to understand and act on the available reasons. Importantly, our ability to be reasons-responsive is not just a matter of how an agent *actually* acts, but also how she *would* act across a range of different circumstances; how she would act in response to changes in reasons across relevantly similar circumstances. McKenna (2013: 154) provides a useful analogy:

[C]onsider a simple example of the sensitivity of a primitive gizmo, a thermostat.

Suppose a thermostat is set at 76 degrees (Fahrenheit) and the room the thermostat is in turns out to be 76 degrees. One might wonder if the thermostat's setting accounts for the temperature of the room. After all, it might be disconnected and so merely a fluke that its setting and the room's temperature are 76 degrees. When we learn that

the room would come up to 78 were the thermostat set to 78, or would come down to 74 were the thermostat set to 74, and so on for numerous other values high and low of 76, we do not just learn something about the way the thermostat would behave; we also learn about how, in the actual scenario when it is set to 76, it does behave; it plays a certain causal role from reliable and suitably sensitive resources.

Just as we assess the thermostat's reliability by examining how its state would vary across circumstances, reasons-responsive accounts assess an agent's control over her actions by examining how an agent would act across different scenarios.

Fischer and Ravizza posit two distinct capacities in their account of reasons-responsiveness. The first capacity involves the agent's ability to *recognize* different reasons in different scenarios. In particular, Fischer and Ravizza maintain that the requisite recognitional capacity involves the ability to simply "recognize the reasons that exist" and to exhibit a "regular and understandable pattern" in doing so.

For instance, imagine that, in the way things actually play out, no one of importance is present when Pablo raises his arm and signals the execution of the civilian. We might also wonder whether Pablo would recognize other sufficient reasons to not do what he does. For instance, let's grant that he would recognize the presence of his 8 year old daughter as a reason to not raise his arm, or if he would recognize a bribe of \$10,000 as a reason to not raise his arm. Furthermore, suppose that he would *not* recognize the following as reasons to not raise his arm: whether he hit a red light that morning, whether it was 2am or 2:01am, whether Canada was north of Mexico. All of this would tell us that Pablo has some capacity to recognize the actual reasons there are, not simply take any old thing as a reason to act. While the threshold for the recognitional capacity is left somewhat vague, the idea is that there is some minimal threshold of being able to recognize reasons, such that the agent has the recognitional capacity required for moral responsibility.

In addition to this, Fischer and Ravizza posit the agent must exhibit a coherent pattern in the kinds of reasons that she recognizes. For example, if Pablo were to recognize a bribe of \$10,000 as a reason to not raise his arm, then he should also recognize a bribe of \$15,000 as a reason not to raise his arm.⁴¹

Recognitional Capacity: An agent *S* possesses the requisite recognitional capacity if and only if *S* exhibits an understandable pattern of reasons-recognition.

The second capacity involves the agent's ability to *act on* the reasons she recognizes as sufficient. We don't just want to know whether an agent can recognize sufficient reason to do otherwise, but also whether she can conform her actions to her judgements. If Pablo could, say, recognize that the presence of his daughter as a sufficient reason to not raise his arm, but never acted on it, he would intuitively lack an important kind of ability to *respond* to that reason. The same goes for any sufficient reason the agent may have to do otherwise. For this reason, Fischer and Ravizza maintain that an agent possesses the relevant *volitional* capacity when she is able to act on those sufficient reasons to otherwise in a suitable proportion of possible worlds.⁴²

Volitional Capacity: An agent *S* possesses the requisite volitional capacity when performing action *A* if and only if, in a suitable proportion of relevantly similar counterfactual scenarios, *S* recognizes sufficient reason to *B* (where $A \neq B$) and *B*'s because of that reason.

⁴¹ Fischer and Ravizza initially pitched their account not in terms of whether the agent was directly reasons-responsive, but whether the *mechanisms* that an agent acted on was reasons-responsive. Various authors have subsequently come to endorse an agent-based account of reasons-responsiveness over a mechanism-based account. My articulation of reasons-responsiveness relies on the former, since it better fits with an agent-causal account.

⁴² Fischer and Ravizza originally only required that an agent act on a sufficient reason to do otherwise in *one* other possible world. Others have come to criticize this as too weak and have instead proposed that an agent act in a suitable proportion of possible worlds (e.g., McKenna (2013), Todd and Tognazzini (2008), and Brink and Nelkin (2013)). Because I'm persuaded that the original formulation is too weak, I adopt the amended version here. Nothing I say below should hinge on this, however.

An agent, then, is said to be reasons-responsive when possesses both the recognitional and volitional capacity.

This concludes our rough sketch of reasons-responsiveness. With just a rough sketch of this account in hand, we are now positioned to see how it can help address cases like (A) and (B) above.

Return to the case of Bella:

- A. Bella has clinical depression, which often produces systematic weakness of will in the form of listlessness and apathy. One day, while experiencing an episode of depression, she fails to keep an important promise that she made.

The concern raised earlier was that agent-causal accounts could not fully account for Bella's impairment. Even if she satisfied all three conditions required by the agent-causalist, it still seemed like an open possibility that she was not (fully) morally responsible. When we turn to the reasons-responsive account, however, there is more to be said. While the exact responsibility status of Bella will turn on the specific details of the case, we seem to at least have the resources to in principle settle whether she is morally responsible or not.

What we want to know, in our present context, is whether Bella is reasons-responsive. To determine this, we need to look at how she would act in relevantly similar counterfactual scenarios. If it turns out that Bella acts on reasons to otherwise in a relatively sparse number of counterfactual scenarios, then Bella would still lack the volitional capacity for reasons-responsiveness. This would mean that she is not morally responsible for failing to keep her promise. On the other hand, if it turned out that Bella both recognizes sufficient reason to do otherwise and acts on those reasons in a suitable proportion of possible worlds, then we would have confirmation that she was reasons-responsive. We could then infer that she is morally responsible for failing to keep her promise.

No doubt the foregoing discussion leaves a number of important points untouched. For instance, the above discussion brackets concerns regarding compatibilism and libertarianism. However, I believe

the discussion still helps to make draw out the advantages that reasons-responsive accounts have over agent-causal accounts when dealing with cases like (A) and (B).

An intuitively appealing way forward for the agent-causalist, then, is to consider whether she can simply borrow those resources from the reasons-responsive account in order to strengthen her own account. Here, then, would be one way to go about integrating the two views:

Agent-Causal Reasons-Responsiveness (ACR): An agent *S* is morally responsible for an action *A* if and only if (i) *A* was directly caused by *S*, (ii) *S*'s causing was not causally necessitated, (iii) *S* possesses the recognitional capacity, and (iv) *S* possesses the volitional capacity.⁴³

There may be various reasons for the agent-causalist to resist this proposal. For one, Fischer and Ravizza present their account of reasons-responsiveness as one that captures the kind of *control* that is required for moral responsibility. In contrast, agent-causalists seem to present their own account of control, one that is grounded in the direct causal involvement of the agent.

Insofar as we aim to integrate the two views, I believe the best way to see the relation between these two accounts of control is to see reasons-responsiveness as *enhancing* the agent-causalist's account of control. As agent-causalists, we may say that direct agent causation secures a vital kind of control for free and responsible action, but nonetheless supplement this control in various ways. In this case, the recognitional and volitional capacities put forth may provide a way securing an even greater degree of control than just agent causation alone.

Another related concern may be that the account of action that Fischer and Ravizza endorse is an event-causal one, whereas the agent-causal account endorses a substance-causal view. Despite this difference, however, there's no reason to think that we cannot simply amend the view to include the

⁴³ I drop the AR condition from above since it is implicit in condition (iv).

agent-causal account of action. What is central to the reasons-responsive account is not the particular view of action it utilizes, but the counterfactual analyses that it gives to being reasons-responsive.

Despite these concerns, then, ACR provides a potentially appealing way to amend agent causation insofar as it allows the agent-causalist address the present gap, and does so in a way without requiring the agent-causalist to give up any of her core commitments.

§5.4 Problems With Integration

Despite the initial appeal of ACR, it faces serious a difficulty that, in my view, undermines its plausibility. My aim in this section is to articulate this difficulty. The difficulty stems from a tension between, on the one hand, the counterfactual conditions required by reasons-responsiveness and, on the other hand, the indeterministic condition required by agent causation. My discussion will proceed in two stages. The first stage is to consider claims that libertarians have made on other fronts which support the view that counterfactuals involving free agents lack truth conditions. The second stage is to draw on recent work on counterfactuals that supports the view that (nearly) all indeterministic counterfactuals lack truth conditions.⁴⁴ While I do not take the following lines of motivation to be conclusive, I think they provide us with sufficiently weighty reason to hold off from thinking that the counterfactuals required for the integration have clear truth conditions. Overall, the conclusion will be that this threatens the entire project integration.

§5.4.1 Libertarian Commitments

Agent-causal libertarians, and libertarians more broadly, all endorse an indeterministic constraint for free and responsible agency:

⁴⁴ Specifically, would counterfactuals are not true, where these are counterfactuals about what a free agent *would* do.

Indeterminism (I): An agent S is morally responsible for performing some action A only if S 's causing A was not itself casually necessitated.

We noted earlier that this constraint committed one to the claim that, up until the moment the action is performed, and holding everything fixed up until that point, it has to be open to the agent to perform the action in question or not perform the action in question. In what follows, I want to discuss two arguments that have been raised against libertarianism, both of which show that many libertarians have endorsed claims that should commit them to saying that counterfactuals about what free libertarian agents would do are not true.

The Roll Back Argument. Start, first, with van Inwagen's (2000) Roll Back Argument. Broadly speaking, the Roll Back Argument can be construed as a type of Luck Objection, where Luck Objections consist of a family of arguments that purport to show that the presence of indeterminism undermines the control that is required for freedom and responsibility. The Roll Back Argument attempts to show this by means of the following kind of reasoning.

Imagine a free agent (in the libertarian sense), Alicia. At time t_1 , Alicia is deliberating about whether to utter some lie. As it turns out, at time t_2 , Alicia freely utters a lie. Now, we are asked to imagine God “rolling back” time to t_1 right before Alicia originally uttered the lie. Given that Alicia has libertarian free will, it is open to her to *not* utter the lie when God rolls back the time. We can imagine, then, God continually rolling back time and letting things play out, in which case sometimes Alicia would utter the lie, and sometimes she would not. van Inwagen argues that, if God continued to do this, the probability distribution of Alicia's actions would have to converge on some definite and set probabilities—perhaps 50% lying and 50% not lying. Since Alicia has no control over this probability distribution with which here actions must converge, she can't be free.

Let's put to the side whether the argument succeeds. What is important to note is that, when we examine responses to the Roll Back Argument by libertarians, they *don't* object to the assumption that Alicia would sometimes perform different actions when God rolls back time. Rather, they object to the assumption that we would be unfree if there were some definite probability distribution with which our actions had to conform to in the long run (see, e.g., Balaguer 2010; O'Connor 2011; Eckstrom 2011; Franklin 2011).⁴⁵ We can reasonably infer from this that these libertarians accept van Inwagen's contention that, if God were to roll back time to the moment right before the agent makes a decision, there would be no settled fact about what the agent *would* do in that circumstance.⁴⁶

If this is correct, the point generalizes to all counterfactuals involving free actions. There is no settled fact about what an agent would do in a situation where they are free. For instance, there would be no settled fact about the following counterfactual:

If Pablo's daughter were present, Pablo would freely not raise his arm.

Pablo *might* raise his arm, but then again he might not.

The larger problem in our present context is that it follows that an agent can never be reasons-responsive (according to the account that we're considering). This is because reasons-responsiveness requires that it be true of the agent that she recognize sufficient reason to otherwise and act on those reasons in a suitable proportion of counterfactual scenarios. But, if it's never true what an agent would

⁴⁵ Buchak (2013) is a notable exception insofar as she argues that van Inwagen has failed provide reason to think that there is some definite probability distribution towards which the decisions would converge. It should be noted, however, that it's unclear whether Buchak fully embraces a Libertarian position.

⁴⁶ If there is some definitive probability distribution, then would the possible worlds bear this out? I don't think so. I think the thought experiment is only a heuristic. As we shall see shortly, van Inwagen thinks there is no settled fact about what we would do in these counterfactual scenarios. At any rate, there also other concerns...maybe with there being infinite sets of possible worlds...and also with the fact that, for any possible world, there will be a duplicate up until that point where the agent freely makes a decision, and where the agent doesn't make the same decision. The thought experiment seems to be, not that there would be distinct possible worlds actualized, but that how often these certain possible worlds would be actualized would converge on a set probability.

do in those counterfactual scenario, then she can never satisfy the conditions for being reasons-responsive.

Frankfurt Style Cases. Turn now to Frankfurt Style Cases (FSC). A few moments ago, I noted an implication of holding to indeterminism, which was: up until the moment the action is performed, and holding everything fixed up until that point, it has to be open to the agent to perform the action in question or not perform the action in question. This commitment (or something closely approximating it) has sometimes been described as the *Principle of Alternate Possibilities*.

(PAP) An agent is morally responsible for an action only if she could have done otherwise than that action.

FSCs consists of a family of different counterexamples directed at PAP. Here is one such generic example:

An agent, Maria, is deliberating about whether to perform some action, such as steal from the store. Unbeknownst to Maria, there is a powerful neuroscientist who has implanted a chip in Maria's brain that is able to recognize her brain states. If the chip were to recognize that Maria will decide to not steal, then the chip would activate and ensure that Maria will steal the item. As it turns out, Maria decides to steal the item on her own and the chip never activates.

The strategy of (many) FSCs is to provide a case where an agent is intuitively morally responsible, yet lacks the ability to do other than what she did. In the above example, we have a 'Frankfurt intervener' in the form of a chip. The role of the Frankfurt intervener is to cut off the agent's ability to do other than what she does. In our example, the Frankfurt intervener would do this by forcing the agent to do the particular action as soon as she gave signs that she wouldn't do it. However, as the way things *actually* play out, the Frankfurt intervener does not activate, and Maria acts 'on her own free will'. So,

while she appears to lack the ability to other than what she did, she still appears to be morally responsible for what she did.

Historically, most libertarians have endorsed some version of PAP.⁴⁷ For our purposes, we need not take a stand on the matter. What is important is to see the dialectic that has emerged regarding FSCs.

Arguably the most popular response to FSCs has been the Dilemma Defense.⁴⁸ As the name suggests, the defense proceeds by raising a dilemma for the proponent of FSCs. It goes something like this: as it stands, FSCs are underspecified as to whether the agent has libertarian freedom or not. (Horn 1) So, first assume that the agent does have libertarian freedom. If so, the Frankfurt intervener will not be able to know beforehand what the agent will do. This is because, given the presence of indeterminism, we cannot know with certainty what the agent will do beforehand. By the time the Frankfurt intervener is able to kick in, the agent will have already *started* to try to do otherwise. So, FSCs fail to show that the agent lacks an alternate possibility. (Horn 2) Instead, assume that the agent does not have libertarian free will. If we assume this, we are merely begging the question against the libertarian. Thus, FSCs fail to show that we lack alternate possibilities if we have libertarian free will. Either way, FSCs fail to undermine PAP.

Let's focus our attention on the first horn of the dilemma and the ensuing dialectic. In response to the first horn of the dilemma defense, critics have conceded that the Frankfurt intervener couldn't know with certainty what the agent would do beforehand. However, because the Frankfurt intervener could still prevent the agent from completing the action, these critics have maintained that what remains is a mere "flicker of freedom"—an alternative possibility that is not robust enough to ground

⁴⁷ Those libertarians who do not endorse PAP generally maintain that, while there may always be some alternate possibility open to the agent, this alternate is not *explanatorily relevant* when it comes to freedom. That is, although it is present, the alternate possibility does not make one free. Those who endorse PAP are typically referred to as 'Leeway' theorists and those who do not endorse it as 'Sourcehood' theorists.

⁴⁸ Sometimes also called the W-Defense, named after David Widerker who first popularized the defense. See Widerker (1995).

moral responsibility.⁴⁹ Libertarians have responded to this in a number of different ways. Some have, for instance, have responded by giving up PAP. These libertarians maintain that, while Libertarianism entails that we have alternate possibilities, they don't make us free; they're explanatorily inert when it comes to moral responsibility. Others have been more optimistic and have instead stood their ground, defending the robustness of these alternate possibilities in FSCs.

Regardless of where one may come down on the issue, most interlocutors (including libertarians and compatibilists alike) grant that there is no way to *know with certainty* what the agent will freely do. The best that the Frankfurt intervener can do is prevent the action from being completed.

While Libertarians may go different ways on the matter, one reason that they might believe that the Frankfurt intervener can't know with certainty what the agent will do is because there is no true fact of the matter about what they will freely do.⁵⁰ But, if there is no there is no true fact about what an agent *will* freely do, then presumably there is no fact about what an agent *would* freely do.⁵¹ This is because what one *will* do can simply be embedded into a counterfactual conditional, thus making that counterfactual false. If it is not true that Pablo will freely raise his arm, then the following counterfactual is also not true:

If Pablo's daughter were present, he would freely raise his arm.

The counterfactual is not true because the consequent cannot be true while the antecedent is true. And, once one concedes that counterfactuals involving free actions cannot be true, one must deny that agents can satisfy the counterfactuals conditions for being reasons-responsive.

⁴⁹ See Pereboom (2014).

⁵⁰ I use the locution "no true fact of the matter" to refer to any position that denies that future facts about free agents can be *true*. This includes positions that say that all such facts are *false* and those that say that all such facts are *indeterminate*.

⁵¹ The inverse may not hold. If there are no facts about what an agent would freely do, there may still be facts about what an agent *will* freely do. Hájek seems to endorse such a view. The reason for this asymmetry may be due to a difference in one's commitment on the metaphysics of time. If one is a presentist, presumably there may be no facts about what will happen in the future (given indeterminism). If one is an eternalist, all events in the history of the universe exist all at once. So, presumably that would be enough to explain what I will do. This may suggest that Libertarians who endorse PAP must also be committed to presentism. I put the issue to the side here, however.

It's important to point out that whether or not one endorses this position may depend on other commitments that one has, especially when it comes to the metaphysics of time. For instance, if one holds to an eternalist theory of time, whereby the universe's complete history exists all at once, then one may think that there can be true facts about future free actions.⁵² Nonetheless, two points are worth noting.

First, insofar as one *is* attracted to a form of presentism, and also attracted to the position that the Frankfurt intervener cannot know what an agent will freely do, the claim that future facts about free actions *cannot* be true provides a plausible explanation for why the Frankfurt intervener cannot know.⁵³ Second, even if one does adopt a form of eternalism, one may think that there is a relevant asymmetry between how future facts about free actions could be true and how the counterfactuals of free actions could be true. Part of the allure for thinking that future facts about free actions could be true is presumably that *one of the facts will be made true* at some point in the future. However, we cannot say the same thing about counterfactuals regarding free actions. So, even the libertarian who is an eternalist may have some reason to sympathize with the claim that *counterfactuals* regarding free actions cannot be true.

To briefly sum up, I have looked at two different arguments involving libertarianism. In both cases, I drew out claims made by libertarians that support the position that counterfactuals involving what free agents would do are not true. Although not all libertarians will accept this conclusion, I believe that many will (or, at least, many will find it plausible). The problem with this is that reasons-responsiveness requires that these counterfactuals be possibly true. Without them, we cannot determine

⁵² If one believes that facts about future free actions can be true, that does not mean that one must also be committed to counterfactuals about free actions being true. One reason may be that future contingents have truth makers, but the counterfactuals don't.

⁵³ Some maintain that the reason why the Frankfurt intervener cannot know beforehand is because there is insufficient *evidence* to determine what the agent will do. There may be a true fact about what she will do, but it's epistemically inaccessible to anyone. I'm not convinced that a Frankfurt intervener couldn't know beforehand what the agent will freely do, but, of course, others disagree.

whether an agent is reasons-responsive, which entails that we cannot determine whether an agent is morally responsible.

§5.4.2 Indeterminacy Itself

In the previous sub-section, we looked at why *libertarians* should deny that counterfactuals involving free actions can be true. Now I want to examine considerations that have been put forth to support the position that (most) indeterministic counterfactuals are not true.⁵⁴ While the claim is broader, it nonetheless supports the same conclusion since counterfactuals about libertarian free actions are a subclass of the all the indeterministic counterfactuals. If we have reason to think that all indeterministic counterfactuals are not true, we have reason to think that counterfactuals involving libertarian free will are not true.

We should start by noting that nothing that was said above to motivate the claim that counterfactuals involving free actions are not true depended on the counterfactuals involving *free actions* in particular. Rather, what seemed to be doing the work is that that counterfactuals concern indeterministic states of affairs (cf. van Inwagen 1997: 231). We should not be surprised to see, then, that we can reach much the same conclusions simply by focusing on indeterminism.

A number of authors who have thought about indeterministic counterfactuals have come to the conclusion that they cannot be true. For instance, Maudlin (2007) develops a counterfactual semantics under which indeterministic counterfactuals cannot be true. While these accounts are worthy of discussion, for space limitations, I focus primarily on recent work from Hájek, who has perhaps done the most to develop and defend this position.

Hájek (2021, 2022, MS) maintains that all indeterministic counterfactuals are false. So, while we have so far stayed neutral between whether they are false or indeterminate, here we will look at

⁵⁴ Counterfactuals with probabilistic consequents may be an exception. More on this below.

considerations for thinking that they are false in particular. Hájek has put forth a number of arguments and considerations for this position, but here we will focus on two of his most prominent arguments.

The first argument centers around the thought that “chance undermines ‘would’,” and the second argument centers around the idea that “might nots undermine ‘woulds’.” The two arguments, though tightly connect, are conceptually distinct. Following Kocurek (2022), I refer to former as the "direct argument" and the latter as the "indirect argument."

The Direct Argument. The direct argument proceeds by trying to draw out an intuitive conflict between *chance* and *would*. When we reflect on what is involved in would claims, we should conclude that they cannot be true. In short, chance undermines would.

The argument is simple straightforward. Let's suppose that I have a coin in front of me, and let's further stipulate that coin flips are irreducibly chancy processes. Now consider the two claims: (i) tossing the coin is a chancy process; and, (ii) if I were to toss this coin, it *would* land heads. In light of this, we're asked to consider the following question: can both of these claims be true at the same time? Hájek maintains that they cannot.

Well, why not? The thought is supposed to be that, when we draw out what each claim involves, a conflict emerges. The basic idea is that, intuitively, chancy processes by their very nature *leave it open* which outcome occurs. If the outcome of an indeterministic process is left open, then we can't say how it would resolve (Hájek MS: 8). Hájek approvingly quotes Richard Jefferey on this matter: If we imagine an unflipped coin, “there is no telling whether the coin would have landed head up on a toss that never takes place. That’s what probability is all about.”

You can also think about it from the other side. If you believe there is something that determines how a chancy process would be resolved, then *what* is that thing that settles it? *What* determines beforehand what the outcome of a chancy process is? Whatever it could be, it would appear to be something entirely mysterious.

In short, when we think about indeterministic counterfactuals, we can see that the chancy element “second-guesses” the determinateness of the counterfactual (Hájek MS: 8). Since chance claims and would claims are in conflict, we have to give up one of them. Since we're presupposing that indeterminism is true, chance claims must stay, and would claims must go.

The Indirect Argument. While the direct argument focuses on a conflict between chance and would, the indirect argument focuses on a conflict between *might-not* and *would*. DeRose (1999) has labelled this the “inescapable clash.” Similar to the direct argument, we're asked to consider two claims. In this case, these are: (i) if I were to toss the coin, it *might not* land heads; and (ii) if I were to toss this coin, it *would* land heads. With these claims in mind, Hájek then provides two distinct lines of reasoning by which we might arrive at the inescapable clash.

Route 1. First, you can motivate the clash by drawing on David Lewis' counterfactual semantics, since the clash falls directly out of it. This is because Lewis identifies *might* and *would* as duals of another. He calls this principle *Duality*:

(Duality) $x \rightarrow y \equiv \sim(x \diamond \rightarrow \sim y)$ (Or, in plain english: “If x were to occur, then y would occur” is equivalent to “it is false that, if x were to occur, then y might not occur.”)

There is a short road from *Duality* to the inescapable clash. To draw this out, let's use our coin flip example. That is, let x refer to flipping a coin and let y refer to the coin landing heads:

(Would) If I were to flip the coin, then it *would* land heads ($x \rightarrow y$)

(Might Not) If I were to flip the coin, then it *might not* land heads ($x \diamond \rightarrow \sim y$)

According to *Duality*, these two claims can't both be true at the same time. This is because *Duality* and *Would* entail the falsity of *Might Not*. *If I were to flip the coin, then it would land heads* entails (from

Duality) *It is false that, if I were to flip the coin, then it might not land heads.* And this is just the denial of *Might Not*.

Or, to make it more pertinent to our overarching discussion. Let's consider an example involving libertarian free will. Let's suppose *if Pablo's daughter were present, he would freely raise his arm* is true. Given *Duality*, it follows that, *it is false that, if Pablo's daughter were present, he might not freely raise his arm*. The problem is that surely the latter counterfactual is true. Given that Pablo has libertarian free will, surely it is true that he *might not* freely raise his arm in the presence of his daughter. Yet, if we endorse this might-not counterfactual, we need to deny the corresponding would counterfactual. That is, we should be committed to saying that, *it is false that, if Pablo's daughter were present, he would freely raise his arm*. And this is just to deny that the counterfactual about what Pablo would freely do is true.

Although *Duality* was first put forth by Lewis, it worth noting that it has been accepted many others as well (e.g., Bigelow and Pargetter 1990; Bennett 2003; Hawthorne 2005; Williams 2008). This is not to say that one must endorse *Duality*—nor have I claimed to defend it. Rather, it is point out that many others have found it plausible as well.

Route 2. The second way to motivate the inescapable clash revolves around the infelicity of uttering sentences such as:

If the coin were tossed, the coin *would* land heads, but it also *might not* land heads .

To many, sentences such as just plain wrong. (Even DeRose (1999) admits as much, which leads him to provide a contextualist solution to the apparent conflict.) One explanation for why this kind of sentence sounds infelicitous is that it is—the would-claim and the might-not-claim contradict one another. When we say that the coin *would* land heads, we seem to be securing that outcome, and closing off other possibilities in the process. Yet, when we allow that it *might not* land heads, we seem to be leaving open the possibility that it would not land heads.

Since we can't keep both kinds of counterfactuals, and since we're presupposing indeterminism is true, would counterfactuals have to go. This leads us back to the conclusion that counterfactuals about what an agent would freely do are false.

§5.5 Probabilities to the Rescue?

Although we have reason to doubt that there can be true counterfactuals about what an agent would in fact do, Hájek and others⁵⁵ have left it open that there may be true counterfactuals about what an *would probably* do. For example, it may be true that:

If Pablo's daughter were present, then he would *probably* raise his arm [e.g., raise his arm with .7 probability].

One reason for thinking that these kinds of counterfactuals may be true is if the laws of nature specify a unique conditional probability in all situations. Since we customarily hold fixed the laws of nature in our counterfactual scenarios, it would appear to follow that there may be a uniform way of settling what the probability of the consequent occurring is. There would be clear truth conditions for these kinds of probabilistic counterfactuals. While I am skeptical that the laws of nature specify such unique conditional probabilities (as is Hájek 2003), others are less skeptical.

At this point, then, one might propose that we use such probabilistic counterfactuals to ground reasons-responsiveness. The idea here would be that, instead of requiring that the agent would in fact act a certain way in a suitable proportion of possible worlds, we could instead require that they would *probably* act a certain way in a suitable proportion of possible worlds.

⁵⁵ See Adams (1977) and Van Inwagen (1997).

I am unsure about what to make of this view. On the one hand, I can see the initial pull for thinking that these counterfactuals may help provide to *some evidence* that the agent is reasons-responsive. But, on the other hand, when we dig deeper, it's unclear to me how to use probabilities in our assessments of responsibility.

To make the case that these probabilities are suitable for playing the role of grounding reasons-responsiveness, consider a variation of thermostat case. Suppose a number of counterfactuals of the following sort are true: if the thermostat were set to 71°F [72°F, 73°F...], then the room would *probably* be 71°F [72°F, 73°F...]. In this case, would we want to say that the thermostat is reliable at regulating the temperature of the room? Our assessment may depend in part on how the probabilities fluctuate across these counterfactual scenarios. If, in some cases, the probabilities are higher, and in others cases the probabilities are lower, it may be hard to get a grip on its overall reliability. Assuming, however, that the probabilities remain at a certain threshold (90% reliability, for example), it would not be implausible to say that the thermostat is reliable. In which case, why not say the same thing about the agent? That is, why not say we can use similar kinds of probabilistic counterfactuals to determine whether an agent is reasons-responsive.

The situation appears much less clear when we turn to assessments of responsibility and probabilities. In the context of considering whether probabilities can play the role of accounting for the scalar nature of responsibility (that people can be more or less responsible), Nelkin (2016) makes the case that intuitions regarding responsibility and probability are quite easily pulled apart. I quote her at length:

To see that probability is ill suited to this task, consider cases that are identical, except for differences in probabilities. Take the case of Drew...Drew takes an extra drink of whiskey before getting in the car to drive home. Suppose the odds that she takes the drink after thinking it over are 50–50, and that she meets the other relevant conditions on responsibility. Presumably, she is blameworthy on most incompatibilist views. But

now, keeping other things fixed, suppose that the odds are 99–1 that she takes the drink. Does this make her more or less blameworthy? One might think that it makes her less blameworthy because it is closer to her being determined to do so. But then should we also say that she is more blameworthy if the odds of her taking the drink are reversed, and only 1–99? This seems counterintuitive. But then how do probabilities affect blameworthiness? Consistent with these judgments is the claim that actions are freer—and thus one is more blameworthy for bad actions—the closer the odds of so acting are to 50–50. But this sort of view faces the challenge that many paradigm free and responsible actions do not seem to be of this sort, and, further, that some that are very likely or very unlikely, are ones for which agents might be blameworthy or praiseworthy to a high degree.

While the concern here is primarily with the degrees of responsibility, there is nonetheless a more basic insight that is relevant here. That insight is that we can easily imagine cases where an agent is morally responsible regardless of what we take the probabilities (that she will act) to be—whether they are low, split, or high (e.g., 1%, 50%, or 99%). But, now, since there is no clear probability *threshold* at which we would be willing to hold an agent blameworthy or not, it's not clear that we should think that there is some probability threshold at which we would count an agent as reasons-responsive (which might in turn ground her being morally responsible).

§5.6 Taking Stock

I have sought to motivate the position that counterfactuals about what a libertarian free agent would do are not true. I have done do this by showing that many libertarians have endorsed positions that either commit themselves to it, or at least make the position plausible. In addition to this, I have

also considered arguments that support the more general position that (nearly) all indeterministic would counterfactuals are false.

It is worth making it explicit that I have not attempted to defend any of these positions myself. I have merely tried to draw attention to a number of different lines of reasoning that many have found plausible, which support the conclusion that important kinds counterfactuals involving free actions are not true. In laying them out, I hope that readers can see the intuitive pull of these cases, especially when taken in their totality, but obviously much more could be said. Though I myself am persuaded by many of these considerations, not everyone who has thought about these matters have been, and some who have not have offered their own solutions to the foregoing puzzles.

Insofar as the agent-causalist takes this challenge seriously, however, they are forced to re-examine what it means to be reasons-responsive. They cannot simply adopt the resources afforded by the standard reasons-responsive accounts, because These reasons-responsive accounts require that certain counterfactuals involving free agents be *true*, if an agent is to be morally responsible.

§5.7 The Scope of the Problem

From the foregoing remarks, one conclusion might be to say so much the worse for agent causation. Reasons-responsive accounts appeared to work just fine on their own. It was only after trying to merge with agent causation that things went sideways. Tempting though it may be, I believe that this conclusion would be premature. In fact, I want to make the case that the problem can generalize to just about any reasons-responsive account that relies on counterfactuals. If this is correct, then we all have reason to rethink what reasons-responsiveness consists in.

In the previous section, we saw that the tension between reasons-responsiveness and agent causation arose from (i) the commitment to indeterminism and (ii) the commitment to certain counterfactual requirements. The tension, then, is not between agent causation and reasons-responsiveness *per se*, but between any account that endorses indeterminism and those counterfactual

requirements. This would include any libertarian view—even event-causal libertarians. Event-causalists agree with agent-causalists on the matter of indeterminism, and they differ from agent-causalists primarily in that they maintain that free actions are caused by certain agent-involving mental events. Insofar as event-causalists are libertarian, they too are going to have to come to grips with the problem of indeterministic counterfactuals.

In fact, the problem extends beyond libertarianism to compatibilism as well. This may be surprising at first, since compatibilists maintain that freedom and responsibility are compatible with determinism, and so don't seem to need to reckon with indeterministic counterfactuals. However, to see that not all compatibilists are immune, we can divide compatibilists into at least two camps. Those who maintain that freedom and responsibility *require* the truth of determinism, and those who do not. Call those in the former camp "strong compatibilists," and those in the latter camp "supercompatibilists." Because supercompatibilists don't require determinism, they must allow indeterminism. Thus, they endorse the position that freedom and responsibility are compatible with both determinism and indeterminism (Vargas 2012).

Most compatibilists are supercompatibilists. One reason for this is that they don't want their account of freedom and responsibility to turn on open scientific questions, like whether the world is fundamentally deterministic or indeterministic. Whatever the reason for their position, supercompatibilists are going to have a difficult time accepting reasons-responsive accounts grounded in counterfactuals. This is because, if the above considerations are correct, we can't be free and responsible if indeterminism is true (assuming the reasons-responsive account). This means that supercompatibilists face a dilemma: either give up *supercompatibilism* and require that freedom requires determinism, or else give up on reasons-responsive accounts that rely on counterfactuals to ground responsibility. I leave it open for the supercompatibilist (reasons-responsive theorist) to decide which horn to take.

§5.8 Where Do We Go From Here?

Let's recap where we're at. So far, I have identified a gap for agent-causalists. The gap is that standard agent-causal accounts can't fully account for various responsibility relevant features, such as when Bella struggles to conform with her will. I sought to fill this gap by supplementing an agent-causal account with elements from a prominent reasons-responsive account. The hope was that, since reasons-responsive accounts appeared to fare better when it comes to dealing with these responsibility-relevant features, agent-causalists could simply borrow the requisite resources from these accounts. However, upon examination, it appeared that this integration was not viable; the indeterministic counterfactual conditionals that would be required to make an agent reasons-responsive cannot be satisfied. Not only that, but it was also argued that this concern threatens nearly all reasons-responsive accounts that rely on counterfactual conditionals to ground moral responsibility.

There persists a gap in the agent-causalist account that needs to be filled. While I argued that a prominent account of reasons-responsiveness is problematic, I remain persuaded that the guiding idea behind reasons-responsive accounts is largely correct. There is a kind of sensitivity to reasons that goes beyond merely acting for reasons that is essential to moral responsibility. The question we want answered is *what does this kind of reasons-responsiveness consist in?*

At first glance, an actual-sequence account of reason-responsiveness presents itself as an attractive alternative for integration, such as the one developed by Sartorio (2016). An actual-sequence view, such as Sartorio's, maintains that freedom and responsibility are wholly grounded in the *actual* causal sequences. Not only this, but the proposal is that we can also get reasons-responsiveness out of an actual-sequence view. The account is interesting and sophisticated, and I do not have the space here to do it justice, but there are good reasons to think that it too will fall prey to a version of the indeterministic counterfactual concern.

Sartorio's actual-sequence view involves positing the existence of "absence causes," where an absence simply refers to the absence of a certain event. For example, right now there is no alien

invasion on earth; there is an *absence* of an alien invasion on earth. To say that there is an absence cause, then, is just to say that an absence can play a causal role in bringing about effects. Most notably, for Sartorio's account, the idea is that the absence of certain reasons can play a causal role in why an agent acts as she does, which in turn can help to ground an agent's reasons-responsiveness.

Putting to the side concerns about the existence of absence causes, Sartorio invokes certain counterfactual constraints on causation. Although she does not provide a reductive counterfactual analysis of causation, she maintains that causes must be difference-makers in the following sense:

Causes as Difference-Makers: If C caused E, then, had C not occurred, the absence of C wouldn't have caused E. (Sartorio 2005: 75; 2016: 94)

The problem is that the latter half of the principle—if the absence of C were to occur, then it would not cause E—invokes a counterfactual. And, if indeterministic counterfactuals are never true, then this causes as difference-makers principle will be empty. It will never be true that, if the absence of C were to occur, then it would not cause E. This means that the principle can never be satisfied. So, while the actual-sequence is initially appealing, because it relies on counterfactuals to ground its actual-sequence claims, it is not likely to be helpful here.

What is needed is a way of capturing reason-responsive in a manner that is thoroughly non-modal. Various authors have recently made the case that reasons-responsive accounts need to make room for ecological (or extrinsic) pertaining to an agent's acts (e.g., Brink and Nelkin 2013; Nelkin and Vargas Forthcoming). Here, one interesting proposal is the fair opportunity model for moral responsibility (Brink and Nelkin 2013). On this model, responsibility is intimately connected to how reasonable it is to expect someone to perform an action. This view says that "what *makes* it reasonable to make particular demands of agents in particular situations is the quality of the relevant opportunities, determined by the agents' capacities and the features of their situations" (Nelkin 2016: 371). Supposing that the reasonableness of a demand can be assessed without making reference to any such modal

features, an agent-causalist may do well to take this proposal seriously. However, whether such a model need not rely on any modal features remains for us an open question; and whether there are other such accounts that may be of help to the agent-causalist deserves further consideration.

In closing, I want to briefly mention a few key takeaways before gesturing at some potential ways forward. First, some remarks about what an agent-causalist can learn from the reasons-responsive theorists we've examined here. One key takeaway is that the agent-causalist would benefit from expanding her conception of what *capacities* are relevant to moral responsibility. As things are typically conceived of by agent-causalists, the primary responsibility-relevant capacity is a causal capacity to be directly causally involved in her actions. Our discussion has shown that this is too impoverished to do the work of making sense of our moral responsibility assessments.

One capacity that deserves more attention from the agent-causalist is the recognitional capacity. Being morally responsible requires a suitably robust capacity to recognize certain sorts of reasons, especially moral ones, when acting. Plausibly, it is not enough that an agent simply (be able to) act on some reason, even if it's moral one. The agent's understanding of her action should be suitably rich in very much the same spirit that F&R describe. Thankfully, it seems to me that the recognitional capacity does not obviously require any counterfactual commitments, so I am optimistic that there could be a straightforward adoption of resources developed elsewhere (cf. Brink and Nelkin 2013). Of course, an agent-causalist may also develop her own account of the relevant recognitional capacity if she so wishes.

In addition to this, the agent-causalist needs to enrich her understanding of what is involved in the relevant volitional capacity. The bare capacity to directly cause an action for a reason is not sufficient. Our capacities can be diminished, but our capacity to directly cause an action, it seems, cannot. Whereas the former implies a kind of scalarity, the agent-causalist picture requires that it be binary.

In considering the cases of Bella and Gael, I am tempted to think that they highlight two different ways that our reasons-responsive capacities can be affected, which in turn suggests two different responses. In the case Bella, doing the right thing requires a great deal more *effort* than is typically expected for a person in her situation; and in the case of Gael, the reasons that should be motivating him don't appear to have the right kind of *pull* or *motivational force* that is typically expected for a person in his situation.

Start with the case of Bella. The thought here is that acting on the right reason for her requires much more effort than is typical. Though it seems to me that philosophers have not generally paid special attention to the notion of effort, Michael Brent (2017, 2022) has recently begun to develop resources that may be useful here. As Brent notes, effort is pervasive in our lives.⁵⁶ We exhibit effort in our physical activities of running a marathon and lifting heavy objects (Brent 2017, 661), as well in our mental activities like focusing our attention, trying to visually imagine various things, and practicing mindfulness and meditation (Brent 2022, 104–105). According to Brent, all activities involve an agent's causing and causally sustaining an action; and causing and causally sustaining actions requires effort (however minimal). Before saying how this might be applied to Bella's case, it is worth noting how our moral responsibility assessments appear to track effort in non-trivial ways.

To use a simple example, suppose that, in order to save someone's life, someone is told to hold a 70 lb. box in the air for an hour. It strikes me as plausible that whether a person is morally responsible, and the degree to which the person morally responsible, is at least partly determined by the amount of effort required to perform that action. Imagine that Aaron is a person of average strength who only rarely goes to the gym, and for which it would require a great deal of effort to sustain the box in the air for that amount of time. I could imagine things playing out in a number of different ways, each with varying assessments of responsibility. But for the moment let's imagine Aaron trying to hold the box,

⁵⁶ Uses of the term effort might be many and varied, and these might come apart in important ways. This would then just be capturing one important aspect, or kind, of effort.

yet getting so fatigued at the 55-minute mark that he gives up. In keeping in line with our libertarian commitments, let's also suppose that it was actually possible for Aaron to hold the box for the whole hour. I can see at least two plausible responses to this situation. The first response is that, even though it was possible for Aaron to complete the task, given the amount of effort it would take to finish the task, he's not morally blameworthy. An alternative response would be that Aaron is morally blameworthy for failing to hold the box until completion, but because it required so much effort for him to do so, his blameworthiness is diminished.

Now contrast Aaron with Amanda who is a world champion weight-lifter. Amanda routinely lifts weights much heavier than 70 lbs. Holding a 70 lb. box for an hour might cause Amanda some discomfort, but it would not require a great deal of effort. If Amanda were to give up at the 55-minute mark, I think that she would be fully morally blameworthy for doing so *because* it would require so little effort for her to do so. If this is right, then one's moral responsibility status is at least partly a function of the amount of effort required to perform an action.⁵⁷

Applying this to the case of Bella, we might explain the fact she is not (fully) morally blameworthy on the same grounds. Conforming her actions to her considered judgement requires effort in much the same way as performing a strenuous physical activity does. Phenomenologically, I think this tracks. Many of us have experienced this on some level, where internal psychological factors make it so that performing activities, even important ones, require a great deal more effort to perform than is typical.

Finally, consider the case of Gael, where I have much less to say. I will only note that there is one way of reading his case where Gael lacks a (or has a diminished) capacity to care and take seriously the relevant reasons. On one way of cashing this out, this diminished capacity to care collapses into a

⁵⁷ I do not mean to suggest that all degrees of effort result in a different degree of moral responsibility. It strikes me as less plausible, for instance, that someone who is only slightly weaker than Amanda that also abruptly gives up at 55 minutes is less blameworthy than Amanda solely because it took them slightly more effort.

diminished recognitional capacity. Gael simply lacks the capacity to recognize the relevant reasons. On an alternative way of cashing things out, Gael has the capacity to recognize the relevant reasons, but he is not *moved by* those reasons. If the latter is the case, this would suggest that there is an additional capacity to be moved and motivated by the relevant reasons that needs further elucidation.

There may be yet other cases that suggest further treatments and revisions about what has been sketched here. Still, I hope to have at least made a case for why this matter needs further treatment, especially for those agent-causal theorists.

Chapter 6

Conclusion

§6.1 Remaining Questions

The central task of this dissertation has been to defend and develop a more plausible form of agent causation. Still, there remain a number of important questions that deserve further treatment which have been unaddressed here. In closing, I would like to outline those issues that seem to me to call to most urgently call for attention.

The first matter concerns the topic of reductionism and agency. What emerged from our discussion in chapter 2 was that agent causation and its relationship to non-reductionism needs reevaluation. The standard story has been that non-reductionism is tied up with the idea of direct substance causation, where the only way to deny agency reductionism is to deny that the causal role of agents is wholly explainable in terms of event causation. On the causal pluralist proposal, what is central to agency non-reduction is not the idea of substance causation; non-reductionism is compatible with event causation so long as the type of events are at the level of the agent. Rather, what is important is that agency not be decomposable to the parts of the agent. Conceivably, one could be an agency reductionist while embracing the view that all causation is substance causation. For example, suppose that fundamental particles are substances, and that all causation is ultimately reducible to the causation occurring at this level (cf. Bird 2016). One would be embracing substance causation while also embracing agency reductionism.

More work needs to be done in order to better make sense of agency reductionism. Agents evidently cannot be reduced to their components, but more should be said about what this exactly amounts to. On this score, metaphysical emergence has begun to receive renewed interest, which in turn has led to the development of richer conceptual tools for thinking through emergence-related questions. The basic idea behind emergence is that there are certain features that a system has as a

whole that are not fully explainable in terms of the features of the system's parts. Some have maintained that agents are emergent entities, while others have advocated for powers (or properties) being emergent. Both of these options seem to me to be appealing prospects for further elucidating the notion of agency non-reductionism (cf. O'Connor 2000 and Wilson 2021).

It is perhaps worth mentioning alternatives that have not received as much attention. In the spirit of neo-Aristotelianism, theories of hylomorphism have been offered as non-reductive accounts of things like agents. Theories of hylomorphism can vary a great deal (Shields 2022), but the basic idea is that all objects are composed of form and matter, which in turn provides a way of characterizing the individuality of particular objects like persons. Or perhaps yet another alternative is the idea that agents are simply prior to their parts. Agents (or their powers) don't emerge from anything, neither is their individuality as an agent explained by something like form and matter, but their existence is simply given a kind of priority over the component parts (cf. Schaffer 2010).

At first glance, one might think that these options amount to bookkeeping within the agent-causalist camp. And while this might turn out to be true, we won't know for certain until we tackle the matter head on. It may turn out, for example, that there are potential costs and benefits or each these different views, especially in the context of free will. Even more, given that we have reasons to rethink the standard picture of agency non-reductionism—one that equates non-reduction to substance causation—it could very well turn out that further complexities arise in considering how each of these frameworks account for a non-reductionist picture of agency.

The second matter is to further clarify how agent causation fits into the empirical picture of the world canvassed in chapter 3. The picture that emerged there, following the lead of scientific pluralists in the philosophy of science literature, was that the different sciences are not wholly reductionist, and the sciences as a whole do not appear to form a perfectly systematic and interconnected disciplinary web as many tend to think. While this picture leaves room for free will, it also leaves a great deal open about how free will fits into this picture of the world. Does, for example,

the emergentist view just sketched above neatly fit into this picture, or does it cut against it? And are there alternative stories to tell that could be developed?

There are additional empirical concerns that ought to be addressed as well. Here are two such related concerns. The first concern is that our practices of holding people morally responsible are costly. Our assessments of moral responsibility are thought to justify all sorts of behavior, such as the appropriateness of praising, blaming, holding accountable, forgiving, punishing, and so on. Given the costliness of these practices, there is a high evidential bar for our theories of moral responsibility. If we are going to hold people morally responsible *on the basis of* something like agent causation, then we need really good evidence that it's true. But, the thought goes, views like agent causation don't meet that evidential bar, so we should give it up.

A related concern is that we should want our theories of freedom and responsibility to be immune to empirical findings. If, for example, one day scientists were to discover that the reductionist picture of the world were true, then we would have to give up all of our practices that depend on this non-reductionist picture of freedom and responsibility. Because this seems like a live possibility, we should not be willing to hang so much on the balance of this non-reductionist view of free will.

The final matter that deserves treatment is to better understand the relationship between freedom and external, especially social, influences. Agent-causal views have often (over)stressed the individuality of human freedom. Just consider this well-known description of agent causation which says that "we have a prerogative which some would attribute only to God: each of us, when we act, is a prime mover unmoved" (Chisholm 1958, 32). The picture painted of the person is one where she is completely stoic and unaffected by any external influences. To be sure, agent-causalists have not wanted to deny that agents can be affected by external influences. I think that it is telling that even Chisholm's seminal defense of agent causation—which, as just sketched, stresses the individuality of human freedom—ends by trying to elucidate the claim that there can be "inclination without necessitation." Still, very little has been said about the way in which agents can be influenced by

external factors on this agent-causal view, and even less has been said about how our agency is socially structured and scaffolded. If agent causation is to broaden its appeal, these matters need to be treated more earnestly and with more care.

§6.2 Final Remarks

The history of agent causation is a turbulent one. There is a case to be made that, for much of the history of philosophy, philosophers widely embraced some form of agent causation. Sometime during the early modern period, however, the view fell out of disfavor and has never quite recovered. Indeed, not too long ago the picture of agent causation was openly mocked in the philosophy literature. But things are looking up for the view. Objections and concerns that were typically taken as obviously true without any serious argumentation are now treated with more seriousness. In addition to this, agent-causal theorists are continuing to provide ever-more sophisticated treatments of the view, helping to lend the view more plausibility.

Amidst all of these exciting developments, it's a curious fact that the view has not gained more support. One plausible explanation for this seems to me to be that agent-causal theorists have been predominantly concerned with abstract metaphysical concerns, like the nature of substance causation and causal powers, and less with matters that are central to our everyday cares and concerns. If agent-causalists want their view to receive more uptake, they must put their sophisticated resources to work to explain and illuminate these central issues. While I have not addressed these matters here in the comprehensive manner that they deserve, I hope to have at least shown that these are issues that are worth taking seriously, and to have gestured at a better way for agent-causalists to think about them.

References

- Adams, R. (1977). "Middle Knowledge and the Problem of Evil," *American Philosophical Quarterly* 14 (2): 109-117.
- Alvarez, M., & Hyman, J. (1998). "Agents and Their Actions," *Philosophy* 73 (2): 219-245.
- Alvarez, M. (Winter 2017) "Reasons for Action: Justification, Motivation, Explanation," *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/>.
- Amaya, S. (2017). "Basic Actions Reloaded," *Philosophy Compass* 12 (9): e12435.
- Armstrong, D. M. (1983), *What is a Law of Nature?*, Cambridge University Press.
- Armstrong, D. M. (2005). "Four disputes about properties," *Synthese*, 144 (3): 309–320.
- Baker, J. (2017), "Rejecting Pereboom's Empirical Objection to Agent Causation," *Synthese* 194 (8): 3085–3100.
- Balaguer, M. (2010). *Free Will as an Open Scientific Problem*. MIT Press, Bradford.
- Bellay, T., Klaus, A., Seshadri, S., & Plenz, D. (2015). "Irregular Spiking of Pyramidal Neurons Organizes as Scale-invariant Neuronal Avalanches in the Awake State," *Elife* 4: e07224.
- Bennett, J. (1988). *Events and Their Names*. Hackett.
- Bernstein, S., & Wilson, J. M. (2016). "Free Will and Mental Causation," *Journal of the American Philosophical Association*, 2(2): 310–331.
- Bird, A. (2007). *Nature's Metaphysics: Laws and Properties*. Oxford University Press.
- Bird, A. (2016) Overpowering: How the Powers Ontology Has Overreached Itself. *Mind* 125 (498): 341-383.
- Bishop, J. (1983). "Agent Causation," *Mind*, 92(1): 61–79.
- Boghossian, P. (2014). "What is inference?," *Philosophical Studies* 169 (1): 1–18.
- Brent, M. (2017). "Agent Causation as a Solution to the Problem of Action." *Canadian Journal of Philosophy* 47 (5): 656–673.
- Brent, M. (2022) "Mental Action and the Power of Effort" in *Mental Action and the Conscious Mind*, 100–121.
- Brink, D. O. (2021) *Fair Opportunity and Responsibility*. Oxford University Press.
- Brink, D. O. & Nelkin, D. K. (2013). "Fairness and the Architecture of Responsibility," *Oxford Studies in Agency and Responsibility* 1: 284–313.

- Buchak, L. (2013). "Free Acts and Chance: Why The Rollback Argument Fails," *Philosophical Quarterly* 63 (250): 20–28.
- Buckareff, A. A. (2011). "How Does Agent-causal Power Work?," *Modern Schoolman*, 88 (1): 105–121.
- Buckareff, A. A. (2017). "A Critique of Substance Causation," *Philosophia*, 45: 1019–1026.
- Cartwright, N. (1999), *The Dappled World: A Study of the Boundaries of Science*, Cambridge University Press.
- Cartwright, N. (2016). "The Dethronement of Laws in Science," In *Rethinking Order: After the Laws of Nature*, 25-51.
- Casati, R., & Varzi, A. C. (2023). "Events" in E. N. Zalta (ed). The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/fall2023/entries/events/>
- Cat, J. (2022, Spring) "The Unity of Science" in Zalta, E. N. (ed). The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/archives/spr2022/entries/scientific-unity/>
- Chang, H. (2012). *Is water H2O?: Evidence, Realism and Pluralism* (Vol. 293). Springer Science & Business Media.
- Chang, H. (2015). "Reductionism and the Relation Between Chemistry and Physics," in Arabatzis T., Renn, J., & Simões A. (eds.) *Relocating the History of Science* (312) Springer: 193–209.
- Chisholm, Roderick (1982). "Human Freedom and the Self." In Gary Watson (ed.), *Free Will*. New York: Oxford University Press.
- Chisholm, Roderick (1970). "Events and Propositions." *Noûs* 4 (1):15-24.
- Chisholm, R. (1976). *Person and Object: A Metaphysical Study*. Open Court.
- Chisholm, R. (1990). "Events Without Times an Essay on Ontology," *Noûs*, 24(3): 413–427.
- Chisholm, R. (1992). "The Basic Ontological Categories," In K. Mulligan (ed). *Language, Truth and Ontology*. Kluwer Academic Publishers: 1–13.
- Chisholm, R. (1994). "Ontologically Dependent Entities," *Philosophy and Phenomenological Research*, 54(3): 499–507.
- Clark, M. J., & Liggins, D. (2012). "Recent Work on Grounding," *Analysis Reviews*, 72(4): 812–823.
- Clarke, R. (1993). "Toward a Credible Agent-causal Account of Free Will," *Noûs* 27 (2): 191-203.
- Clarke, R. (2003), *Libertarian Accounts of Free Will*, Oxford University Press USA.
- Clarke, R. (2017), "Free Will, Agent Causation, and "Disappearing Agents"," *Noûs*: 76–96.

- Dainton, B. (2010). *Time and Space*. McGill-Queens University Press.
- Davidson, D. (1963). "Actions, Reasons, and Causes," *Journal of Philosophy* 60 (23): 685.
- Davidson, D. (1985). *Reply to Quine on Events*. Basil Blackwell.
- Demarest, H. (2017). Powerful Properties, Powerless Laws. In J. D. Jacobs (Ed.), *Causal Powers*: 38–53.
- DeRose, K. (1999). "Can It Be That It Would Have Been Even Though It Might Not Have Been?," *Philosophical Perspectives* 13: 385-413.
- Dupré, J. (1993), *The Disorder of Things: Metaphysical Foundations of the Disunity of Science*, Harvard University Press.
- Fischer, J.M., & Ravizza, M. (1999). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge University Press.
- Franklin, C. E. (2011). "Farewell to the Luck (and Mind) Argument," *Philosophical Studies* 156 (2): 199-230.
- Franklin, C. E. (2016). "If Anyone Should Be An Agent-causalist, Then Everyone Should Be An Agent-causalist," *Mind*, 125(500): 1101–1131.
- Franklin, C. E. (2018). *A Minimal Libertarianism: Free Will and the Promise of Reduction*. Oxford University Press.
- Franklin, C. E. (2019). "The Heart of Libertarianism: Fundamentality and the Will," *Social Philosophy and Policy*, 36(1): 72–92.
- Strawson, P. F. (1962). Freedom and Resentment. *Proceedings of the British Academy* 48: 187–211.
- Friedman, A. K., Weiss, K. R., & Cropper, E. C. (2015). "Specificity of Repetition Priming: The Role of Chemical Coding," *Journal of Neuroscience* (35): 6326–6334.
- Giere, R. (1999). *Science without laws*. Chicago: University of Chicago Press.
- Haji, I. (2008). *Moral Responsibility, Authenticity, and Education*. Routledge.
- Hájek, A. (2003). "What Conditional Probability Could Not Be," *Synthese* 137 (3):273-323.
- Hájek, A. (2020). "Contra Counterfactualism," *Synthese* 199 (1-2):181-210.
- Hájek, A. (2021). "Counterfactual Scepticism and Antecedent-contextualism," *Synthese* 199 (1-2): 637-659.
- Hájek, A. MS: "Most Counterfactuals Are False," Unpublished manuscript.
- Hettema, H. (2012). *Reducing Chemistry to Physics: Limits, Models, Consequences*. Createspace.

- Hofer, C. (2003). "For Fundamentalism," *Philosophy of Science* 70 (5): 1401-1412.
- Horst, Steven W. (2011). *Laws, Mind, and Free Will*. MIT Press.
- Hyman, J. (2015). *Action, Knowledge, and Will*. Oxford University Press.
- Jacobs, J. D., & O'Connor, T. (2013). "Agent Causation in a Neo-aristotelian Metaphysics," In S. C. Gibb & R. Ingthorsson (Eds.), *Mental Causation and Ontology*. Oxford University Press.
- Kane, R. (1996). *The Significance of Free Will*. Oxford University Press.
- Kelley, M. (2024). "How to Perform a Nonbasic Action." *Noûs* 58 (1): 106–125
- Kim, J. (1973). "Causation, Nomic Subsumption, and the Concept of Event," *Journal of Philosophy*, 70(8): 217–236.
- Klaes, C., Schneegans, S., Schöner, G., & Gail, A. (2012). "Sensorimotor Learning Biases Choice Behavior: A learning Neural Field Model for Decision Making," *PLoS Computational Biology*, (8): e1002774.
- Kocurek, A. (2022). "Does Chance Undermine Would?," *Mind*, 131(523): 747-787.
- Kuykendall, D. (2019). "Powerful Substances Because of Powerless Powers," *Journal of the American Philosophical Association* 5 (3): 339-356.
- Kuykendall, D. (2021). "Agent Causation, Realist Metaphysics of Powers, and the Reducibility Objection," *Philosophia*, 49(4): 1563–1581.
- Levy, N. (2011). *Hard luck: How luck undermines free will and moral responsibility*. Oxford University Press
- Lewis, D. (1973). *Counterfactuals*. Blackwell.
- Lewis, D. (1986). "Causation," In D. Lewis (ed). *Philosophical Papers Vol. II*: 158–172. Oxford University Press.
- Lombard, L. B. (1986). *Events: A Metaphysical Study*. Routledge.
- Lombardi, O. (2015). The Ontological Autonomy of the Chemical World: Facing the Criticisms. In *Philosophy of Chemistry*: 23–38. Springer, Dordrecht.
- Lowe, E. J. (2003). *Personal Agency*. Cambridge University Press.
- Ludwig, D., & Stéphanie R., (2021, Winter) "Scientific Pluralism," in E. N. Zalta. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), <https://plato.stanford.edu/archives/win2021/entries/scientific-pluralism/>
- Martin, E., & Bechtel, W. (2016). Multiple Orders in Biology. In *Rethinking Order: After the Laws of Nature*: 75–98.

- Markosian, N. (1998). "Brutal Composition," *Philosophical Studies* 92(3): 211–249.
- Markosian, N. (1999). "A Compatibilist Version of the theory of Agent Causation," *Pacific Philosophical Quarterly* 80(3): 257–277.
- Martinez, J. (2023). "Causal Pluralism: Agent Causation Without the Panicky Metaphysics," *Synthese*
- Maudlin, T. (2007). *The Metaphysics Within Physics*.
- McKenna, M. (2013) "Reasons-Responsiveness, Agents, and Mechanisms," *Oxford Studies in Agency and Responsibility* 1: 151–183.
- Merricks, T. (2001). *Objects and Persons*. New York: Oxford University Press.
- McLaughlin, B. & Bennett, K., (2018, Winter) "Supervenience", in E. N. Zalta (ed). The Stanford Encyclopedia of Philosophy, <https://plato.stanford.edu/archives/win2018/entries/supervenience/>
- Mele, A. R. (2006). *Free Will and Luck*. Oxford University Press.
- Mellor, D. H. (2004). *Probability: A Philosophical Introduction*. Routledge.
- Merricks, T. (2001). *Objects and Persons*. Oxford University Press.
- Montague, R. (1969). "On the Nature of Certain Philosophical Entities," *Monist* 53 (2): 159–194.
- Mumford, S. (1998). *Dispositions*. Clarendon Press.
- Mumford, S., & Anjum, R. L. (2011). *Getting Causes From Powers*. Oxford University Press.
- Nagel, E. (1961). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Harcourt, Brace & World.
- Nelkin, D. K. (2004). "Irrelevant Alternatives and Frankfurt Counterfactuals," *Philosophical Studies* 121 (1): 1–25.
- Nelkin, D. K. (2011). *Making Sense of Freedom and Responsibility*. Oxford University Press.
- Nelkin, D. K. (2016). "Difficulty and Degrees of Moral Praiseworthiness and Blameworthiness," *Noûs* 50 (2): 356–378.
- Nelkin, D. K., & Manuel V. (2023). "Responsibility and Reasons-Responsiveness," in T. Cyr, A. Law, & N. A. Tognazzini (eds). *Freedom, Responsibility, and Value: Essays in Honor of John Martin Fischer*. New York: Routledge.
- O'Connor, T. (2000). *Persons and Causes: The Metaphysics of Free Will*. Oxford University Press.
- O'Connor, T. (2003). "Review of Derk Pereboom, Living Without Free Will," *Philosophical Quarterly* 53 (210): 308–310.

- O'Connor, T. (2009). Degrees of freedom. *Philosophical Explorations* 12 (2): 119 – 125.
- O'Connor, T. (2011). "Agent-Causal Theories," in R. Kane (ed.), *The Oxford Handbook of Free Will: Second Edition*. New York: Oxford University Press: 309–328
- O'Connor, T. (2021). "Free Will in a Network of Interacting Causes," in *Neo-Aristotelian Metaphysics and the Theology of Nature*. Routledge: 151–168.
- O'Connor, T. & Churchill, J. R. (2004). "Reasons Explanation and Agent Control: In Search of an Integrated Account," *Philosophical Topics* 32 (1): 241–253.
- Paul, S. (2020). *Philosophy of Action: A Contemporary Introduction*. Routledge.
- Pereboom, D. (2001). *Living Without Free will*. Cambridge University Press.
- Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*, Oxford University Press.
- Pereboom, D. (2014). "The Disappearing Agent Objection to Event-causal Libertarianism," *Philosophical Studies* 1: 1–11.
- Rice, Rebekah L. H. (2011). "Agent Causation and Acting for Reasons," *American Philosophical Quarterly* 48 (4): 333–346.
- Rosenberg, A. (2008). *Darwinian Reductionism: Or, How to Stop Worrying and Love Molecular Biology*. University of Chicago Press.
- Runyan, J. (2017). Agent-causal Libertarianism, Statistical Neural Laws and Wild Coincidences, *Synthese* 195 (10): 4563–4580.
- Ryle, G. (1949). *The Concept of Mind: 60th Anniversary Edition*. Hutchinson & Co.
- Sartorio, C. (2016). *Causation and Free Will*. Oxford, United Kingdom: Oxford University Press UK.
- Quine, W. V. (1985). Events and Reification. In E. Lepore & B. McLaughlin (Eds.), *Actions and events: Perspectives on the Philosophy of Davidson*. Blackwell: 162–71.
- Sehon, S. (2005). *Teleological Realism: Mind, Agency, and Explanation*. Cambridge MA: Bradford Book/MIT Press.
- Sehon, S. (2016). *Free Will and Action Explanation: A Non-Causal, Compatibilist Account*. Oxford, United Kingdom: Oxford University Press UK.
- Schaffer, J. (2010). Monism: The Priority of the Whole. *Philosophical Review* 119 (1): 31-76.
- Shields, C. (2022). Hylomorphisms. *Ancient Philosophy Today* 4 (1): 96-127.
- Simons, P. (2003). "Events," in M. J. Loux & D. W. Zimmerman (eds). *The Oxford Handbook of Metaphysics*. Oxford University Press.

- Skow, B. (2018). *Causation, Explanation, and the Metaphysics of Aspect*. Oxford University Press.
- Speak, D. (2004). "Toward an Axiological Defense of Libertarianism," *Philosophical Topics* 32 (1): 353–369.
- Steward, H. (2012). *A metaphysics for Freedom*. Oxford University Press.
- Taylor, R. (1992). *Metaphysics*. Englewood Cliffs.
- Taggart, C. P. (2020). "A Response to the Problem of Wild Coincidences," *Synthese* 12: 1-15.
- Thomson, J. J. (2003). "Causation: Omissions," *Philosophy and Phenomenological Research*, 66 (1): 81–103.
- Todd, P. & Tognazzini, N. A. (2008). "A Problem for Guidance Control." *Philosophical Quarterly* 58 (233): 685–692.
- van Inwagen, P. (1997). "Against Middle Knowledge," *Midwest Studies in Philosophy* 21: 225–236.
- van Inwagen, P. (1990). *Material Beings*. Cornell University Press.
- van Inwagen, P. (2000). "Free will Remains a Mystery," *Philosophical Perspectives* 14:1–20.
- Vargas, M. (2012). "Why the Luck Problem Isn't," *Philosophical Issues* 22 (1): 419–436.
- Vargas, M. (2013). *Building Better Beings: A Theory of Moral Responsibility*. Oxford University Press.
- Velleman, J. D. (1992). "What Happens When Someone Acts?," *Mind*, 101 (403): 461–481.
- Vihvelin, K. (2013). *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. New York, New York: Oxford University Press USA.
- Wallace, R. Jay (1994). *Responsibility and the Moral Sentiments*. Cambridge, Mass.: Harvard University Press.
- Whittle, A. (2016). "A Defense of Substance Causation," *Journal of the American Philosophical Association* 1: 1–20.
- Widerker, D. (1995). "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities," *Philosophical Review* 104 (2): 247–61.
- Widerker, D. (2018). "In Defense of Non-causal Libertarianism," *American Philosophical Quarterly*, 55(1): 1–14.
- Wilson, G. M. (1980). *The Intentionality of Human Action*. Stanford University Press.
- Wilson, J. M. (2021). *Metaphysical Emergence*. Oxford University Press.
- Wolf, S. (1990). *Freedom Within Reason*. New York: Oxford University Press USA.