Powers, Manifestations, and their Essences

A.J. van Westen

3681092

Utrecht University

A thesis presented for the degree of Research Master of Arts in Philosophy
Defended on the 27th of November 2015

Supervisor: Albert Visser

Utrecht University

Second Supervisor: John Heil

Washington University in St. Louis

Universiteit Utrecht

Preface

It is a baffling insight of the physical sciences that the vast variety of things we see around us, and interact with in daily life are made up of constituents that come only in a small number of varieties, and that those constituents are themselves very much unlike the objects we know. The topic of this thesis is one of the most fundamental categories of our understanding: causation. On my view, causation is what happens when these particles that make up our universe *do* things. The constituent objects of the universe have powers to bring about certain effects when they are in the right conditions. This causal interaction between these fundamental entities is what explains all change, all stability, everything that happens. As such, powers and manifestations of these powers are essential to our understanding of causation.

It is a main tenet of my thesis that these powers are essentially powers towards certain effects. A power is a power to bring about a certain type of effect under certain circumstances. If it were to bring about some other effect, it would be a different power. This is why manifestations give the powers their essences; manifestations are essential to our understanding of powers. Ultimately, causation is a matter of powers, manifestations, and their essences.

In philosophy much progress is made by discussion and counterargument. Theories are criticized in journals, at conferences and within philosophy departments. A philosophical theory does not come to maturity in isolation. In this thesis therefore I will engage with other representatives of the type of view I am developing, as well as philosophers from outside this small area of philosophy. However, my thesis does not have the character of a literature study or a critical footnote to another theory. My attempt is to develop a metaphysics of powers that has certain qualities that I believe are valuable in philosophical theories. The theory has to have the philosophical merit of giving an insightful account of causation. I believe that the idea that powers are the basis of causation is not a trivial one. Also, I have tried to present a theory of causation in terms of powers in a convincing way. By showing how the theory deals with problems an examples I try to make it plausible that causal powers exist and can do the things we expect of causation. Furthermore, I aim for this theory to be in harmony with what we know about the world. Science is the source of some of our best knowledge about the world. Therefore I strive for my story to fit with science. Science is not a homogeneous whole and areas of science are sometimes disconnected in subject matter, method, and conclusions. My account is to accord with fundamental science, mostly physics.

In this thesis I have tried to strike a balance between doing analytic philosophy with the rigour that is afforded by semi-formal methods and readability. Most if not all arguments are made in running prose, and the definitions and theses that are written down in a more formal way serve to make clear what *precisely* is meant in the running text. This method is part of an attempt to write a thesis that is compactly written, interesting, and accessible to philosophers of various backgrounds.

On another stylistic note, most philosophy written in the digital age still takes all of its stylistic cues from the format of a printed document. This thesis is meant to be primarily read not on paper, but on the screen of a computer or tablet. The file you are reading now is a pdf (which stands for printed document format, ironically) that is enabled for certain features not available in printed documents. The main difference is the way in which cross-referencing is done in this thesis. Some terms, especially names of sections, equations, and principles, are formatted as hyper-links. These links show up as dark red text. Clicking on the words will take you to the relevant section in the thesis. On some applications for pdf-viewing, such as *Skim*, hovering the mouse over the link will give you a small preview window that shows the section or equation referenced.

Chapter 1 has as its goal to favorably incline the reader towards bearing with me as I construct a theory of causation in terms of powers. In order to do so the chapter contains discussion of some of the intuitive and methodological advantages that a powers theory of causation has over broadly Humean theories of causation. With neo-Humeanism as the accepted paradigm for causation, some discussion of its pitfalls is requisite in an alternative account.

Chapter 2 is where serious discussion of powers starts. In this chapter I present a picture of what powers can be. This is done by showing the relation between powers and properties. Also I present a standpoint on reductionism that informs my preference for physics as the paradigmatic science to engage with as a philosopher of causation. Furthermore Chapter 2 contains a presentation of a number of related problems for dispositional theories of causation. These supposed counterexamples to dispositional theories have received much attention in the literature. In order to solve the problems in Chapter 4 more metaphysics has to be done in Chapter 3.

The third chapter is the engine of the thesis: in it I spell out exactly what powers have to do with causation. I give a theory of causation that has powers as causes and manifestations as effects. The strength of my particular version of such a powers metapysics lies in its carefull formulation of what a power is, what a cause is and what an effect is. I use the example of gravitational attraction to show how my metaphysics applies to real-world causation.

In Chapter 4 more is said about the implications of the metaphysics presented in Chapters 2 and 3. I discuss how this powers theory can deal with the counterexamples presented in Chapter 2, how it deals with spontaneous or probabilistic effect, and what its relation to modality is. Also I discuss the logical properties of the causal relation, and the possibility of causation by absences as well as spontaneous effects.

Contents

Chapter 1

Why Powers?

Many philosophical theories are in one way or another informed by skepticism about their subject. Philosophers such as Descartes, Kant, and Hume have all used skeptical positions as a means to carve out the space in which theory can be constructed. For my subject, causation, Hume is the Great Skeptic. Most of the theories of causation that have been popular since the 20th century have been constructed in the space carved out by Hume's skepticism about causation.[1] I will forfeit this tradition and instead give a distinctly non-Humean account of causation. This is not because I have found an argument that counters skepticism, but rather to see how far theory can take us if we bracket skepticism.

In this first chapter I will say a thing or two about broadly neo-Humean theories of causation, in order to provide some contrast to my theory. However, the worth of the powers view will show itself mainly on its own terms. It is not in contrast with neo-Humean theories that we see what the powers view is worth, but rather when we simply develop it and see how it functions. There might not be a knock-down argument against an ontology without powers, and I believe that there is no definitive argument to the contrary either. The answer to the question "why powers?" therefore is one that is motivated by examples, considerations of theoretical aesthetics, and mostly on an evaluation of the ability of such an ontology to handle well-known and new philosophical problems. This is my method, as opposed to directly countering a non-powers account of the metaphysics of causation.

This chapter is much too short to do all those things at great length. The content is therefore restricted to a number of general and important points that have been the reasons for me to choose a metaphysics with powers over one without. I feel the obligation to give you, the reader, some arguments for choosing my position over another as a preliminary to the detailed account. However, the strongest advocate for my position will be the position itself, that I believe to have great plausibility when worked out in detail, as in the later chapters of this thesis.[2]

---

[1]Although I will argue later in this chapter that Hume's skepticism doesn't actually support the neo-Humean theories of causation.

[2]Of course, others have argued in favor of a non-Humean metaphysics of causation. Notable recent examples are Bird (2007), Heil (2012), Mumford and Anjum (2011). This eases the load this first chapter has to carry, since others have already given arguments in favor of a powers view. The roots of dispositionalism can be traced from Aristotle via Locke to 20th century authors like Mellor, Harrè and Madden, Shoemaker, and Martin (Mellor, 1974)(Harré and Madden, 1975)(Shoemaker, 1980)(Martin, 1994).

1.1 What are powers?

The term *powers* in the current debate has been used as synonymous with *dispositions*. I endorse a slightly different use of the term. I use the term *dispositions* for a wider category than *powers*. Properties are dispositional if they can be specified as:

<p style="text-align:center">*x* would *E* under circumstance *M*.</p>

Where *x* is the object that carries the dispositional property, *E* is a description of the manifestation of the property, and *M* specifies the manifestation conditions of the disposition.

Dispositions therefore have a directed character in the sense that they are dispositions *to*. The disposition-schema above is the disposition of *x* to *E*. In order to account for the fact that *x* doesn't always *E*, there are manifestation conditions included in the disposition. *x* only *E*'s when *M* obtains. Powers, on my use of the term, are a subspecies of dispositions. Powers are the fundamental dispositions that are responsible for causation. When I say powers are *fundamental* I mean that they figure in ontology. They are real features of the world, that cannot be analyzed into smaller parts or more basic elements. This goes against many philosophers, most prominently Armstrong (1996), who believe that dispositional properties have a non-dispositional (categorical) basis. I reject such a view: it is my position that there are non-reducible dispositional properties, and these are responsible for all instances of causation. As such, the claim that powers exist is much stronger than the claim that dispositions exist. To say all causation is the work of powers, and only fundamental entities count as powers, is strongly reductionist. There are arguments for and against reductionist theories of many kinds, but I will not get into them. I will say more about what I mean by *reduction* and *fundamental* in Chapter 3.

If powers exist, and there is at least a conceptual distinction between powers and dispositions, are there dispositions that are not powers? Powers are fundamental elements of the ontology I present. There may also be non-fundamental causal dispositions. One example of such a disposition that is not a power because it is not fundamental, is the favorite example of a disposition in the literature: the fragility of a glass. A glass is fragile if and only if it has the disposition to break when being struck with a certain amount of force. However, I don't believe fragility is a fundamental property of things in the world. Glass is fragile because it is composed of (mostly) silica in a strongly bound tetrahedral network structure. The structure and behavior of liquids and amorphous solids like glass could be accurately described using quantum mechanics, but since there are so many bodies (particles) involved, it is computationally intensive. The difficulty with describing the fragility of say, a wine glass in quantum-mechanical terms is due to a lack of computational power, not due to the limits of the mechanics itself. Since the property of fragility of a specific glass could in principle be exhaustively characterized by the quantum-mechanical description of the particles in the amorphous solid, the fragility of the wineglass is not a fundamental property that should be included in our ontology.[3]

There is a conceptual possibility of another kind of disposition that is not a power. The fragility of the wine glass was not a power because it was not fundamental. Other dispositions may not be powers because they do not cause anything. However, I cannot think of any fundamental but non-causal dispositions. If observation presupposes any

---

[3]Much has been written on the reduction of macroscopic object-descriptions to descriptions in terms of elementary particles. There is a number of arguments against this kind of reduction. I believe at least some of these arguments will also be problematic for the type of predicate-reduction I propose. I cannot fully argue against those arguments. The best I can do is give a more or less detailed account of what I mean by reduction and what it entails, as I will do in Chapter 2.

kind of causation, then non-causal dispositions are unobservable. Even if someone comes up with an example, philosophers might be inclined to invoke the *eleatic principle* that only things that play causal roles should be allowed a place in our ontology (Oddie, 1982).

Powers are properties on my view. They are dispositional properties of objects, to be precise. But properties are of a kind, the individual properties belong to *types* or universals. For example, the fragility of this specific glass that I have in front of me is an instantiation of the property of fragility that can be found in many objects. In order to distinguish individual powers and properties from their types, I will use lower case letters such as m for powers of objects, and the corresponding upper case M for the type. I will use the same convention to distinguish between event types and individual events. Most of the talk in this thesis is about individual powers of objects, but sometimes generalizations like 'flu causes fever' are useful albeit somewhat misleading, since causation is between individuals, and not between types.

Although I do not believe the relata of causation to be events, as can be read in Chapter 3, I will use event-speak for my discussion of other theories of causation. This is because most other theories of causation *do* suppose the relata to be events, and because my idea of the relata of causation is incompatible with non-powers accounts. More on the exact conception of powers I endorse will be said in chapters 2 and 3.

1.2   Humean accounts

The strongest reason for accepting powers as fundamental elements of our ontology is the theory of causation that can be offered with the help of powers. Most theories of causation that have enjoyed popularity in this century and the last have been strongly influenced by David Hume's writings on causation. Hume famously wrote that causation cannot be observed in nature, and that ascriptions of causation are a habit of the human mind (Hume, 1748). Hume's influential idea is that nature features no necessary connections, but only particular matters of fact. The necessary connections we see in nature (such as causal laws of the form F's cause G's) are not really there: we project them onto nature. What's really there is just regularity in the particular matters of fact of which nature consists. There is no necessity for a F to cause a G, but it is the case that for every F that happens, a G happens shortly afterwards and nearby.

Causation according to Hume is a necessary relation that cannot be grounded in the particular matters of fact. The neo-Humean metaphysics that is popular in modern-day analytical philosophy posits that all there really is are these particular matters of fact.

> Humean supervenience is named in honor of the greater denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing after another. (But it is no part of the thesis that these local matters are mental.) We have geometry: a system of external relations of spatio-temporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter or æther or fields, maybe both. And at these points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that (Lewis, 1986b, p.ix–x).

More general things, like causal relations or modalities such as necessity and possibility are grounded in these particular matters. The foremost exponent of this school of thought

is David Lewis, whose work formed the basis for the popular view that holds that modality is grounded in particular matters of fact in possible worlds, and that causation is grounded in matters of fact about counterfactuals (which are also dealt with in terms of possible worlds) (Lewis, 1986a)(Lewis, 1973). On this view, necessarily ϕ just means that ϕ is true in all possible worlds. According to Lewis, those possible worlds are real worlds that are causally and spatiotemporally disconnected from our world. As such, ϕ's being true in a possible world w_x is a particular matter of fact about w_x. Necessarily ϕ is made true by the sum of these matters of fact across all possible worlds. Such modal statements are to be translated in terms of quantification over worlds(Lewis, 1968). The relation between these general statements and these particular matters of fact has been dubbed *Humean supervenience*.

This representation of David Lewis' standpoints is not undisputed. The received view of Lewis' theory is that local matters of particular fact in other worlds are the truthmakers for modal truths, as is held by D.M. Armstrong(Armstrong, 2004). However, John Heil has argued that Lewis believed that modal truths are made true by the arrangement of qualities in *our* world, and the similarity or dissimilarity of this arrangement to the arrangements of other possible worlds(Heil, 2015, p.44). This difference is not of major concern to my thesis, as my arguments will hold against both interpretations of Lewis' view. For I believe truthmakers for modal truths do not lie in other worlds, nor in similarity relations between worlds. The truthmakers are in this world – modality is a local matter.[4]

1.2.1 Regularity theory

On Hume's view, causation became not a matter of necessity, but a matter of *constant conjunction*. A F causes a G as long as whenever an F happens, a G follows. That we think F causes G is because we associate them using the principle of causation, that is, we project causation onto the constant conjunction of F and G.[5]

The regularity theory of causation is that f causes g if and only if f and g are spatiotemporally contiguous, f precedes g and F events always cause G events.[6]

One problem with regularity which is also an argument for a powers account is that regularity appears when there is no causation, i.e. spurious regularity. Regularity accounts have been met with many counterexamples where regularity occurs but intuitively, causation doesn't. In response regularity theories have added conditions and clauses to repair the damage.

A second problem for regularity theorists is that there could be causation without regularity. Moreover, there could be unique causal situations, that do not fall under a generalization. These objects would then have to be attributed a type with an extension of one object. Speaking of regularity only makes sense in the context of talking about events of certain types causing events of certain types, for regularity in single instances would *automatically* be perfect. So if there were a causal event that is the only of its type, it can hardly be said there's an interesting regularity. One might argue there is

---

[4]Heil (2015) also argues the amended interpretation of Lewis' account of modal truths is still subject to criticism from an Aristotelian point of view.

[5]There is considerable debate on how to interpret Hume's writing on causation. The way I present it here is as a regularity theory of causation, with a projectionist flavor. Some argue Hume is a realist about causation or a pure projectivist. I will not get into exegesis of Hume's writing here. I choose to present Hume in such a way that it provides maximum contrast with my position.

[6]There is debate on whether events are the relata of causation, and how to individuate them if they are. I will say more on the relata of causation in Chapter 3 and some more on events in Chapter 4. For now I will follow most theorists of causation in their use of events as the relata of causation.

perfect regularity, since if f caused g and f is the only F ever to exist, then in all cases of F's occurring, G followed. But this seems a rather uninformative way of talking about regularity. There may be solutions for this last problem, but there is an intuition at the base of the problem that won't go away. That is the idea that causation is ultimately a matter of tokens, not types. To claim so might not be much more than directly contradicting regularity accounts, but there is another argument to be made: if a theory of causation can handle both token causation and type causation, while another can deal only in types, then the former seems a better choice.

From the point of view of a powers account of causation, it is easy to see why regularity theories run in to problems with spurious regularity and singular cases. For on a powers view, regularity is not the constituent of causation, but rather a symptom of it. For if powers cause things, and certain types of powers cause certain types of things, then regularity is a consequence of the causal events in the world. An analogy can be made with Plato's famous question from the Eutyphro dialog whether the pious is loved by the gods because it is pious, or whether it is pious because it is loved by the gods. In the case of causation we can ask whether there is causation because of the regularity of causes and effects, or whether there is regularity of causes and effects because there is causation.

1.2.2 Counterfactual dependence

Lewis has proposed a somewhat different Humean view: it's not regularity, but counterfactual dependence that defines causation (Lewis, 1973). On this view, f causes g if and only if, had f not happened, g would not have happened. In other words, cause-statements are really statements of subjunctive conditionals, with a counterfactual statement in the antecedent. This view has enjoyed more popularity than the regularist's view, but it has similarly been plagued by numerous counterexamples.[7] Most of these fall into two categories: cases of preemption and cases of overdetermination. Preemption cases are cases where f causes g, but if f hadn't obtained, some backup cause f' would have caused g. Imagine I pour you a glass of water, but if I hadn't, someone else would have poured you the glass. In such cases, the counterfactual dependence test fails: even if f hadn't happened, g would still have.

Overdetermination cases are cases where multiple causes contribute to the production of the effect, and each alone would have been sufficient for the effect to occur. Say I cause a 200ml glass to overflow by pouring the contents of a 220ml glass with my right hand and another 220ml glass with my left hand into the smaller glass. Just pouring one of the two 220ml glasses into the 200ml one would have caused it to overflow. But I poured both, and it seems wrong to say that only my right hand caused the glass to overflow, or only my left hand, since water from both hands entered the glass. In this case, f_1 and f_2 are both individually sufficient for g, but are both involved in the bringing about of g. The counterfactual dependence test fails, since if only one of the causes had occurred, the effect would still have happened.

Just like the regularity view, the counterfactual view has been modified in attempts to counter the counterexamples. From the point of view of a powers theory of causation, it is clear why the counterfactual dependence test doesn't give sufficient conditions for causation. The problem is the same as the problems of the regularity view: the order of analysis is wrong. A powers theorist would argue that subjunctive conditionals about causation *are made true by* the productive relation that holds between causes and effects.

---

[7]For an overview see Paul (2009)

True causal counterfactuals are not what constitutes causation, they are true *because* there is causation.

1.2.3 Humean metaphysics

The problems with the Humean views of causation are a consequence of reversing the order of analysis. This reversal is the hallmark of neo-Humean metaphysics. Because everything has to be grounded in particular matters of fact, and these matters of fact are thought to be non-modal, all modality reduces to quantification over possible worlds. According to neo-Humean metaphysics, the truth of the subjunctive conditional 'if f hadn't happened, neither would g have' lies not in anything about f or g, but rather in the state of affairs in possible worlds where f *didn't* happen. Lewis endorses such reductions for all modal properties; everything boils down to non-modal particular matters of fact. The truth of a counterfactual of the form 'if f hadn't happened, g wouldn't have' depends on the particular matters of fact in a possible world that differs from our own in that f didn't happen, but otherwise is maximally similar. If in that world g didn't happen, then the subjunctive conditional is true.

The embargo on modality *in the world* is at the same time the characteristic feature of neo-Humean metaphysics and the cause of the problems of Humean views of causation. A powers view doesn't have the same trouble with regularity or counterfactual dependence precisely because the metaphysical underpinning of such a view is anti-Humean. On the powers view I propose, modality is not to be conceived in terms of quantification over possible worlds, but rather as being *in the world*. To make these vague metaphysical statements a bit more precise: the truthmakers for modal statements such as subjunctive conditionals lie not in states of affairs in possible worlds, but rather in the states of affairs in this world. To be even more precise, the truthmakers of subjunctive conditionals are dispositional properties of objects in this world. So the claim that if f hadn't happened, g wouldn't have happened is made true by the arrangement of powers in the world that made it so that f caused g. This is the radical difference between a metaphysics that involves fundamental dispositions and a neo-Humean metaphysics: modality is in the world, and not across possible worlds.

In addition, my powers account is firstly an account of token causation, and type generalizations such as 'smoking causes cancer' are secondary. The regularity account has to explain f's causing g by saying how F's cause G's. On my account, token causation holds independently of any types. Causation is the consequence of powers, and powers are individuated in single objects. It is only secondarily that powers belong to types. More on this in Chapter 2.

I believe there is much to be said for a metaphysics that puts modality into the world. In this thesis I cannot argue the case fully, but I can show how a metaphysics that does so can deliver us a good account of what causation is. Furthermore, I simply find it theoretically more appealing to put our modality into this world than to build a system as metaphysically committing as Lewis' modal realism, which holds that all possible worlds exist. A further intuitive argument for anti-Humean metaphysics is given by Strawson (1987). If there aren't necessary connections in nature, then the regularity modality produces (such as the regularity between causes and effects) seems a mere coincidence.To this intuition I add that on a neo-Humean account if something is true in all possible worlds, this is a brute fact, while on an anti-Humean metaphysics, this fact could be explained by the powers in the world. Strawson says there must be something that explains regularity in nature, or else the regularity seems coincidental (Strawson, 1987).

1.3 Skepticism

Hume thought necessary connections cannot be observed in nature.[8] From this thought and an empiricist methodology that does not want to include unobservables we get theories such as the regularity theory of causation. It makes reference only to things that are in principle accessible to us even if you believe in Humean skepticism. It might be thought that this is an advantage of neo-Humean theories. But that is not the case, since neo-Humean theories go further than the epistemological claim that necessary connections cannot be observed in nature: *they deny there are any.* A such, neo-Humeanism is a substantive *metaphysical* thesis, not a merely epistemological one. Just as skepticism about the external world warrants neither the conclusion that there is an external world nor that there is no external world, skepticism about necessary connections in nature warrant neither the avowal or denial of their existence. What this means, is that Humean skepticism does not form an argument for neo-Humean metaphysics.

I do not take a position on the truth or falsity of Humean skepticism. Instead I will develop my theory without regard for skepticism; for I do not believe it to form a counterargument to any metaphysics of causation. The general thought behind this method is that the limits of our perceptive abilities or of our understanding do not necessarily describe the limits of our world. This is a robustly realist intuition. From this intuition, doing metaphysics within the confines of skepticism might distort our image of the world as it really is. When we let skepticism inform our metaphysical investigations, we might find only the boundaries of our skepticism and not the way nature really is. So I must ask the reader to bracket all skepticism about causation, necessity in nature, and the like, in order to proceed.

Anti-Humean metaphysics do not counter skepticism on its own terms, but when properly understood, the skeptic's challenge cannot be given an internal answer, only an external one.[9] The skeptic's position is strong, because it *cannot* be countered on its own terms. That doesn't mean the skeptic's position is compelling: maybe there are good reasons for not accepting the skeptic's position other than a counterargument on its own terms.

In any case, since neo-Humean *metaphysics* is a substantive doctrine, that is not supported by Humean skepticism on the assumption that the limits of our epistemic capacities do not describe the limits of our world, we are free to develop an alternative without further reference to skepticism. Neo-Humean metaphysics is equally unsupported by Humean skepticism. Since skepticism does not give us a prima facie method of judging the neo-Humean theory against a powers theory, which some would call Aristotelian, we can develop both in isolation and see which one does better at giving a metaphysics of causation. Both neo-Humean accounts and Aristotelian accounts such as a powers account of causation stand in need of a justification that is independent of skepticism. In fact I believe neo-Humean theories of causation to be implausible from the viewpoint of common sense and science, giving Aristotelian theories the upper hand in the comparison of theories of causation.

---

[8]Some writers have thought that causation is not unobservable: maybe we feel it when we cause something intentionally, or maybe we feel the causal *oomph* of events (Ducasse, 1926) (Anscombe, 1993). For a discussion of such views see Beebee (2009).

[9]This is the same position as Barry Stroud defends in regard to skepticism about the existence of the external world (Stroud, 1984).

1.4 Powers and science

There are many different views of the relation between philosophy and empirical sciences (or as non-philosophers say, 'science'). It is not my goal to develop a full theory of the relation between philosophy and the entirety of science. I merely wish to emphasize one thesis for which I have no strong argument: that philosophy should work together with science. This is not to say that a priori reasoning cannot constrain science. In fact it does so: if science finds contradictory conclusions when repeating an experiment, we reason there was a difference between the experiments or the two different results are probabilistic in the sense that the chance of either respectively occurring is smaller that 1 and larger than 0. In this line of reasoning we apply the rule of non-contradiction. We simply cannot believe that a conclusion and its contradiction are true. As such, science is constrained by the rules of logic. Scientific inference is inference and therefore must follow inference rules.

The arguments in philosophy do not proceed from logical laws only. In order to give content to arguments, these inferential structures are 'loaded' with premises. These can be intuitions, results from empirical investigation, assumptions, etc. Science tells us what we know about the world. If the premises of a philosophical argument contradict science, then we are loading our arguments with the wrong ammunition. The arguments will turn out unsound. Therefore we must constrain the premises of philosophy to those that do not contradict the best theories science has to offer.

There is a further constraint on philosophy that I wish to endorse; that the premises do not contradict scientific knowledge, is only a minimum requirement. A philosopher could go further, and in my opinion should strive to do so, by giving a theory that not only employs non-contradictory premises, but actually *fits* science. What that means exactly will remain implicit. It would be useful to develop some kind of criterion that would determine whether a philosophical theory fits a scientific theory, but this thesis is not the place to do so. Nancy Cartwright and John Pemberton have attempted to show that powers (albeit somewhat differently conceived than mine) are essential in successful scientific explanation (Cartwright and Pemberton, 2013). For now I will take recourse to showing rather than explaining. In Chapter 3, where my main theses are presented, it will become clear how closely my theory of causation is related to scientific models of interaction of objects. It is my view that a powers account of causation is more fit to integration with laws of nature about the interactions of objects than neo-Humean accounts.[10]

---

[10]A similar argument has been made about accounts that represent causation in terms of *mechanisms*, such as Craver and Darden (2013). These mechanistic accounts fit well with biological and neurophysiological theories. I am not unsympathetic to these accounts as accounts of causal explanation in biological sciences, but I do not believe they succeed in giving a fundamental metaphysics of causation.

Chapter 2

Powers as properties

2.1 Individuation of powers

In Chapter 1 I have been talking about dispositions as properties. If powers are dispositions, powers must be properties.[1] Properties are instantiated universals – as universals they can apply to multiple objects, but the instances are bound to a single object. Since I believe the instances to be primary, and the universals secondary, when I speak of powers or properties, by default I mean instances in a certain object. If every power belongs to exactly one object, then there is a mapping for every power to a unique object:

$$\mathcal{X}_{(p)} = x \qquad \text{OBJECT FUNCTION}$$

Where p is a power, x is an object and $\mathcal{X}$ is the OBJECT FUNCTION that 'reads off' the objects from the powers.

Every power has a certain kind of manifestation connected to it. It is a power to do something of a certain kind under certain conditions. Notice the manifestation is a manifestation *type*. The glass doesn't have the power to undergo one specific spatiotemporally located breaking, it has the power to do something that's *a* breaking. Powers are essentially connected to their manifestations. Even if the manifestation conditions are the same, if the manifestations are different, the powers are distinct.[2] This is not hard to see: the power to melt at 100°C is distinct from the power to evaporate at 100°C. This can be expressed as a function from powers to manifestation types:

$$\mathcal{E}_{(p)} = E \qquad \text{MANIFESTATION FUNCTION}$$

It is important to note here that every power is connected to a manifestation type E, but not always to an instantiated manifestation e. It might be the case that an object has a power that never manifests, due to the fact that its manifestation conditions never obtain. For example, we could have a charged particle that doesn't have any other charged particles in its light cone. In such a case it would never manifest attraction or repulsion due to its charge.

Intuitively it might seem that powers that have the same manifestations, but different manifestation conditions, could still be the same power. For example, the power of nitroglycerin to explode when subjected to flame might be thought to be the same power

---

[1] In section 3.5.2 I discuss why we might want to say that powers are subsumed under properties, not identical to them. Nevertheless powers are *ways things are*.

[2] In section 3.3.1 I will say something about powers having manifestations of the same kind but different strengths over time.

as the power of nitroglycerin to explode when subjected to shock. However, in other cases the intuitions point the other way: take the power of ovalbumin (a protein found in eggs) to solidify when cooled to below freezing, and its power to solidify when heated to >80°C. When subjected to heat, the protein solidifies by the breaking of bonds. Freezing ovalbumin does not break the bonds. A consequence is that frozen egg whites will retake their liquid form when thawed while cooked egg whites will remain solid when cooled. In this case there seem to be two different powers at work.

In reality I think neither example elucidates much about the individuation of powers, as the powers in the example are not fundamental. The explosive disposition of nitroglycerin has as its truthmaker the microphysical structure that will expand by a factor ∼1200 (at room temperature and normal pressure) when detonated. This microphysical process, when seen at the level of the individual molecule, has much clearer manifestation conditions than the macroscopic process has. The molecular behavior of nitroglycerin is a consequence of the interaction and arrangement of the atoms that it consists of. The atoms, in turn, interact the way they do because of their constitution of fundamental particles and the powers of these particles. These fundamental powers have clearer manifestation conditions than vague dispositions such as the disposition to solidify. I believe it will become clear that no disjunctive manifestation conditions need to be allowed when we include only fundamental powers in our ontology, as I will argue for in 2.2.4 and 2.2.5. If we accept that powers have only one manifestation condition, then there is a MANIFESTATION CONDITION FUNCTION $\mathcal{M}$ from powers to their manifestation conditions. Just like manifestations, the manifestation condition is a *type*, because a power is not a disposition to a manifestation under a single specific spatiotemporally located manifestation condition, but under a manifestation condition of a certain *kind*.

$$\mathcal{M}_{(p)} = M \qquad\qquad \text{MANIFESTATION CONDITION FUNCTION}$$

Powers are individuated by their manifestation types, their objects and their manifestation condition types. This makes me an essentialist in the sense that the manifestation type and the manifestation condition give part of the essence of a power.[3]

I do not think that a metaphysics that involves powers commits one to a certain theory of properties or universals. However, my theory of causation, in a way, does. For I take causation to be a matter of individual objects and their powers. As such, the power-instances are the relevant units, while the power-universals are of secondary importance. Furthermore as a naturalist metaphysician, I find accepting universals as separate entities (i.e. platonism) altogether unappealing. I think these tenets make my theory of powers compatible with several theories of properties, including tropes and various forms of nominalism (Armstrong, 1989)(Campbell, 1990). My theory, since it posits essences of powers, is not compatible with any theory of universals that insists on quidditism.[4]

Although compatible with different ideas about universals, a power with an essence enables us to construct sets of powers that have the same type. The universal, or *power type*, that corresponds with a particular power is represented by an uppercase letter. This letter stands for the powers that share their manifestation types and manifestation conditions. Since powers are powers of something to do something under certain circumstances, there can be a class of powers that are powers to do the same thing under the same circumstances. In this way we gain the power type, since it is the power, abstracted from the object.

---

[3]This view is also defended by Ellis (2010) and other dispositional essentialists.
[4]See 4.4

$$P = \{p \mid \mathcal{M}_{(p)} = M_0 \ \& \ \mathcal{E}_{(p)} = E_0\} \qquad \text{TYPING}$$

But we need to add the proviso that $P \neq \varnothing$, since we do not want a power type without possible corresponding powers.

TYPING is not very committing, as it does not mean that every possible set of powers constitutes a property. Some theories of universals, such as a radical version of class nominalism, would have that every set of objects constitutes a universal (Armstrong, 1989, p.8-10)(Armstrong, 2010, p.7). Typing does not commit us to an analogous position, since not every set of powers has shared manifestation types and manifestation condition types. As such, TYPING is a constraint on the formation of universals from particulars. Neither does TYPING commit me to the thesis that powers are what they are because they belong to certain types. Some class nominalists about universals would have that particulars bear certain properties in virtue of belonging to some universal (Armstrong, 2010, p.9). I am not committed to an analogous position by TYPING, since the order is the other way around. Powers belong to a certain power type in virtue of possessing certain manifestation types and manifestation conditions. Powers belong to a class of powers precisely because they are what they are – class membership is determined by the instances, not the other way around. This is to say that particulars, on my view, have metaphysical priority over universals.

The method of creating types from powers given above specifies there is a function from powers to universals, or *power types*, since every power has both a manifestation conditions type and a manifestation type. This means that for every power there is a power type such that the power belongs to it.

$$\mathcal{P}_{(p)} = P \qquad \text{TYPE FUNCTION}$$

Since power-type formation is restricted by the manifestation conditions and manifestations, the individuation of power types is dependent of the individuation of these manifestation condition types and manifestation types. Now, we could imagine two powers with the same manifestation where one has manifestation conditions that include the manifestation conditions of the other. Take the property of *being brittle*. We can conceive of being brittle as a special kind of being fragile. If being fragile means being disposed to break when stressed, and being brittle means being disposed to break when bent, and bending is just one type of stress that can be applied, then brittleness is a sub-type of fragility. The manifestation conditions (being bent) form a special case of the manifestation conditions (being stressed) of fragility.

If this could be the case, then there could be power types that are subtypes of other power types. Then a power that is in the subtype, would also be in the meta-type. That would mean that there is no function such as TYPE FUNCTION that reads off a unique type from the power. That would interfere with my attempt to provide non-ambiguous criteria of individuation for powers. However, I believe the problem is solved by what the manifestation condition types look like for fundamental powers. Manifestation conditions are spelled out in terms of a well-formed scientific theory. If this is carefully done, than manifestation condition types can be mutually exclusive so that every power is in precisely one type.

Alternatively, if our theory of what powers there really are (science) tells us that there really are powers that have multiple manifestation conditions, we must individuate powers more weakly. That means that there is a function from powers to all the manifestation conditions that power has. The power is then individuated by the object it is a power of, its manifestation and the combination of its manifestation conditions.

Of course, an object can have more than one power, for example, a glass can be both fragile and transparent. I think the powers account can fit both bundle theories, where objects are conceived of as bundles of properties (or tropes, etc.), and substrate theories, where properties are seen as belonging to some object.[5]

An object cannot carry two instances of the same power type. For if an object had two powers of the same type, they would share manifestation conditions and manifestations. When the manifestation conditions obtain, both powers would manifest in exactly the same way simultaneously. Such a scenario seems to me a radical case of overdetermination. There would also be no way to judge the object had these two powers, and not just the one of them, since powers are observed through their manifestation. Since the manifestation of the two powers would be identical in character and simultaneous, a double manifestation would be indistinguishable from a single manifestation. However, this point depends how finely we carve objects and manifestations. There are cases in which it may be possible for an object two have two powers of the same type, if they occur in different regions of the object. For example, a cylindrical rod could have the power to make circular impressions with its one end, and also with its other end. I think such cases are harder to come up with when we look only at fundamental objects, but this is not a knock-down argument.[6] In the end it comes down to where we want to deal with the problem: if we allow for objects to have multiple powers of the same type, individuating powers becomes difficult. Alternatively, we could simply say this is impossible, and that for any example we can think of, we must simply individuate objects in a more fine-grained manner. But this is not a solution, but rather a shifting of the problem from the domain of powers to the domain of objects.

A power is the instantiation of a power type in an object. If an object cannot carry two instances of the same power, giving the power type and the object it is a power of, provides adequate criteria for the individuation of powers. However, the success of such criteria depends on having criteria for the individuation of objects and of power types. The power types are sets of powers that share manifestation conditions and manifestations. The success of individuating powers depends of explicating what the manifestation conditions and manifestations of fundamental powers are. I will be engaging that task in Chapter 3.

The causal work is done not by power types but by the powers. Power types would be strange causal actors indeed: they are only sets of powers.[7] The sets are formed according to the manifestation conditions and manifestations of individual powers. Therefore to say that powers are individuated by their objects and types is a roundabout way of saying they are individuated by their object, manifestation conditions, and manifestations.

$$p_1 = p_2 \iff \mathcal{M}_{(p_1)} = \mathcal{M}_{(p_2)} \,\&\,$$
$$\mathcal{E}_{(p_1)} = \mathcal{E}_{(p_2)} \,\&\,$$
$$\mathcal{X}_{(p_1)} = \mathcal{X}_{(p_2)}$$

Identity criterion

Which is, seeing how the power types are constructed, essentially the same thing as:

---

[5]I do not believe such a substrate theory would imply the existence of bare particulars (Armstrong, 1989, p.95). I find the bare particulars argument against substrate theories unconvincing.

[6]John Heil suggested that a field could be configured in the same way in different places. If fields are objects, this could be an example of a fundamental object having two powers of the same type.

[7]Power types are only sets of powers, but the formation of these sets is governed by TYPING. Therefore being a member of a set implies having a certain essence. Thereby these sets become a lot like, if not completely the same as natural kinds.

$$p_1 = p_2 \Longleftrightarrow \mathcal{P}_{(p_1)} = \mathcal{P}_{(p_2)} \ \& \qquad \text{Identity criterion}^*$$
$$\mathcal{X}_{(p_1)} = \mathcal{X}_{(p_2)}$$

2.2 Categorical and dispositional properties

Dispositional properties are often contrasted with categorical properties. Categorical properties are usually simply defined as non-dispositional properties. For example, the property of being triangular is often thought to be non-dispositional as it doesn't seem to be properly analyzed when we give it a dispositional conditional analysis. Some authors, most notably Armstrong, think all dispositional properties have categorical properties as their bases. For example, it would be said that the fragility of the glass is dependent on its strongly bound tetrahedral network structure of (mostly) silica molecules. If this structure is a categorical property, then the disposition would have a categorical basis. Other authors think properties are purely dispositional, or both dispositional and categorical in some sense.

I endorse the idea that properties can be essentially dispositional. This is in conflict with Armstrong, but in accord with Mumford and Anjum, Heil and Martin, and others who call themselves 'dispositionalists.'. I am not sure I believe all properties are purely dispositional, as Mumford and Anjum do.[8] My position is closer to Heil and Martin's, although I find the representation of their idea as an *identity* of categorical and dispositional properties objectionable. I will argue for my position by reviewing the other positions mentioned.

2.2.1 Nomic necessitation

On Armstrong's view, all dispositions have categorical bases. Causation is governed by laws, and the laws relate universals to each other. For example, a causal law could be that F's cause G's. Then there is a relation N between F and G that Armstrong calls *nomic necessitation*. This higher-order relation means that if F and G are related by nomic necessitation, then Fx brings about Gx.[9] In other words, if an object x is F, and $N(F, G)$, then Gx. It is unclear what exactly nomic necessitation should be, and how a relation between universals can explain causation between particulars. Lewis points out that just calling the relation 'necessitation' isn't good enough:

> But I say that N deserves the name of 'necessitation' only if, somehow, it really can enter into the requisite necessary connections. It can't enter into them just by bearing a name, any more than one can have mighty biceps just by being called 'Armstrong' (Lewis, 1983, p.366).

Armstrong thinks the necessary connections between universals are provided by nomic necessitation because the laws confer a *soft* modal force on first-order properties, that is strong enough to support counterfactuals and necessitate the effects. So if $N(F, G)$, then Fx necessitates Gx *because* N holds between F and G. The modal force is soft, because it isn't quite metaphysical necessity, since the higher-order relation itself is contingent. However, it is stronger than just contingent regularity, since there must be a difference,

---

[8]See section 3.5.2

[9]Armstrong's view that laws are higher-order properties, i.e. relations among universals, is shared by Dretske (1977) and Tooley (1977).

according to Armstrong between lawlike regularity and non-lawlike regularity (Armstrong, 1997, p.223). So, according to Armstrong, that properties are related in a lawlike manner is contingent, which properties are related is contingent, but the relation itself introduces a soft necessity between universals that is stronger than contingent regularity.

Dispositions gain their modal force from these laws. The dispositional character is imposed on a categorical property by the higher order relation of nomic necessitation. As such, Armstrong's account of dispositions is a combination of categorical bases and laws of nature that relate these categorical properties to yield dispositional features.

Lewis' quip posed the question how the higher-order relation can account for necessitation at the level of individuals. The higher-order relation is just a new name for necessitation, not an explanation. Something similar has been argued by Van Fraassen (1989), who says it is impossible to identify at the same time an acceptable account of what the nomic necessitation relation is and the way it works on particulars. Armstrong has replied that nomic necessitation is inferred from observed regularities; it is the best explanation of lawlike regularity (Armstrong, 1993). Fraassen (1993) has in turn argued we are not justified in inferring a higher-order relation from regularity in relations between individuals.

Alexander Bird has a stronger argument than Van Fraassen. According to Bird, Armstrong cannot account for the necessitating influence of the higher order relation on the individual instantiations of the properties involved without accepting these individual property-instances as themselves powerful. That is, without deferring to what's essentially a powers account of natural laws (Bird, 2007, p.91–97). The gist of the argument is as follows. Suppose

1. $N(F, G)$ entails perfect regularity between F's and G's: $R(F, G)$

But that would give N non-trivial modal character.

2. Properties cannot have a non-trivial modal character, or they wouldn't be categorical.

So 1 must be false, by Armstrong's idea that all properties are really categorical

3. If 1 is false, then there are worlds where $N(F, G)$ but not all Fx cause Gx.

Which would mean that there are worlds which have exceptions to laws of nature. This is in conflict with the spirit of nomic *necessitation*.[10] So Armstrong is left with either an account that can't account for the necessity, soft or otherwise, of the consequences of laws of nature, or with an account that allows non-trivial modal properties, i.e. dispositions without a categorical basis.

2.2.2 Eleatic principle

There is also a simpler argument against the categoricalist: an appeal to the eleatic principle. For if it's the disposition that does the causing, and not the categorical basis, then why would we accept the categorical basis in our ontology that is restricted by the eleatic principle and therefore contains only things which have causes or effects?

There is another reason not to depend too much on categorical properties: if detection is a causal process, then only things that have effects can be detection.[11] Therefore, if

---

[10]Bird's argument continues with a counter to two possible solutions, which are omitted here. (Bird, 2007, see p.93–96).

[11]Heil makes this argument and I say more about it in 2.2.3.

we accept a dispositional theory of causation, only powerful objects can be detected. The underlying categorical bases are undetectable, except for their dispositional form. I consider it a possibility for anyone who holds a dispositional theory of causation to reject categorical properties altogether. At any rate, for a powers-theorist of causation, categorical properties do not seem the area to focus on, for in a theory of causation, why dwell on non-causal properties?

Many properties that are thought to be categorical also have dispositional aspects. For example, triangular objects have a disposition to leave a triangular impression on soft materials when pressed down. If we assume that causation is a dispositional matter for a moment, then all the causal roles the categorical property of triangularity plays are explained by the dispositional aspect of the property. For a theory of causation, what is the interest of the non-causal aspects? If causation is a matter of dispositions, we need not worry much about categorical properties.

Could there be properties that are neither categorical nor dispositional? One option would be hybrid properties such as the disjunctive property of being square or fragile (McKitrick, 2003, footnote 9). I think such hybrid properties are a byproduct of bad philosophical methodology. *Properties* are features of the world, and their existence is therefore determined by *what there is* in the world. *Predicates*, on the other hand, are linguistic units. What predicates there are is a question of what terms there are in a language. That a predicate exists, does not mean a corresponding property also exists. We can even come up with predicates that correspond with impossible properties, such as the property of being square and round.

In *philenglish*, the language philosophers speak, we have a rule of compositionality that allows us to create compound predicates by connecting two predicates with a logical connective. In this way we can we get the predicate *being square and round* or *being square or fragile*. However, in the world, we cannot simply take two properties, put a connective between them, and say they're now a single property. I find it highly unlikely that disjunctive properties actually exist in the world. It is a common mistake in analytic philosophy to confuse predicates and properties, and reason too readily from the existence of predicates to the existence of properties.[12]

2.2.3 Identity theory of powers

Martin and Heil have proposed an identity theory of categorical and dispositional properties (Martin, 1997, p.215)(Martin and Heil, 1999, p.45-46)(Heil, 2003, ch.11)(Heil, 2005, p.346)(Heil, 2012, p58-62). According to Heil (2012, p.59), speaking about 'categorical' properties is a somewhat misleading way to speak of properties that are here and now, and not purely conditional. The purely conditional view of dispositions, held by Gilbert Ryle, says the disposition is nothing actual about the object, but purely how it would behave in certain circumstances. (Ryle, 1949)

The Martin-Heil view is that properties are not distinctly categorical or dispositional, but rather both. This is incompatible with the standard definition found in most discussions of dispositions, that defines categorical properties as simply non-dispositional. As such Heil and Martin prefer to use the term *qualities* over *categorical properties*. The idea then is that all real properties have the here-and-now nature of categorical properties: they are not purely conditional. But the account goes further: where Armstrong would have it properties are dispositional in virtue of their categorical basis combined with laws of nature, Heil and Martin hold that properties are dispositional essentially. On the

---

[12]See also section 2.2.4.

Armstrong view, if the laws of nature had been different, the categorical property could have a different dispositional nature. Since the laws of nature are contingent, according to Armstrong, the connection between dispositional properties and their categorical basis is also contingent.

Heil's argument draws on the point made in 2.2.2, that non-powerful properties would be unobservable. That is a strong argument against anyone who endorses a dispositional theory of causation, without accepting all properties are dispositional, but not quite against those who hold other theories of causation, or believe, like Armstrong, that dispositions *do the work* but have categorical bases. If non-powerful properties are unobservable, that would mean that even common examples of categorical properties such as triangularity would be dispositional, for otherwise we would have had no perception of triangularity.

In the end the picture of properties that emerges form the Heil-Martin view is that they are powerful qualities. That means they are ways things are here and now, which accounts for the qualitative aspects, but that these ways things are here and now can be dispositional ways things are. The connection between the property and the disposition cannot be severed as it can in Armstrong's case. Since the dispositional aspect of the property is part of what makes that property that property, part of it's essence, if the property had a different dispositional content, it would have been a different property.

I agree with this idea, although I find the equation of categorical properties and qualities somewhat problematic, since they are commonly defined in a mutually exclusive manner. However, that might be no fault of Heil's or Martin's, since they propose to use 'qualities' as an alternative to 'categorical properties', and not as a synonym. As such it can be that the identification of qualities with dispositions is not contradictory.

I would say that properties that have no dispositional essence whatsoever have no place in our ontology, on the basis of the eleatic principle, and Bird's arguments against Armstrong.[13] But that is not to say that all there is to properties is a purely conditional nature. Powers, as I see them are not *virtual* features, or only conditional, they are features of objects here and now. Spatiotemporal features of objects, such as location or shape, have sometimes been classified as categorical. It seems obvious that location and shape play important roles in causal processes, and can figure in dispositions. As such, they are dispositional property types, *and* actual. Whether this means that all properties are both dispositional (in a way) and categorical (in a way) as Heil and Martin think, I don't know. I think we could get by with just powers, as long as we do not fall for Ryle's view of them as purely dispositional and non-actual. However, I think that if there is any disagreement between my view and Martin and Heil's, that it is a merely or mostly verbal disagreement.

Sydney Shoemaker has proposed a view that properties are sets of powers. What's categorical about the property then is just that it bundles powers and is not itself a single power (Shoemaker, 1980). Mumford and Anjum have accepted this view for all properties (Mumford and Anjum, 2011). I will say more about this view in section 3.5.2.

2.2.4 Predicate or property

Within the everyday linguistic practice of talking about objects it is strange to say this glass doesn't exist, and there are only particles arranged glass-wise, or that sentences such as 'this glass is transparent' cannot be true. The reduction of macroscopic object-descriptions to descriptions of complexes of particles brings us to a way of speaking that

---

[13]In 3.5.2 I give a reason to accept categoricals, but only in connection with powers.

has such detail that it is not of much use in our everyday interaction with objects: we cannot see the individual particles, feel them, or do anything with them. But this thesis is not on the pragmatics of philosophical concepts. For the metaphysics of objects it really doesn't matter that much how we talk about objects. The questions are about *what there is* and *how it is* not about *how we talk*. There is a mismatch between the nouns in our languages and the objects in the world: some things we have names for don't really exist, and there are things that exist that we do not have names for or that we don't even know about.

The story about properties is not any different from the story about objects. Surely the existence of properties in the world cannot be directly derived from the fact that we ascribe certain predicates (the terms for properties). This would be a mistake equivalent to taking our linguistic practices with object-words to inform ontology. Similarly, we should not take the predicates we use to imply there really exist such properties. In fact, most predicates we use do not correspond with the properties of the real constituents of the world. This is because there is a mismatch between properties, that is, real features of objects in reality, and predicates, that is, the linguistic units that are attached to subjects in propositions.

Some form of this view on the mismatch between properties and predicates is held by most analytic philosophers that endorse a realist method that rejects the linguisticized metaphysics that became popular in the 20th century in the wake of Wittgenstein, Quine and Carnap. Lewis characterizes the difference between predicates and properties as the difference between a *sparse* view of properties and an *abundant* view of properties. Sparse properties are those that really exist, abundant properties are all properties we ascribe to objects, i.e. properties as corresponding to every predicate.[14]

Most philosophers working in this area, such as Armstrong, Bird, Heil, and Mumford, endorse similar distinctions. I prefer to use the pair *properties* and *predicates* over *sparse* and *abundant* properties, in order to preserve the term *properties* for things in the world, and let the term *predicates* handle the linguistic component. There can be predicates that directly refer to properties, and there are many predicates that do not (directly) refer to properties.

On our best model of the world, an electron is a fundamental particle. Electrons have the property of *negative charge* The predicate *has negative charge* refers to the property of having negative charge. Since an electron has negative charge, the sentence 'This electron has negative charge' is made true by the negative charge of that electron. The predicate 'is a circular square' does not pick out any property in the world. The sentence 'This piece of paper is a circular square' is therefore false, since the piece of paper can not be a circular square. These are the simple cases, where the predicate directly refers to a property in the world, or where the predicate does not refer to a property at all.

Examples get more complicated when we refer to objects or predicates that are not fundamental, but the sentences still seem true. Take the sentence 'this peach is sweet'. Peaches are not among the fundamental constituents of the world, but rather agglomerates of such constituents. Sweetness is not a property of the fundamental constituents of the world. Yet the sentence 'This peach is sweet' can be true. And while it might be the case that not *all* truths need a truthmaker (think of tautologies or mathematical truths), it seems to me that at the very minimum non-general empirical truths have a truthmaker.

---

[14]Of the dispositionalists Bird (2007) and Heil (2012) accept only sparse properties while Mumford and Anjum (2011) and Kistler (2007) embrace more abundant properties. Cartwright is an anti-realist about scientific models and therefore thinks even physics doesn't usually deal in 'real' (sparse) properties (Cartwright, 1983).

The truthmaker for 'this peach is sweet' must be some complex of fundamental entities and their properties. To give a total description of the truthmaker of 'this peach is sweet' would be very hard, but seeing that there must be a truthmaker is not.

2.2.5   Reduction and truthmakers

In the previous sections I have hinted at a picture of the world, and of causation in particular that many would call reductionist.[15] I've put this in terms of truths and their truthmakers. I want to clear up what kind of reduction I endorse, and how it would work.[16]

First of all, it must be stated that there is no revisionary goal for ordinary language in this thesis. Some philosophers, like Paul Churchland, have advocated a bottom-up approach where ordinary language should follow the contours of scientific knowledge, so that the predicates in language carve the world at the joints of nature (Churchland, 1981). My aim is rather to have philosophers and scientists carve the worlds at the joints of nature, and to distinguish between predicates and properties. While there might be much *folk physics* at work in ordinary language, it is often more useful and certainly easier to speak of 'objects' that are on the same scale as our bodies, and 'properties' that are consistent over these macroscopic objects, rather than multiply realized at the fundamental level.[17]

Second, an important point must be made concerning ontological reduction. In the literature on reduction, a distinction between ontological and epistemological or linguistic reduction is sometimes made. I find these terms obfuscatory. For if ontological reduction is to mean reducing particular entities to other particular entities, or particular properties to other particular properties, then it is impossible on my account. For all there is, are the fundamental entities, and their fundamental properties. There is nothing to reduce ontologically once one accepts that there are only fundamentals. The non-fundamental stuff like peaches, glasses, and fragility, only exist as linguistic entities that feature in sentences that are made true by complexes of fundamentals. Then is my aim to reduce one way of speaking to another way of speaking? To replace predicates with other predicates? No. Language is fine as it is. The goal is not translation of predicates ordinary language to those of fundamental physics. I am not trying to get rid of anything, not entities, not properties, not predicates. The 'reduction' I endorse merely aims to establish connections between sentences and their truthmakers. My goal is to give a theory where microphysical objects and properties can be the truthmakers for cause-sentences. The path of reduction leads not from predicates to other predicates, nor from entities to other entities, but rather from language to reality.

The reduction I wish to endorse tries to home in on the truthmakers for sentences by replacing sentences with ones more directly referring to the truthmakers. For example, we say of a glass that it is fragile. But this property-ascription to this object could be reduced to the powers of the elementary particles that make up the glass. To get at the truthmakers of sentences, we must take a path down their meaning and to the world.

---

[15]Many dispositionalists would not endorse the reductionism I advocate. See Kistler (2007) for arguments against such reductionism.

[16]Much of this paragraph and the previous is inspired by, and in agreement with, Heil's view as presented in Heil (2003) and Heil (2012). The refutation of levels of reality I read in these books resonated with what had become my standpoint after thinking and reading about philosophy of mind, especially the work of Donald Davidson (1980).

[17]Scare quotes around *objects* and *properties* are meant to indicate that they are here used in a ordinary language sense; that these words are really non-directly referring predicates and nouns. This means that 'properties' here refers to universals, and not instantiations as elsewhere.

Fundamentality is a concept that is often appealed to but perhaps not easy to define accurately. I will spell out fundamentality in terms of the relation between descriptions and their truthmakers. Roughly, more fundamental descriptions *stand closer to their truthmakers*. In the first chapter I claimed that only fundamental dispositions are powers. Combine this with the view presented in 2.2 that all properties that matter are powerful properties, and we get the view that all true predicate-ascriptions are made true by powers. As such we get a reduction from subject-predicate sentences to objects and powers.

Take the following sentence:

'This paint is red' is true $\rightarrow$ this paint is red

The left half of the sentence is a predicate attribution, a description. We might think the right half is the truthmaker for that description. The truthmaker for the paint being red is that the paint is red. The sentence is bound to this something in the world by its meaning. But is the formulation I've given of this truthmaker really what makes this sentence true? We could exchange the right side of the Tarski-sentence in order to get:

'This paint is red' is true $\rightarrow$ this paint reflects mostly red light

If this sentence is true, we've now reduced the predicate *being red* to reflective properties of the substance in question. We could take it a step further:

'This paint reflects mostly red light' is true $\rightarrow$ this paint reflects light in the $\sim 700 nm$ range with a higher amplitude than other frequencies.

The idea is that there is a point where the description cannot be reduced further. At this point the truthmaker on the right side of the implication will simply be the same as the sentence on the left side, but without the quotation marks. Only then have we arrived at predicates and nouns that directly refer to the fundamental constituents. This kind of description would be a *fundamental description*: one that has its direct and literal meaning as its truthmaker. But if the method of arriving at the truthmakers of sentences is through substitutions of synonymous descriptions, how do we know we are making the right substitutions? Synonymous descriptions are freely substitutable (in a semantically transparent context). So then why would we say that 'this paint reflects light in the $\sim 700 nm$ range with a higher amplitude than other frequencies' is more fundamental than 'this paint is red'? There has to be some kind of asymmetry we can appeal to in order to make sure that, when reducing, we only go *downwards*.

The other, perhaps more promising answer, is simply to defer to mereology and science. This particular stroke of paint is made of particles. Science tells us what particles there are. The particles are part of the paint, but the paint is not part of the particles. The mereological part-whole relation gives us the direction of reduction because it is asymmetric. Science gives us a point where to stop, since physics discovers whether a particle is fundamental or not. The discovery of the substructure of the atom gave us a reason to think reduction to atoms is not full reduction to fundamental particles.

The descriptions of what makes the sentence true that are put in the terms of these fundamental scientific theories stand closer to their truthmakers than descriptions like 'red paint'. In fact, they stand as close to their truthmakers as possible. When I said reduction was a question of getting from the truths to the truthmakers, this is what I meant. My appeal to science, specifically physics, is debatable. Physics, like other

sciences is a way of describing the world.[18] I assume for the course of this thesis that fundamental physics is that way of describing the world that aims at describing in terms that directly refer to their truthmakers.

2.3 Finkish problems

2.3.1 Finking

Dispositions are usually analysed as conditionals of the form:

$$o \text{ would } F \text{ under circumstance } c.$$

Finks, masks and mimicries are counterexamples to such an analysis of dispositions. Finkish dispositions are situations where the circumstance that would normally be sufficient for the manifestation of the disposition obtains, but the manifestation itself does not occur. C.B. Martins original example was of an *electro fink* (Martin, 1994).[19] We say a wire is live when it is disposed to conduct electricity when touched by a conductor. An electro fink is a device that is attached to a wire and senses when a conductor comes close. If that happens, the electro fink turns the dead wire into a live wire or vice versa. Now imagine a wire that is dead and has an electro fink attached. If a conductor comes near the wire becomes live by the operation of the electro fink, and therefore conducts electricity. As such, the wire is disposed to conduct electricity when touched by a conductor. But the start of the scenario stipulated that the wire to which the fink was attached was dead, and therefore not disposed to conduct electricity. An inverse situation is also possible: the wire to which the fink was attached is live, but the proximity of the conductor activates the fink, so that the wire does not conduct electricity when the conductor touches it. In the first example the manifestation conditions and the manifestation occur despite the absence of the disposition. In the second case the disposition is not manifested despite the manifestation conditions obtaining. As such the conditional analysis of dispositions gives the wrong outcome in cases where a fink is present. Choi and Fara (2014) point out that Martin's counterexamples are examples of what Shope (1978) calls the *conditional fallacy*. The conditional fallacy is to see the truth of some statement as dependent on a subjunctive conditional, and failing to see that truth is (in some cases) dependent on the truth of the antecedent of the subjunctive conditional. In the case of the electro-fink, the truth of 'the wire is live' is dependent on the truth of the antecedent (the wire is touched by a conductor). The counterexamples depend on a deviant causal chain: the antecedent of the conditional in combination with the fink cause the opposite from the normal situation, i.e. non-conductance instead of conduction and vice versa.

2.3.2 Masking and mimicry

Masking is a line of counterexample similar to finking: a *mask* prevents the disposition from manifesting even though the manifestation conditions obtain.[20] An example is a

---

[18]To get from the way physical description work to the thesis that statements of physics give us direct access to truthmakers of statements of special sciences requires further argument, which I shall not give here. If it turns out there are strongly emergent objects or properties, the truthmaking bases of which are not mereologically simple, I would modify my account to include such strongly emergent objects or properties as fundamentals.

[19]The extensive discussion of finks in the literature is probably due to Lewis (1997). Note not everyone is convinced the counterexamples are as effective as claimed (Bonevac et al., 2012).

[20]Masks are called *antidotes* by Bird.

fragile glass packaged in bubble wrap. The glass has the disposition to break when being struck, but when it is struck, the bubble wrap prevents the striking from causing the breaking. The difference between *masks* and *finks* is that finks remove the disposition from the object (the wire is no longer disposed to conduct when the fink renders it dead), while masks block the manifestation without changing the dispositions of the glass.[21] Mimickers are things that seem to manifest a certain disposition under certain conditions, but do not possess that disposition. For example, a Styrofoam cup does not have the disposition to break when struck, but it makes a distinctive sound. The Hater of Styrofoam, who lurks near, will react to the sound by swooping in and breaking the cup. If we suppose the hater of Styrofoam is always near, then there is a true generalization that Styrofoam cups break when they are struck. This seems a lot like a disposition to break when struck (Lewis, 1997).

2.3.3 Non-standard analysis

Another way to attempt to repair dispositions is by giving a different analysis of dispositions. For example, dispositions could be analyzed as the disposition of x to E when M in the absence of finks, masks and mimickers. But that seems rather ad hoc: for what is common to finks, masks and mimickers other that that they are counterexamples to dispositions? Therefore the new analysis would stipulate, and not prove, that dispositions cannot be finked, masked or mimicked. Lewis (1997) has suggested amending the simple conditional analysis with the clause that states only intrinsic dispositions can count as counterexamples, so that the extrinsic circumstance of the wire (that a fink is attached) does not form a counterexample. But new examples with 'tricky triangles' involving intrinsic finks have been developed (Choi, 2005).

A different solution is to appeal not to the conditional as an analysis for dispositions, but rather analyze them in a comparative manner (Manley and Wasserman, 2007). If x has a disposition to e when m, that means that x is more disposed to e in situation m than y is. This approach gets rid of the finking problems the conditional analysis has, since conditionals do not admit of degree. By turning to degree, Manley and Wasserman are shifting to some sort of notion of standard cases, thereby eliminating finks and masks and mimickers. However, it seems that such a gradable approach works by quantifying over outcomes of individual scenarios. For 'is more disposed to' just means 'would e more often'. But which scenarios where x does e and y does not e count towards the grading? This seems impossible to specify without a preconceived causal notion.[22]

An approach that is somewhat similar to Manley and Wasserman's is to say dispositions are *habituals*. Fara (2005) analyzes dispositions as habituals – generalizations that allow exceptions: habitually following a healthy diet allows eating an unhealthy snack every now and then. But Fara's account has similar pitfalls to that of Manley and Wasserman: habituals are quantifications over particular instances. Yli-Vakkuri (2010) argues that Fara's habituals are just disguised *ceteris paribus* clauses that are impossible to spell out without making the amendments *ad hoc*. Another worry about habituals is that they seem to be type-properties. Many token dispositions, as instantiated in a single object, can only manifest once. For example, the fragile glass can only manifest its disposition to break once. So in what sense does it have the *habit* to break? For something that can happen only once, saying it's a habit of tendency seems strange.[23]

---

[21] Masking is introduced by Johnston (1992), but the best examples come from Bird (1998) who uses the term 'antidote'.

[22] For this criticism of Manley and Wasserman see Choi and Fara (2014).

[23] This point is made in Mumford and Anjum (2011, p.6)

Another solution is simply to claim that with masking and finking disposition is interfered with in the time between asserting the disposition and the occurrence of the manifestation condition. Say an object o has a disposition p at time t_1, and the manifestation occurs at t_3. The fink acts at t_2 and removes the disposition, such that at t_3, the disposition is no longer present when the manifestation conditions occur. This would be an effective strategy if all finks, maskers and mimickers depended on there being some span of time between t_1 and t_3. But for some examples, t_1, t_2 and t_3 coincide – the assertion of power p is true at all those times, and the fink works instantly. In Martin's formulation of the electro-fink scenario, the fink acts instantaneously, thereby blocking the time-gap solution (Martin, 1994, p.3). Lewis thinks the time-gap is not of importance, and the fink need not act instantaneously for the counterexample to be problematic (Lewis, 1997).

Mumford and Anjum have proposed introducing a *sui generis* modality that is stronger than contingency but weaker than necessity to deal with these cases. Under this modality, exceptions would be permitted. However, it seems metaphysically dubious to introduce a different modality just motivated by counterexamples that are in a way accidental. I will say more on this in Chapter 4.

2.3.4   Arrangements

All approaches to counter finks seem to fail in some respect. I think it might be impossible to solve the puzzle of finking, masking and mimicry through changing our method of analysis. Since finking is a special case of the conditional fallacy, and we can always come up with cases where the antecedent of the conditional that is supposed to analyze the disposition makes the analysandum true, there might not be a solution in the strain attempted above. If changing the analysis doesn't work, maybe we just need to restrict dispositions to certain types of manifestations or manifestation conditions. One such approach is Martin and Heil's idea that dispositions only manifest when they meet a *manifestation partner.*

Heil says powers have to act on something. The negative charge of an electron attracts a positively charged particle, which means there is a mutual manifestation of powers. On the electron side we have the power to attract positively charged particles, on the positively charged side we have the power to attract negatively charged particles. On the Martin and Heil conception of powers, mutual manifestation partners are needed for a power to manifest itself.

This mutual manifestation partner needs to have a power that enables the first power to manifest itself. This power can, but need not be the same power as in its partner. For example, one electron can manifest repulsion when in the vicinity of another electron, as they both have the power to repel negatively charged objects. However, an electron may manifest attraction when paired with a partner that is positively charged – has a different power. Mutual manifestations are not restricted to two powers. Multiple powers may be involved in manifestations.

Heil thinks finking and related problems can be overcome by two arguments. First, he claims that nothing is a fink by nature. This can be understood in two ways. On the one hand it can mean that all finks are extrinsic:

> Describing a system as finkish, or describing something as an antidote, an inhibitor, a preventer, a blocker–or, for that matter, as a facilitator–is to deploy what Descartes calls 'extrinsic denominators'. Nothing is, of itself, any of these things (Heil, 2012, p.130).

Although usually Heil's style of non-technical writing and broad historical context is pleasant, in this case it is problematic that he doesn't engage with newer and stronger counterexamples. For Choi has come up with finkish cases where the fink is intrinsic to the object (Choi, 2005). If there are such things as intrinsic finks then this argument of Heil's does not hold.

If we take Heil's assertion not to mean that all finks are extrinsic, but rather that there are no finks at the fundamental level, or that there is no natural or real property that is finkish, then Choi's counterexamples can be relocated to the realm of fantasy – it would simply be claimed that such intrinsic finks are not real properties. I believe such a statement needs arguing for, since it is a substantial metaphysical claim. In Chapter 4 I will say more about this.

The second argument against finks is that, even in the case of finks, everything acts as it is supposed to on the basis of the powers that are present.

> When you look at finkish systems dispassionately, when you consider them as physical systems rather than as counter-examples or potential counter-examples to philosophical theses, they manifest themselves exactly as they ought to manifest themselves given their dispositional composition (Heil, 2012, p.129).

The wire in the electro-fink scenario may be live, and yet not conduct electricity when touched with a conductor, but if this happens, it is the result of other powers present in the scenario. In this case the powers of the electro-fink, the circuit breaker. (Heil, 2012, p.129) I like to put it this way: the *arrangement* of powers in the scenario is such that the wire does not conduct electricity when touched with a conductor.

When we try to see the conductor and the fink that's attached as a single system that is disposed to do whatever the conductor was disposed to do before we attached the fink, it seems that the conductor fails to manifest its disposition even though the manifestation conditions obtain. However, when we look at the conductor+fink system as a single system with a certain arrangement powers, then it works exactly the way it is disposed to work. Wires with finks attached to them simply have different dispositions that wires that do not have finks attached. The presence of the powers of the fink change the manifestation of the power of the conductor.

The move on Heil's part here seems to be to look at the entire situation rather than at the individual disposition of the wire in the fink-example. When we take all present dispositions into account, things happen just as they should. However, the counterexample of finking was aimed not at the total state of affairs, but rather at the conditional analysis of the single disposition of the wire. Finks were just a specific version of the conditional fallacy of failing to see that the antecedent of the conditional analysis influences the truth of the analysandum. By looking at the whole picture we include this relation between the truth of the antecedent of the analysans and the truth of the analysandum. If we do that, the outcome of the example is not surprising. We now see that the coming close of the conductor, that is, the manifestation condition of the analysandum *causes*, by route of the electro-fink, the falsity of the analysandum. So if we take all causal connections into account, the fink isn't finkish.

But this seems to do no justice to the counterexample. The idea of the counterexample was not to spell out a physically impossible situation, but rather to cast a doubt on the conditional analysis of dispositions. Now we may say that the mutual manifestation strategy of Martin and Heil gives us a way of spelling out the conditional analysis so that finking is no counterexample. For if we say that adding an electro-fink to the wire

scenario is simply adding another mutual manifestation partner to the arrangement, a manifestation partner that *prevents* manifestation of the disposition to conduct, then we have a way out of the finkishness.

Or do we? If we want to spell out exactly what a specific power is on the mutual manifestation model, we must give an account of what manifestations the power would have when paired with different partners. For example, to spell out what the negative charge of an electron is, we must spell out that it will manifest attraction when paired with a positively charged partner, and manifest repulsion when paired with a negatively charged partner. But how would we exclude all blocking partners such as finks from the list of positive manifestation partners without having to take recourse to just saying they are blocking partners? Is there some general characteristic that manifestation partners that block manifestation have, but that partners that facilitate manifestation don't have?

So while the strategy of appealing to arrangements of mutual manifestation partners might be successful at countering the finkish and related counterexamples, it is not prima facie so. A further account of the manifestation partners of specific powers needs to be given in order for the strategy to be successful across the board.

A small problem for the mutual manifestation account would be if there were dispositions that manifest in the absence of any power. What if there are powers that do not need reciprocation? But I think the mutual manifestation account does not exclude the possibility of powers that will manifest themselves *unless* some partner that prevents it from doing so occurs. Some powers might not need mutual manifestation partners, even on the Martin and Heil view. The decay of radioactive particles could be the product of a power that needs no manifestation partner.

Chapter 3

Causes as Powers

In the previous chapters I have constructed an ontological framework in which we can build a metaphysics of causation based on powers of objects. This ontological framework consists of fundamental objects, fundamental and essentially powerful properties of these objects, manifestations, and manifestation conditions. In this chapter the goal is to put all these parts together into a theory of causation.

3.1 Causes are powers

On my view, causes are not of the same kind as effects. The causes are powers – the effects are not. An instance of causation consists of at least one object with a powerful property, the manifestation conditions of that power obtaining, and the manifestation of the power, which is an effect. As such, an instance of causation on my view has the following form:

$$p \rightarrowtail e$$
$$\vdots$$
$$m$$

where p is a token power, which specifies a power type which specifies a manifestation condition type and a manifestation type. The m represents a token manifestation condition of the type M, which belongs to the power. The holding of the manifestation condition in combination with the power, causes the manifestation e, which is a token of the manifestation type E. Since the holding of the manifestation condition is an event, and events have spatiotemporal location, by specifying m we have specified that the instance of causation happens at a certain time. In this way we can individuate the repeated manifestations of a single power.

So causes involve objects and their powers. Since the powers are instantiated in the objects, and the instantiations are individuated by the objects, we need not make direct reference to the objects in spelling out the causal relata. There is no need to say that it was the power type P, which is instantiated in object type X that caused effect type E. For one, it seems mistaken to say the power *type* caused the particular effect, rather than its instantiation. This could be worked around by saying that it is the power type P *as instantiated* in x that caused E. But why make separate reference to the type and the object, if we can simply refer to the power-instantiation p of power type P in object x. To say P caused E would be a category mistake not only on the side of causes, but also

on the side of effects. For causes don't cause effect types, but specific token effects. This is why is argue that the proper analysis for a single instance of causation is $[p, m] \rightarrowtail e$. [1]

3.2 Manifestation conditions

If we believe that all instances of causation are driven by powers, and powers are fundamental dispositions held by fundamental objects, then the manifestation conditions must be conditions that function at the level of these objects. The manifestation conditions in conjunction with the manifestation types describe how the powerful objects behave or would behave. As such, manifestation conditions and manifestation types are closely related to natural science. Some branches of physics are concerned with the properties and interactions of fundamental particles, and with what particles are fundamental. In formulating proper manifestation conditions and manifestation types of real powers that exist in our world, we need to take into account the discoveries of these branches of science.

On my view, the fundamental powers that all causation derives from are microphysical properties of the stuff of nature. For example, the negative charge of an electron is the power to attract positively charged particles, and to reject negatively charged particles.[2] Or the mass of an object is its power to attract other objects with a mass. The manifestation conditions of these powers are given in the laws of nature associated with them. If, for the sake of simplicity at this stage, we accept that Newton's gravitational force equation gives a correct account of how massive objects interact qua massy, then we may use it in our account of the power of mass.

$$F = \frac{Gm_x m_y}{r_{xy}^2} \qquad \text{NEWTON'S GRAVITATIONAL FORCE EQUATION}$$

Now this equation does not describe a single state of affairs of which we can say that it obtains or does not obtain. Neither does it describe a type of states of affairs of which we can say that a token obtains *now*. So it is not obvious that the equation can function as the manifestation condition for a power, since powers were analyzed in Chapter 1 as being of the form:

$$x \text{ would } E \text{ under circumstance } M.$$

Simply replacing M in this formulation by NEWTON'S GRAVITATIONAL FORCE EQUATION will not give us an intelligible result.

The manifestation conditions of the power of gravitational attraction consist in the *filling in* of the unknowns of the equation that describes the law of nature. That is, the manifestation conditions are the values of the variables and constants in the laws of nature. NEWTON'S GRAVITATIONAL FORCE EQUATION states the gravitational power of object x over object y. The manifestation of this power depends on:

- Object x having a mass, that is, filling in the variable m_x

---

[1] On some notions of universals, such as Armstrong's, the universals P and E could be wholly present in the instantiation. This would make a statement like P caused E less objectionable than on a Platonist or radical class nominalist conception of universals. But still it seems to me unnecessary to appeal to the universalisation of this particular way this particular object is in order to explain this particular instance of causation.

[2] This is a single-track power, because a single description (Coulomb's law) can describe both the attraction and the repulsion. Negative charge is described by the variable for the charge of the object in the description of the manifestation taking a negative value.

- Object y being at a certain distance from x, that is, filling in the variable r

- The gravitational constant G taking a certain value.

- There being another massive object, that is, filling in the variable m_y

So we may say that the object x has an attraction F of $G\frac{m_x m_y}{r^2}$ over y when x has a mass m_x and y has a mass m_y and the distance between the objects is r and the gravitational constant is G. As such the manifestation conditions of the power of x lie in its own mass, the mass of y, the value of G and the value of r. This result is something we can input in the power. The manifestation condition type of the power type mass could be spelled out as the conjunction of the following requirements:

1. Object x has a nonzero mass

2. Object y has a nonzero mass

3. There is a nonzero distance r between x and y

4. The gravitational constant is not 0.

If these conditions hold, and there is an object x with this power, then the effect will occur.

3.3 Effects are manifestations

On the other side of the causal relation we have the effect. It is a dogma of the received view that effects happen *after* their causes, and that causation is asymmetrical. Some examples of causation do not seem to fit the received view well. Many cases involve simultaneous causation, and situations remaining stable rather than changes following one another. If we see a chair with legs that stand at an angle to the floor, we see a case of simultaneous causation. The seat connects the legs in such a way that they don't fall over, while the legs support the seat. The causation seems to be going both ways. And the effect and the cause do not seem to follow each other, but rather to exist simultaneously. This simultaneous causation involving powers that balance each other is a paradigmatic case that a good theory of causation needs to be able to explain. Causal powers not only explain why the world is a dynamic mosaic, but also why there is stability and persistence of situations. The example of gravitational attraction is perfectly suited to show this: attraction is alway mutual, continuous and simultaneous. It is my view that such manifestations may also change in character while remaining a single manifestation. I will argue for this in 3.3.1.

With the manifestation conditions of the power of gravitational attraction defined above, let's continue with the manifestation of the example power. The basic idea of my account of manifestations is that manifestations of powers are the forces these powers exert. On our example, the manifestation e of the gravitational power p of x is the force x exerts on y. Thus the manifestation type can be described by the gravitation law:

$$E = \left(\frac{Gm_x m_y}{r^2}\right)$$

where G, m_x, m_y and R all stand for specific values. But of course, in the singular instance of power p, the effect is not the manifestation type E but rather the manifestation e.

There is a further problem with this notation: the variables of this equation take numbers as values. Which means that what is on the right side of the equation is a number. The identity sign then gives us the result that e is a number. But a manifestation can surely not just be a number? We need to make some distinctions in order to avoid this category mistake. We also have to be careful not to make the mistake of thinking that $m_x = p$. This would also be a category mistake since m_x is just a number attached to the mass of x. Just as 42195 meters is not identical to the distance between the battle of Marathon and Athens. Rather, 42195 meters is the number (measurement) attached to that physical distance. We could have a measurement function from p to m_x in order to link the number to the power.

In order to improve the formulation of the law, we need to specify that e is taken to stand for a force, an actual something that can be a manifestation. The right side of the equation merely gives the strength of the force. The relation therefore between the right side and the left side is not one of identity. I will therefore use a different symbol ($=:$) to denote the relation between a manifestation and the equation that defines it.

$$e =: \left(\frac{G m_x m_y}{r^2} \right) \qquad \text{\textsc{Manifestation as force}}$$

Yet this notation is still not complete, for a force has not only a strength but also a direction.

$$e =: \vec{F}_{xy} = -\frac{G m_x m_y}{|r_{xy}|^2} \, \hat{r}_{xy} \qquad \text{\textsc{Manifestation as force}}$$

This notation determines the direction and length of the vector that represents the force. The negative sign indicates that it is always an attractive force. Again, the manifestation is not equal to the vector, since a vector is just a number with a direction. Rather, the effect is represented by it.[3]

The manifestation of the power p that x has is the attractive force that x exerts on other massive objects. The manifestation type specification tells us how great that force will be, depending on the circumstances, and what the direction will be. If the mass of x takes a greater value, the force x exerts on other massive objects will increase, etc.

Modern fundamental physics doesn't usually deal in forces, but talks about fields or gauge bosons to fulfill the function I have now ascribed to a force. For now I will stick with my example that uses forces in order to keep things clear and simple. However, a theory of powers that is fully integrated with (some part of) physics would probably talk about manifestations as fields rather than forces. To integrate powers into the Standard Model of particles, we may need to ascribe most powers to gauge bosons or say gauge bosons are an explanation of how powers *produce* their effects. This thesis is too short to attempt such an integration.

3.3.1 Are manifestations events?

Notice the manifestation on my view isn't the objects moving towards each other, or accelerating in the direction of the other object. The force originating from one object and acting upon another is the manifestation of the power. The connection between the mass and the force is tight – there is no way for an object to have a mass, be at some

---

[3]Mumford and Anjum (2011, ch.2) represent causes as vectors whereas I use this representational device for manifestations. I find their representation of causes as vectors rather metaphorical. It is motivated mostly by the idea that a power may not be strong enough to have an occurrent effect and that powers combined might be stronger that the minimally sufficient force for the occurrent effect to happen.

distance from another massive object, and not exert this gravitational force. In other words, the force of massive objects cannot be prevented. The effect of the causal power of mass, in combination with the manifestation conditions of their being another massive object around, make the effect. The effect is the force that x exerts on y at the time that these values of the manifestation conditions hold. So the effect of this specific instance of causation is the exertion of F by x over y.

Superficially, that seems to make the effect an event, as events are the usual category in ontology in which to put *happenings*. The holding of a force between two objects is like a happening. But what exactly is happening in the case of two massive bodies attracting one another? There is a force that has a determinate strength and a direction, which can be read off from the instantiated manifestation conditions.

But is the power of x to attract y individuated by the specific values that enter into the manifestation conditions? That is, if x and y are moving away from each other (for example, under the influence of electromagnetic repulsion), does that mean that the gravitational attraction of x over y at time t is a different manifestation from the gravitational attraction of x over y at time $t + 1$, which is a millisecond later than t?

If the manifestation given in the definition of the power referred to a single token manifestation, then the power to attract y would be a power distinct from the power to attract x. Maybe the power to attract y at t_1 would even be distinct from the power to attract y at t_2. This would mean imply an enormous number of powers in x, many of which are very similar. Also it seems unlikely that the property of x to attract y at some specific time t in the future is a real natural feature of x. Rather, we should say that x has a power to a certain *type* of manifestation, so that we may say that all these different token manifestations are consequences of the same power of x.

When a power manifests, what happens is not a manifestation type, but rather a concrete manifestation, a token. It is a token of the type referred to in the power description, that is manifested, made concrete. A power's manifesting can be seen as the instantiation of a manifestation type. That seems to mean that if the nature of the manifestation changes, there are really multiple manifestations of the same power going on successively. If the manifestation at t and at $t + 1$ are of the same type, then maybe we can say the power remains at work and there is an infinity of manifestations falling under the same type. But this is not ontologically parsimonious, since we are left with an infinity of manifestations any time a power manifests for longer than an instant (in a dynamic situation).

If we try to avoid infinite manifestations resulting from a single power by saying powers, we might be tempted to think powers can only be properly attributed only to times slices of objects. In this scenario a power of the time slice would cause a time slice manifestation. But that would be a mistake. First of all, even if a variation in distance means a distinct power, that would not imply powers can only be attributed to times slices or objects at specific moments. For in a non-dynamic situation, the power could reasonably be attributed to an object during a certain span of time that is longer than a moment. If the distance between x and y were to remain unchanged for the duration between t and $t + 1$, and start to change only at $t + 1$, then we could still attribute the power to the object for the duration of t to $t + 1$. So we have the option to individuate manifestations as wide as the static situation permits us.

Yet, I still think it would be a mistake to individuate manifestations so narrowly. For if powers are intrinsic properties of the objects, and the manifestation condition type features variables rather than definite values of the distance between objects and their respective masses, then it seems strange to have manifestations that do not allow for

change of the value of these variables. How can something extrinsic like the distance between two massive objects matter to the individuation of the manifestations of these intrinsic powers? The manifestation conditions only demanded that there be a nonzero value to the variables given in the law. So there will be a manifestation any time there x has a nonzero mass, y has a nonzero mass, and x and y are a some distance from each other.[4]

The most elegant option, I think, is for a single manifestation of the power of mass not to be bound to a single value of F. If we go this route, the massive power of x will manifest as long as x has a mass and there is another massive object around. The nature of the manifestation may change, as the force may change in direction or strength depending on the circumstances. The *quantities* involved in the power need not be stable. This makes manifestations an ontological category of exceptional flexibility.

Imagine x and y having a stable and equal nonzero mass :

$$m_x = m_y \neq 0$$

And imagine y moving through space relative to x such that the distance between x and y over time is described by a sinusoid curve. In this case, the force of attraction of x over y would vary in a derivative manner. In this case there would be one power with one manifestation, since the manifestation conditions hold continually and the power manifests continually. But the manifestation of a power changes in nature over time, since it changes in strength. And perhaps the distance between x and y changes sinusoidally because y follows some kind of elliptical path around x. In that case the direction of the force would also change. We can represent the force as an arrow originating from x, pointing towards y, and going around in circles following y. At the same time the arrow is getting longer and shorter continuously. On my view, x is 'doing' one manifesting, as long as there is an arrow. That means, the power manifestation only stops when x's mass drops to zero, or when y disappears from the scenario, or when G becomes zero, or when the distance between x and y becomes infinite.

Now do we still think the effect is an event on this account? Say that x and y are both positively charged. They would repel one another as long as they remain charged. Since our example universe contains only x and y we will assume that they remain charged.[5] The gravitational attraction between x and y also persists. Now we have powers of gravitational attraction counteracting powers of electromagnetic repulsion. If these are events, then they are co-instantiated spatiotemporally. According to the most common method of individuating events, this would mean they are identical. That, to me, is an absurd conclusion. Gravity and electromagnetic repulsion are very different power types, and their manifestations ought to be distinguishable. On another method of individuating events they are instances of properties in objects at a certain time (Kim, 1993). On such a view of events are close to states of affairs. This is an acceptable option for what effects are, as long as we allow the times that individuate events to include regions of time and not just time slices. And we need to have a view of properties that

---

[4]Whether this distance must be non-infinite and/or nonzero depends on some assumptions about the physics of x and y. For example, if x and y were bosons, they could occupy the same state, since bosons do not obey the Pauli exclusion principle. But, as we are using Newtonian mechanics as an example here, we will simply assume for now that x and y simply do not occupy the same place in space, and are at a finite distance from one another.

[5]It is impossible for there to be nothing but two charged objects in a universe, since mutual repulsion of positive charge is electromagnetic interaction, which is mediated by photons. Since photons themselves are not charged, if there are two positively charged particles repelling one another, then there must be photons. For the sake of simplicity we will ignore this fact.

allows for them to change over time, as I have argued. Note that for Kim the notion of a property is closer to what I would call predicates, and two properties are identical according to Kim whenever they are co-extensional. While Kim's idea of events might fit with my account of effect, it would need to be loaded with a different view of properties.

Intuitively the word 'event' seems more suited towards dynamic situations – disturbances in the equilibrium, so to say. I want my account of causation to be able to explain static situations as well as dynamic ones. The equilibrium needs causes to remain in balance (at least sometimes). If events are to be effects, than we have to have a method of individuating them that fits well with this idea of spatiotemporally co-instantiated happenings over time that can be either dynamic or static.

In spelling out what powers are, I said they are connected to certain types of manifestations. The idea behind this thesis is that a single power can manifest at various times or can even manifest more than once in some cases. For example, the massive power of x is a power to attract y categorically, meaning that it is not a power to attract y at some specific time, or just for one moment. It will always attract y as long as the manifestation conditions hold and x retains the power. Although I don't think this *continued* attraction implies that there are multiple token manifestations, *repeated* attraction would. Also, if x attracts both y and z by being massive, then there seem to be two manifestations of the same power at the same time.

Perhaps the term 'events' isn't very well suited for effects. Effects as I conceive of them can be extended in time eternally, can be spatiotemporally co-instantiated, can vary in nature over time. When in need for a term that is closely related to events, but not quite the same, philosophers sometimes appeal to *processes*, but that also doesn't seem adequate here, since processes involve change and have some kind of *end*, whereas an attractive force obtaining might not great change and might not have a final state, like a process intuitively does.

If manifestations are not events, what are they? On Armstrong's view, effects are states of affairs.[6] This position is somewhat attractive since states of affairs allow all kinds of entities and can vary from very simple to very complex. They might even be extended and variable through time, which would fit well with manifestations. However, states of affairs are rather hard to individuate. For which parts of the way the world is belong to the effect of *this* causal interaction? How do we limit the size of effects if they are states of affairs? I find it hard to think of any criterion that could limit the state of affairs to only the relevant parts of the world that were affected by this instance of causation, without appealing to a pre-established notion of what the effect of this particular cause is.

I can't think of a category that is commonly invoked in analytic philosophy that would fit well with my account of manifestations, but I think any category that would fit the bill should allow for the following traits:

- Extended in time

- Spatially co-instantiated with others of the same category

- Qualitative change over time (or not)

- Involving external relations between objects

---

[6] I am fairly optimistic about the possibility of using states of affairs as the category for manifestations to fall under. Armstrong's notion of a state of affairs is closely related to Kim's notion of events.

Since powers work on objects and their properties when they manifest, a viable candidate for the category of manifestations could be something like a collection of property-states over time. For example, the effect x has on y in our gravity example is a variable attractive force (which is an extrinsic property of y) over a period of time. The 'snapshot' of the attractive force on y over time is the effect of x's massive power. A final option is to count manifestations as their own category in ontology. This would require more argumentative work, but I have given some pointers as to what to look for when specifying the category.

3.4   Symmetry of attraction

In the case of gravity, the power of attraction x exerts on y is always equal in force to the power that y exerts on x. The direction of attraction is opposite. This is easily seen when we take the vectorized version of the gravitational law and switch around every occurrence of x and y. Since the multiplication and division that occur in the law are commutative, the resultant strength of the force is equal in both directions. The vector direction comes out opposite when we switch around x and y, for obvious reasons.

For any case in which there is an object manifesting a power of gravitational attraction, there is a partner object also manifesting this power. This can be read off from the manifestation conditions of the power. Take the power p of x to E when M. If this power is manifesting, that means the manifestation conditions are instantiated, and so is the effect. Therefore we have an instance of causation that looks like this in my notation:

$$p \rightarrowtail e$$
$$\vdots$$
$$m$$

Contained in the power is the value of the mass of x which we could name m_x. In a manifestation condition type there is an indefinite designator that says there must be some object with a mass other than x for there to be gravitational attraction. In the instantiated manifestation condition there is an actual object that has a mass. This object is named y. There is reference not merely to the object, but to its mass m_y. Also from the instantiated manifestation condition we can read off the distance r between x and y, and we can calculate the vector direction. As such, the power that y has to attract x, together with its magnitude of force, and its direction, can be read off from the complete formulation of the manifestation of p. Since all this information is contained in p, it is necessary that in every case of gravitational attraction there is an attraction of equal force and opposite direction originating at y.

This necessary interconnection of the manifestation of the power of one object with the power of another object brings out an interesting and attractive feature of the metaphysics of powers. Since powers are directed towards their manifestation, they point outwards. Powers supply metaphysical links between an object and the rest of the world. The idea that powers are directed at their manifestations has sometimes been linked to Brentano's concept of *intentionality*. Whether this is a good conceptual match or not, it is true that dispositions are *directed* at their manifestations. But some powers, such as gravitational attraction, are not only directed at their manifestations, but also at other powers, as I have been claiming in this section. This makes the power not only intentional in itself, but also directed towards other objects with powers. As such the picture that emerges from this powers view is un-Leibnizian; objects are not windowless monads, but

rather elements connected with other elements.[7] On the neo-Humean account, properties and objects are disconnected like monads.

3.5 One variable, two laws

So far I have been talking about the mass of x as identical with the power to attract other massive objects. However, in Newtonian physics, mass not only figures as a variable in the gravitational law, but also in the second law of motion:

$$F = ma \qquad \text{ACCELERATION LAW}$$

It seems that if this law is about powers, then this is a different power than the one described in NEWTON'S GRAVITATIONAL FORCE EQUATION. There is a puzzle here: if mass is a power, then how can it figure in different laws that describe different manifestation types and manifestation conditions.

My answer is that the law $F = ma$ does not define a power or a manifestation. The ACCELERATION LAW does not describe a manifestation, but rather the occurrent behavior that results from a certain manifestation. As I have argued in 3.3, it is not the occurrent behavior, i.e. the moving of objects, that form the manifestations of powers. Rather, the manifestation of the power of gravitational attraction is the force a massive object exerts on another massive object. This idea creates a gap between what effects are and what we can simply observe. By saying powers have forces as manifestations, we have posited a level of manifestations between the level of powers and the level of occurrent effects.[8]

3.5.1 Occurrent effects

If we have a situation where one power is perfectly balanced by another, we see no movement. The *net force* is zero. But how do we explain the fact that the powers do not actually cause anything to move or change? In my view, this is not strange at all. Powers are not only to explain instances of occurrent dynamic situations, but also of occurrent stability. But there is more to say: the occurrent effect of a force of strength x and a force of the same strength in opposite direction is that nothing moves. But why wouldn't the occurrent effect of this situation be that the object moves sideways or that it disappears? Of course, we know how to add forces: it is vector addition. [9]
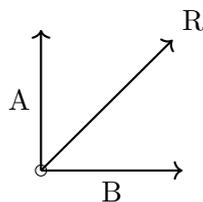
If we add the vectors representing forces A and B working on the object, we get the resultant vector R, which describes the net force. The philosophical question here is what

---

[7] (Leibniz, 1714)

[8] Note that manifestations are the consequence of a single power while occurrent effects may be the consequence of multiple powers. Compare (McKitrick, 2013).

[9] Cartwright is worried that vector addition is something *we* do, and not something in the world. On her view the component forces are only metaphorical, and only the resultant force is real. To me this is an absurd position. While the component forces do not have their direct occurrent effects, their reality is what explains the resultant force and thereby the resultant occurrent effect. See Cartwright (1983, p.58–59) and Mumford and Anjum (2011, p.39–44).

the status of these rules of vector addition are. A simple answer is just that it is a brute fact that forces combine according to the rules of vector addition. But that would make it pure coincidence that they always do that. Another answer is that there is a governing law that makes it the case that occurrent manifestations are like resultant vectors. But one of the great upsides of a powers metaphysics was to do away with governing laws and instead have laws be made true by powers. So to accept governing *combination laws* that make certain combinations of manifestations result in certain occurrent effects, would be to give away much that we have gained.

The right way out of this problem seems to me to say that not only have powers essential connections to manifestation types, but also that manifestation types have essential connections to occurrent effect types. The rules of vector addition are contained in what a force is. It is part of the essence of a force that it can be canceled out by an opposing force of equal strength, etc.

Now back to the ACCELERATION LAW. The F in the acceleration law denotes not a manifestation of a single power, but rather a resultant vector. The acceleration law does not describe the manifestation of the power of mass, but rather gives us a description of the occurrent effect. Again, to avoid scary ungrounded governing laws, we must accept that the ACCELERATION LAW is a part of the essence of manifestations.

A result of this position is that, in a way, NEWTON'S GRAVITATIONAL FORCE EQUATION is a different kind of law of nature, since it is grounded in how powers manifest, while the ACCELERATION LAW is grounded in the essences of manifestation types. I think this actually concords nicely with what the ACCELERATION LAW seems to be even before considering powers and grounding laws of nature.[10] Newton's three laws of motion don't seem to say anything about what objects interact with each other or what forces can be exerted. The laws describe what happens *when* a force is exerted on an object (and when it is not, in the case of the first law of motion). Also, that the F in NEWTON'S GRAVITATIONAL FORCE EQUATION denotes a single force while the F in the ACCELERATION LAW refers to a net force van be seen to be true independently of what you believe about powers, their manifestations, and their essences. Of course, nowadays the laws of motion have been replaced by more generally valid laws of conservation of momentum, angular momentum, and energy. But, like with the gravitational law, Newton's laws of motion are merely an example for my theory. I think the more general conservation laws are also laws of the type that are grounded in the essences of manifestation types rather than directly in powers.

3.5.2 Linking values across powers

The question how m could appear in two different laws has been solved by appealing to the difference between manifestations and occurrent effects. But we can imagine a harder case: what if one variable were to appear in two totally unrelated laws, both of which are grounded in powers and describe manifestations rather than occurrent effects. We could, for example, imagine that the mass of an object not only enables it to exert force on other massive objects, but also gives it the ability to attract photons (which have no mass).[11] In this case we could not appeal to the distinction between manifestations and

---

[10]Lewis Creary has argued the division between laws about forces and laws about results in (Creary, 1981). Cartwright rejects this view on the basis that many laws science actually uses don't allow for simple vector addition(Cartwright, 1983, p.63-73). I believe my strategy of allowing only fundamental powers avoids this pitfall to a large degree.

[11]This is of course a fictitious example. On most actual cases where a single variable appears in multiple laws, the laws are related in some way. This fictional example is designed to create the hardest case

occurrent effects, since attracting photons is a manifestation and not an occurrent effect.

One way to solve the puzzle of multiple laws featuring the same powers is to say that the laws do not really refer to the same property of objects. On this option, the m in the gravitational law is not the same as the m in the photon-law. This would solve the problem neatly: different laws refer to different properties and each power-type is neatly described with the help of a single law. But there is something off about this idea: for if the mass of gravitation and the mass of photon attraction are two different properties, how come they alway take the same value in single objects? Or to put it differently: could an object have had a mass of gravitation that was larger or smaller than its mass of photon attraction? The answer to this must, I think, be a firm 'no'. For this would make it a staggering coincidence that every object in our imaginary world has the same mass of gravitation and of photon attraction. It could even be said that this solution is just denying a premise of my problem: that the *same* variable appears in both laws. Denying there is such an actual case is fine, but denying that there is a possible case in which a single property enables an object to do two totally unrelated things seems somewhat dubious.

A second option for solving the hard puzzle is to appeal to multi-track dispositions. Some dispositionalists believe in dispositions that can have different manifestations under different conditions. I find it hard to see how multi-track dispositions would be individuated. On my account powers are individuated by appeal to their manifestations and manifestation conditions. I find it both unlikely and metaphysically sloppy that these conditions or manifestation types could be disjunctive. On the Martin and Heil account of powers such a situation would be possible, since the power could manifest in very different ways when paired with different manifestation partners (i.e. an photon or a massive object). I think this opens the door to having powers that do wildly different things under different circumstances. Thus the individuation of powers can become much too coarse-grained on such an account. The beauty of having powers with essentially connected manifestations is that powers are powers to do certain types of things. If the type of things a power can do becomes too diverse, I think the account loses metaphysical plausibility as well as explanatory power. Although I have already acknowledged that powers can manifest differently when paired with different other powers (such as with a larger and smaller mass), these manifestation must remain similar to a high degree, and the manifestation partners that enable a single power must also be highly similar. Therefore my solution does not involve multi-track dispositions in the sense that very different manifestations can be the result of one power.

My proposed solution to this problem is to appeal to a property of the object that is not identical with the power, but which gives the value to the power-variable. This is a true categorical property since it does not have powers. Its only role in this metaphysics is to 'bundle' different powers and ascribe them the same value.[12] If we go back to the example of mass enabling both attraction of massive objects and of photons, we can see how the solution works.

The idea is that there is a categorical property $\mathfrak{M}$ that underlies both the power to attract photons and the power to attract massive objects. The claim is then that there is a function C from the power p to the categorical property it belongs to.

---

possible, where the two laws are unrelated besides sharing a variable.

[12] Mumford and Anjum (2011) have argued that properties are bundles of powers. Their reason for thinking that is simply that properties such as 'being hot' provide an object with the power to warm you, to make water boil, etc. So it seems that their power-bundles lead to a much thicker notion of properties that underly the power than my idea of value-fixing does. I think this is a product of their acceptance of non-fundamental powers and properties.

$$\mathcal{C}_p = \mathfrak{M} \hspace{4cm} \text{\textsc{categorical function}}$$

We can posit there is an underlying categorical for every power. Possibly, some powers share their underlying categorical, and therefore variables in the power-description necessarily take the same value. This would be the case for the mass attracting power p that x has and the photon-attracting power q it also has. The mass-variable m appears in both the manifestation types of p and of q, and it necessarily takes the same value which is provided by $\mathfrak{M}$. On this scenario what it means for an object to have mass is to have both the powers p and q and also the categorical $\mathfrak{M}$.

The downside of this solution is that where before there were only powerful properties in the ontology, now we've added categorical properties that underly the powers. But these categoricals are properties on a very thin notion of properties: they don't do anything except provide equality of values between some powers. If we accept this solution we need to consider whether powers are still properties, or whether they are subsumed under these categoricals, and the categoricals are the only real properties. I have no definite answer to the problem of multiple laws sharing variables. I think the distinction between manifestations and occurrent effect takes us a long way, but this harder problem shows that it might not be all the way. Solving the problem all the way may be ontologically costly because we need necessary connections between powers that are otherwise disconnected.

Chapter 4

Aspects of causation

Over the course of the previous chapters I have developed an account of causation in terms of powers. I have contrasted this account's metaphysical basis with neo-Humean accounts of causation, given the elements needed in our ontology in order to build this account, and shown how these elements combine to give an account of causation. In this chapter I will explore some philosophical consequences of my way of conceptualizing causes and effects. Also, I will tentatively explore some further lines of research that are closely related to the above, such as the grounding of modality and probabilistic causation. Perhaps most importantly I will show in this chapter that the specific account of powers I have given is not prone to the most powerful criticisms of dispositional accounts of causation: the problems of finking, masking and mimicry.

4.1 Powers without finks

In Chapter 2, I listed a number of problems for dispositional accounts of causation. These counterexamples are supposed to show that it is sometimes possible to have a disposition and a manifestation condition, but not get the manifestation. If such situations can be real, then powers do not necessitate their effects. There are various strategies for countering this point. The first group of strategies concedes that there are finks, masks or mimickers, but tries to find a way to make dispositional account of causation work without the necessary connection between power and manifestation. Such approaches can proceed through a non-conditional analysis of dispositions, effectively abandoning what is known in the literature as the DSM-model of dispositions. The DSM-model is the idea that a **D**isposition comes with a **S**timulus condition and a **M**anifestation. I endorse the DSM model (although I use other letters) and therefore this option is not open to me. Also I think it unnecessary to take this route. Another way of countering the problems with finks, masks and mimickers, is by replacing the necessary connection between disposition and manifestation with some other connection. This is an option endorsed by various authors who say dispositions merely *tend* towards their manifestation, or who say there is a different modality that lies somewhere between pure contingency and necessity that describes the connection between disposition and manifestation. I fear this second strategy either shoots itself in the foot, or is metaphysically unacceptable. The *tending* strategy shoots itself in the foot because if dispositions merely tend towards manifestations, there is no guarantee that anything will get caused. On a theory of causation where things merely tend, all causation in the universe could stop if we hit a stroke of bad luck on the dice roll that merely tending is. Also, causal explanations would lose much of their value. We could say a power p caused e, but if the presence of p

and its manifestation conditions is not sufficient for *e*'s occurrence, then the explanation that *e* happened because of *p* does not tell us that *e had* to happen because of *p*. Also, we lose counterfactual-supporting qualities of the causation relation.

Some of the same arguments apply to the other option: accepting a new modality for dispositional production. Besides worries whether this modality would give sufficient support to explanation and counterfactuals, there are other worries. The immediate reaction I have to the idea of there being some modality stronger than contingency and weaker than necessity is that it is metaphysically unacceptable. In the first place, necessity and contingency are some of the most important notions of analytical metaphysics. Adding another metaphysical modality would be a great cost in terms of theoretical parsimony, or metaphysical commitments. But not all philosophers care much about parsimony or extent of commitment. Mumford and Anjum argue that the in-between dispositional modality is more familiar to us that either necessity or possibility. They offer very little positive definition of the modality, let alone a formal semantics. Their argument for the modality is informed mostly by the possibility of interference with macroscopic powers, and what the modality is supposed to be is enlightened mostly by analogy with intentionality and normativity. The motivation is ill-conceived because we can deal with interference in other ways. The explanation by analogy seems to me a strategy that runs the risk of explaining *obscurum per obscurius*, even though Mumford and Anjum claim that explaining dispositionality in terms of necessity is explaining the familiar by the obscure (Mumford and Anjum, 2011). Even if one believes the in-between modality is really more familiar to us, that doesn't make it any *clearer* to us.

Since I wish to keep the DSM-model, and the necessary connection between dispositions and manifestations, it seems only a direct refutation of the counterexamples will work. And this is what I will attempt to give in the next sections.

4.1.1 Finks

Finks can be put into two categories: extrinsic finks and intrinsic finks. Martin's electro-fink is an extrinsic fink. It is not an intrinsic property of the conductor that makes it finkish. It is the addition of the fink-switch (what lewis calls the circuit breaker) that makes it finkish. John Heil makes a convincing argument (discussed in section 2.3.4) against extrinsic finks. The idea is that the finkishness of the conductor only appears when we take a certain perspective on the case.

Choi has come up with a strong finking case Heil's argument does not apply to: an object that has intrinsic properties that make it finkish. The example he chooses is that of a *tricky triangle*. A tricky triangle is a triangle, and therefore has a disposition to make triangular impressions.[1] But the tricky triangle also has the tricky disposition that whenever someone picks it up with the intention of making impressions with it, it turns square. To put it more abstractly, it is an object that has a certain disposition p_1 with manifestation conditions m_1, but whenever m_1 obtains, it looses p_1.[2] I think it doubtful that we would want to say the object really has p_1 in the first place, if it is consistently tricky, but for now we will grant this point as a premise of the counterexample.

---

[1] I have altered the example from counting corners to making shape-impressions because counting corners seems to have too much to do with the disposition of the person who does the counting.

[2] Alexander Bird seems to think that, for such a fink to be a counterexample, it must retain its original power up to the moment the manifestation starts. Then there is no gap between manifestation and obtaining of the dispositions, and therefore fundamental finks are impossible (Bird, 2007, p.61). I take a different strategy of argument that does not turn on the lack of a time-gap between disposition and manifestation.

My first argument against intrinsic Finks (hereafter simply Fink) is one I call the extensional argument. It is very simple. In previous sections of this thesis I have stressed the importance of the idea that, for a metaphysics of causation, we need only causation by fundamental powers of fundamental entities. Therefore, a Fink is only a counterexample to my theory of causation, if it is a fundamental Fink.

The most widely accepted physical account of what there is and what it does, is the Standard Model of particles (SM). The interactions between particles as described by SM are to me the paradigmatic examples of causal situations. The extensional argument is simply that no type of particle on SM and no type of interaction on SM looks to me capable of being a Fink. I can simply not think of an analogous case to the electro-fink in the context of particle physics. Now I am not a particle physicist, and my knowledge of the subject is very limited. Because of this, and of the nature of the argument, the extensional argument is not very strong. But I do think it is strong enough to shift the burden of proof to those who believe finking forms a counterexample. None of the counterexamples I have encountered are fundamental finks.[3] It is up to the opponent of a fundamental powers theory to come up with fundamental finks.

Besides the extensional argument I also have an intensional argument against Finks. This argument is essentially the rejection of the premise that the tricky triangle really has the disposition to make triangular impressions. My intensional argument against intrinsic finks is that if it has intrinsic properties that make it so that it consistently manifests in the tricky rather than the normal way, that the object is then disposed not towards the normal manifestation, but towards the tricky one. Suppose electrons were tricky particles that really had the disposition to attract negatively charged particles. But, as soon as any negatively charged particle appears in the universe, the disposition of the electron changes to having the disposition to attract positively charged particles (which it then does). I would say it is something of an absurdity to think this tricky electron has a disposition towards attracting negatively charged particles, and not positive ones. Furthermore, if there were such a thing as a tricky electron, we could never find out that it is tricky, for we could never get it to manifest its non-tricky, normal disposition.

Please notice that the intentional argument is not that if a disposition never manifests, it is not there. I believe there might be dispositions that never manifest because the manifestation conditions never hold. But if an object's dispositions consistently manifest *differently* whenever the manifestation conditions hold, I think we should conclude simply that we were confused about the dispositions we thought it had. So if being a triangle entails making triangular impressions, a tricky triangle is simply not a triangle.[4]

4.1.2 Mimicry

In cases of mimicry the manifestation conditions of a certain power hold, and the manifestation obtains, but it is caused by some other power. Heil's argument of perspective seems to work well against mimickers, since in the case of mimicry things happen exactly as they should, if you include the actual cause of the manifestation in your perspective. I don't think intrinsic mimickers are a possibility, so Heil's position seems sufficient to do away with mimickers.

For those who don't agree, there are more arguments against mimickers. For one, the extensional argument against finks seems to apply just as well to mimickers. I think if

---

[3]I have encountered fundamental Masks. More on those in the section on Masks.

[4]If the intentional argument is successful, then the extensional argument need not carry any weight. I have chosen to present multiple arguments for the same thesis (each more or less sufficient) in anticipation of rejection of at least some parts of the arguments.

finks are not a problem for DSM-model, than mimickers certainly aren't.

4.1.3   Masking

Masks don't change the disposition of an object the way finks do, so my intensional argument against finks won't work against masks. What we *can* say of masks is that it is perspectival whether something is a mask or not. For the masking case I will use the example of Millikan's oil drop experiment. In this experiment a small drop of charged oil is suspended between two electrodes. The electrodes create a field that is strong enough to counteract the gravitational attraction of the earth working on the drop. The voltage across the plates is adjusted until the drop is suspended in equilibrium between the plates.[5] In this experiment the gravitational powers of the earth and the drop are masked by the electrical field. The falling of the drop due to gravity is prevented by the other powers at work.

My account has no trouble dealing with these masking cases. The way I spelled out what a power's manifestation is, it is the force exerted, not the movement made. Through this strategy all masking problems can be averted, since the exertion of the force cannot be blocked or interfered with. Nancy Cartwright has espoused a view similar to mine, where forces are the manifestations. The question of how forces become movements can be answered by appeal to combination rules for forces. For the forces in the Millikan experiment the combination rules would be the rules of vector addition. The problem with these combination rules is that their metaphysical status is debatable: are they governing laws in the Humean style, or can they be grounded in the manifestation types of powers? In 3.5.1 I have argued that they can and must be.

4.2   Non-reductive reductionism

As discussed in Chapter 2, my account of causation could be called reductionist, since I think all causing is done by the fundamental powers of fundamental objects. Also I have given arguments why my position is not reductionist: if there are no levels, there is no reduction. I do not believe there to be levels of reality. Also I am not advocating any revision of the way we talk about causation outside of disciplines that are concerned with ontology, with the way things really are.

These considerations are all about one sense of 'reduction': the sense in which it is used in discussions on physicalism. But there is another sense in which we may speak about reductive accounts of causation. What is meant in that case is that the concept 'cause' can be reduced to some other concept or process or entity. That is, that causation is fully explained in non-causal terms. An example of such a reductive theory of causation is a neo-Humean regularity theory. If causation is nothing more than certain type-regularities in the sequence of events, then that *a* caused *b* means nothing more than that *a* followed *b* and that *a* and *b* belong to types the instances of which always follow one another. On such an account 'causation' is reduced to regularity, which is a non-causal notion.

On my account, causation is explained by powers. These powers produce their manifestations. Some authors have insisted that powers theorists give a further story of what this production is. If we try to do so, neo-Humean answers are close at hand: maybe

---

[5]The purpose of the experiment was to determine the charge of the droplet. By doing the experiment with many drops Millikan and Fletcher discovered the charges of the drops were multiples of a certain value. In this way they estimated the charge of a single electron to within 1% of the value now accepted.

production is some kind of regularity between powers and effects, or maybe the necessity of production lies in the relation between the actual world and possible worlds, or maybe production is some higher-order relation between power types and their manifestation types. All these answers would bring us right back to the pitfalls of neo-Humeanism and devaluate what we've accomplished in constructing a theory of causal powers. Although I am not certain that any attempt to explain further the productive capacity of powers will lead to neo-Humeanism, I am afraid it will be hard to avoid.[6]

The real question a powers theorist needs to ask first is whether it is even needed to find a further explanation of the production that powers are capable of. The main reason to look for such a further explanation is that if the analysis of causation contains the analysis of powers, and the analysis of powers contains the notion of production, and production is a causal term, then causation has not been explained. The contention is that the explanans contains the explanandum. This problem is similar to Molière's joke that opium makes the man sleepy because[7] of its soporific quality. The joke turns on the idea that the doctor who gives this explanation hasn't actually explained anything and has been ultimately uninformative.[8]

Yet there is a way in which it is not uninformative to say the man is sleepy because of the soporific power of opium. If you say that the man is sleepy because of the soporific power of *opium*, you are attributing the man's sleepiness to the fact that he took opium, rather than to a lack of sleep or to watching a boring performance of a 17th century play. As such the explanation is informative since it gives contrast with other possible causes. But let's say that everybody agreed that it was the opium that caused the man's sleepiness, and that the question is what it is about the opium that causes the sleepiness.

Is it then really uninformative to say that opium makes one sleepy because it has the power to make people sleepy? I do not think so. For one, to say that opium in general makes people sleepy provides a truthmaker for subjunctive conditionals about *other* people that could have taken the opium instead of the man that did. This is because the soporific power of opium is not just the power to affect a single person at a single time but rather a power that references manifestation condition types and manifestation types. Furthermore, the explanation that the man became sleepy because of the soporific power of opium is only uninformative if you take the soporific power to be a disposition in the Gricean sense where the having of the dispositions consists *solely* in the fact that it will act a certain way in the future (or some other modal context). In that case saying the man is sleepy because of the soporific power is saying that the man is sleepy now because in the past he would be sleepy in the future. That is either thoroughly uninformative or wrong, depending on your views about the reality of the future.

Furthermore, by saying it is the soporific power of opium, and not, say some higher order relation between properties, the doctor has given us a metaphysical claim: the substance behaves the way it does because it has a certain dispositional nature. The disposition is in the opium even when it is not being consumed.

Powers are what we invoked to explain causation, and they have done so successfully in the sense that the analysis of a causal situation has changed from surface-skimming strategies such as a regularity theory to an analysis that tells us how things do what they do (because they are what they are). In conclusion I do not think that charges of uninformativeness leveled at dispositional accounts of causation are a problem for such

---

[6]The non-reductionist account gets the slightly more positive sounding name 'causal primitivism' in Mumford and Anjum (2011).

[7]I sometimes use 'because' to signify causation. In the contexts where I do so it will be apparent that I mean it in a causal way, and not as reason-giving or material implication.

[8]For a discussion of this problem along a different line than I take see Michon (2007).

accounts.

There is a further reason why I am not too worried about this objection. In my variety of a powers theory of causation, there are only fundamental causes. While in the case of non-fundamental causal claims it might make sense in many situations to ask for a further explanation, there seems little reason to do so in fundamental cases. We might want to know why the car crashed into the tree. Perhaps we can explain the fact by the speed and mass of the car combined with the traction of the car on the road surface and the direction of movement. This would be a useful reductive explanation. Now imagine a minimal universe with only two bodies exerting mutual gravitational attraction. We can explain the attraction by appealing to the mass of the bodies, and the distance between them. This is what constitutes the power on my view. If someone were to ask what it is about the mass that makes it an attractive power, I think they have misunderstood what powers are. It simply doesn't make much sense to ask the 'further' question once you have arrived at the most basic level of explanation.

If this causal 'rest product' that is production is still thought to invalidate powers as causal explanation, then I will have to be content with the fact that my account is non-reductive in the second sense mentioned above. The way I have put it, production is simply what a power does. There is no need for further explanation. The real question is not what this productive capacity of a power is, for that is just what a power is, but rather whether there are powers. I think it is a plausible thesis that there are, as I have been arguing.

4.3   Chancy events

If powers are like conditionals, then they are also determinate in that, if the manifestation conditions hold, the power will manifest. Mathematically, Newtonian mechanics is not a deterministic model, and moreover Newtonian mechanics has been superseded by relativity and quantum mechanics. In quantum mechanics, many predictions and processes are not usually interpreted as being deterministic. The example of a chancy event that is most often cited in the literature on causation is that of radium decay. Radium has no stable isotope. The most common radium isotope is ^{226}Ra. This isotope has a half-life of 1600 years, and decays into the radon isotope ^{222}Rn. When ^{226}Ra decays into ^{222}Rn it emits an alpha particle. As a stochastic process, alpha decay is not predictable, it is not a product of an external influence.

Chancy events like this are an interesting case for theories of causation. We might wonder if there is even causation if there is no external trigger of the decay. John Heil seems to take this route: he calls radioactive particles 'unmoved movers'. I think this is a viable option to a certain degree, because it allows us not to worry about these special cases as cases of causation. However, I also find it unsatisfactory to a certain degree, because I want to say that radium-226 has the *power* to decay into radon and an alpha particle. If it is a power of radium to do so, then, on my account of powers, it must have manifestation conditions.

A different option for me would be to say that the decay happens only when a certain manifestation condition obtains, but this manifestation condition is like a statistical light switch, that just turns on with a certain probability but is not predictable. This option is attractive because it retains the necessitation between power and manifestation, given that the manifestation condition holds. However, it also seems somewhat ad-hoc, especially since there is nothing we know of that could fulfill this function of manifestation conditions. Radium decay is truly random, meaning that there could be two situations

that are identical in every respect except for that in one situation the radium decays, and in the other it does not. Introducing a probabilistically obtaining manifestation condition that is not external to the radium particle seems somewhat concocted.

A further option is to say that the manifestation condition for the power to decay always obtains. We could think of the manifestation condition as something that obtains vacuously, or as something that obtains throughout our universe. *That* the power will manifest is certain, but *when* it will do so is not. But this answer is unsatisfactory for two reasons. First, introducing some nondeterministically long delay between the obtaining of the manifestation conditions and the obtaining of the manifestation is both ad hoc and doesn't really solve the problem that there seems to be no cause for the radium decay happening *now*. Second, and worse, is that probability does not give certainty. We know that for a large number of radium-226 particles, after 1600 years half will be decayed. And that gives us a pretty good idea of the chance that a certain particle will decay within a certain span of time. But that is not a guarantee that it *will* actually decay at some point.

The solution (insofar as it can be called one) I favor is to say that for certain powers, the manifestation conditions are empty, or obtain vacuously. The example for ^{226}Ra could be that object x has the power to decay to ^{222}Rn under the condition that x has the atomic number 88 (which it always has.) If the manifestation condition always obtains, why isn't the radium always decaying? My answer would be to put the indeterminacy in the productive aspect of the power. In this way we could have probabilistic powers, by saying that for some powers, if the manifestation conditions hold, the manifestation follows with a certain probability. We can express this by putting some indicator of the probability above the $\rightarrowtail$ in our graphical representations of causation. The downside of this solution is that it sacrifices the necessity of causation for a certain class of powers. However, I think philosophy should take cues from science, and if science says it is probabilistic and not necessary, then philosophy should have no qualms putting probability into the production relation.

4.4 Modality

There is a special relation between powers and necessity. On neo-Humean accounts of modality, possible worlds provide the truthmakers for modal statements. They do so either directly (the states of affairs in possible worlds are the truthmakers) or indirectly (the similarity between some subset of possible worlds and the actual world provides the truthmaker). A powers account has the advantage that powers can put modal force in the world. A metaphysics that is realist about dispositions provides a connection between facts and modality. We can say that when there is a power and the manifestation conditions obtain, then the manifestation is necessitated. In this case, the power provides a truthmaker for the subjunctive conditional that if the manifestation conditions obtain, so will the manifestation. The power is not identical with the modal fact: that could lead us straight back to Humeanism (identifying a disposition with counterfactual dependence). That powers can supply us with truthmakers for modal statements is one of the greatest accomplishments of a metaphysics containing real dispositions.

Also, since the powers are essentially connected to their manifestation condition types and their manifestations, we may make the modal claim that if an object x has some power p then necessarily it is disposed to E when M. This is a modal statement that restricts which worlds are possible, and which are not. For if some object in a possible world had p but not the disposition to E when M that would lead to contradiction, since

having p *is* having the disposition to E when M. From the essential connection between powers and their manifestation types and manifestation conditions it follows that powers are the powers to do those same things under the same circumstances in all possible worlds. If there are different manifestation conditions or a different manifestation type, it is a different power.

This position is opposed to a position known in the literature as *quidditism*. Quidditism is analogous to a position about objects known as *haecceitism*, but applied to properties. Haecceitism is the view that objects have their properties contingently (they are all accidental properties). As such, even though I am a bipedal mammal in the actual world, there is a possible world in which I am not. According to haecceitists the only property that holds essentially of objects is their self-identity. This property is called the haecceity of the object. So according to this theory, I have a property of being Allert van Westen (although it is contingent that I have that name), and this is the only property I have essentially.

Quidditism is the analogous position for properties rather than objects. Take the property in the actual world of being all red. Let's call this property P. According to quidditists, P could have been the property of being all blue. Property P is essentially identical with itself, but its further characteristics are accidental. When you extend this position to powers, the power p, which, in the actual world is a power to E when M, could have been the power to E' when M'.[9]

It will be clear from my exposition in Chapters 2 and 3 that I am not a quidditist. By individuating powers through their manifestations and manifestation conditions, it becomes impossible for a power to be qualitatively different from how it is. I see this as a great advantage of my position. For in this way powers not only explain how things come about in our actual world, but also what would happen if we put the object with those powers in some other situation, or what the world would be like if some objects had different power. In other words, powers can support counterfactual reasoning because of their special modal status. Furthermore, as powers are dispositions, it seems intuitively perfectly fine to individuate them with reference to their manifestations and manifestation conditions.

4.5   Counterfactual dependence

On a well-known neo-Humean account, causation is reduced to counterfactual dependence.[10] The claim is that (at least) counterfactual dependency is necessary for causation. There could also be a stronger claim, that counterfactual dependency is sufficient for causation. The stronger claim that counterfactual dependence is a sufficient condition for causation is much more controversial. The claim that counterfactual dependence is sufficient for causation would be to say that, if it is true that had a not happened, b would not have happened, then a caused b. An easy counterexample to this thesis is a scenario where a and b are identical. If a had not happened, b would not have happened, since they are identical. Therefore, if counterfactual dependence is sufficient for causation, a caused b. However, this would imply that every a causes itself, for everything is identical

---

[9]Note that if you reject identity across worlds, it is vacuous that p could not have been any different from how it is, since that would mean that p is in that different way in some other possible world. But since you reject identity across worlds, p is not in any other worlds than the one it is in. Note that many of the arguments in this section would need reformulation if we were to reject trans-world identity in favor of, for example, counterpart theory. I think all the arguments would still hold.

[10](Lewis, 1973) (Lewis, 2004a)

with itself, and the identity supports the relevant counterfactual.[11]

The weaker and more plausible thesis that counterfactual dependency is necessary for causation goes as follows: if a caused b, then, had a not occurred, b would not have occurred.[12]

The idea that counterfactual dependency is necessary for causation is widely criticized with counterexamples. Take cases where, if a doesn't cause b, then c does the job. Think of a back-up generator in a hospital: if the main power supply fails ($\neg a$), it causes the backup generator to start (c), which causes the electrical equipment to function (b). In the case where the main power supply works, it should be said that a caused b, since it was the main power supply that was causally responsible for the functioning equipment. However, the conditional that $\neg a \to \neg b$ doesn't hold. For, had a not happened, b would still have happened because of the backup generator.[13]

On the powers account of causation presented in the previous chapters, it is clear that counterfactual dependence is not sufficient for causation. But is it necessary? I think that powers are counterfactual supporting, in the sense that powers may provide the truthmakers for these subjunctive conditionals. On the neo-Humean counterfactual dependence view, the counterfactual dependence is what implies the cause-statement. On my view, the causal situation involving the power makes it the case that the counterfactual holds. This is connected to the modal nature of dispositions.However, does that mean that a case of causation implies the effect would not have obtained had the causing power, not been present? I do not think so. There is no guarantee that there aren't *backup powers*, that is, powers the manifestation condition of which are such that the absence of other powers is a condition for their manifestation. Say there is a power p of x to E when M. There is a counterfactual situation where M doesn't hold, and p doesn't manifest. Is it then necessary that E does not obtain in the counterfactual situation? It would, if counterfactual dependence were a necessary condition for causation. But we may add to the scenario that x has a second power q to E when N. If the manifestation condition type N is realized by the situation where M does not hold, then y will necessitate E. The possibility of an object having two powers towards the same manifestation, but with different manifestation conditions, makes it the case that there could be backup powers.

But in this example, p and q are powers of the same object, namely x. They are also powers towards the same manifestation type, E. What makes them different powers is the difference in manifestation conditions. We could entertain the thought that p and q are really the same power, and that the manifestation conditions of this power take the form of a disjunction, such that the power p/q manifests when M or N. On this scenario, we would have restored counterfactual dependence, barring that q had a backup power too.

I see no good reason for either option: insist on unity of p and q in this case or sacrificing counterfactual dependence. But we may construct a slightly more difficult case in which sacrificing counterfactual dependence seems the more natural option.

---

[11]In order to make this argument work for the counterpart theorist, we would have to change it to reflect the impossibility of trans-world identity. Take $a\star$ to be the counterpart of a, in the closest world where a counterpart of a does not happen. The sufficiency thesis then is: if it is true that if $a\star$ does not happen, then $b\star$ will not happen, then a caused b. Now replace the identity relation in my example with the relation of being the counterpart in the closest world where the original does not happen, and you can reconstruct the argument.

[12]Note I am presuming that a and b can appear in two possible worlds. The counterpart theory alternative would be to replace a and b with adequate counterparts when evaluating the counterfactual.

[13]Again, note that if one thinks that b in the counterfactual situation cannot be identical with b in the actual situation because it has a different causal history or simply because it is in another possible world, then the example could be adapted to involve counterparts of the right sort.

If it is possible for two distinct powers of distinct object to cause the same manifestation type, then we can construct a scenario where p is a power of x to produce E when M, and q is a power of y to produce E when N, where N would obtain in some counterfactual situation where M does not. In this case, identifying p and q is impossible, since they are powers of different objects. While disjunctions in the manifestation conditions might be metaphysically allowable, powers shared between objects would not fit in the ontology presented in Chapters 2 and 3 at all. The option that p/q is a power of the object x or y is not a metaphysically possible statement. It may express an epistemological uncertainty when we simply don't know which of the two objects the power belongs to, but to think of one power actually being had by two objects is confused.

As such, we may be forced to give up counterfactual dependency even as a necessary condition of causation. However, this argument rests on the possibility for two distinct powers to produce the same event. What 'same' means in this context is difficult, since the one power would produce that manifestation in the actual world, and the other power would produce it only in the relevant counterfactual situation. So 'same' in this context should be interpreted as some appropriate relation of trans-world identity, or some counterpart relation. I do not wish to decide on this issue, but it is possible that considerations about the impossibility of trans-world identity make it so that the effect of p could never be the effect of y. If that is the case, we may have counterfactual dependency after all, since the backup power could not truly be a power towards the same manifestation. Arguments of this kind appear in non-dispositional discussions of the counterfactual dependency of causation. The idea is that the effect of the backup-generator from our earlier example, could not be qualitatively or numerically identical with the effect of the main power supply. If the effect of the backup cause is simply not the effect of the first cause, then counterfactual dependency holds.

Another way we might save counterfactual dependency is by appealing to the entire arrangement of powers present. The necessity of counterfactual dependency could be spelled out as the conditional that if the case of causation G holds, then the effect is counterfactually dependent on the cause. If we take the description G to include not only the power, but also the backup power, then the necessity of counterfactual dependency holds. However, this solution turns on what we include in the causal situation G, and 'causal situation' is not a term I have defined. I think this solution might run the risk of becoming circular: if we stipulate that we include in G precisely those situations that would verify the necessity of counterfactual, the solution becomes true vacuously. In its most extreme form, we would include all powers in the universe in G. Then the solution would be true, since if all powers in the universe had not existed, then the effect would not have been caused. But this is a rather uninteresting result. All it tells us is that effects are products of some part of our entire universe. So for this solution to have any currency, we would have to com up with a non-liberal and non-circular way to decide what is relevantly a part of the causal situation.

In conclusion, I think it is not fully decided whether causation implies counterfactual dependence of the effect on the cause, but that backup powers are not easily ruled out. Therefore it seems to me that for now I can not claim counterfactual dependency is a necessary condition for causation on my powers theory.

4.6 Transitivity

If causation is a relation, we can ask what the logical properties of the relation are. The usual conception of causation that we find in the literature is that it is an irreflexive,

anti-symmetrical, and perhaps transitive relation. I say 'perhaps' because transitivity is the most debated logical property where causation is concerned. When one accepts the idea that causation is this irreflexive anti-symmetrical and transitive relation between events, combined with a causal version of the principle of sufficient reason (denying unmoved movers), a picture emerges of the world as strings of events linked together with causal glue. I think this view is largely mistaken. Much has to be in place in order for causes to happen: objects and properties have to be arranged in just the right way for a certain event to follow. Furthermore, some discoveries of physics don't seem to jive with the causal variety of the principle of sufficient reason; some processes, such as radioactive decay, seem to happen spontaneously. Although there might be a causal story behind radioactive decay, the idea that some spatiotemporally contiguous event causes it seems to be in contradiction of the conclusions of empirical research.

4.6.1   Transitivity on the received view

On the most common view of causation, it is a relation between *events*. One event may cause another, which may then cause yet another. The question of transitivity is whether it is then always true that the first event caused the third.

> If event e_1 causes e_2 and e_2 causes e_3, then it is true that e_1 causes e_3 TRANSITIVITY

A weaker claim is also possible:

> In some but not all cases if e_1 causes e_2 and e_2 causes e_3 it is the case that e_1 causes e_3.   TRANSITIVITY$^-$

In day-to-day causal explanation we quite often speak of indirect causes as causes. If you missed the bus because you overslept, and missing the bus caused you to be late for work, then it wouldn't be unusual to say that oversleeping caused you to be late for work. It seems that our normal way of giving causal explanations at least endorses TRANSITIVITY$^-$. Of course, our normal way of giving causal explanations need not map onto the way things are, and metaphysical thesis should not be derived from how some people talk, but It will be interesting to see how much of the intuition we can accommodate on the powers view.

4.6.2   Transitivity on the powers view

For transitivity to hold, the relata must be of the same type. Since on my powers view causes are of a different type than effects, there can be no transitivity.

Alexander Bird has this idea of a network of powers: salt has the power to dissolve in water, saline water had the power to clean wounds, etc. This is not transitivity of causation, since you wouldn't say the earlier powers caused the latter. Powers produce manifestations which produce situations that include other powers, but that does not mean the effect of a power is a power too.

4.6.3   Sophisticated transitivity

Although there might not be transitivity on either view of causation, I think the event-view captures a certain intuition about chains of causal processes that the powers view does not display as clearly. If one throws a rock at a window, the window breaks and the glass falls on the floor, which makes a certain sound. On the view of causation as a transitive relation between events, the intuition that the sound produced had something

to do with throwing the rock at the window is easily captured: for throwing the rock is what *caused* the sound, through the intermediary effect of breaking the window. On the powers view this is harder to see: the rock doesn't seem to have the power to make the sound of falling glass. I want to capture the intuition that lies at the base of thinking causation might be transitive, without saying that causation is actually transitive. First of all, because this is impossible since cause and effect are of a different kind, and secondly because even if they are (as on the event-view), there seem to be problems with transitivity. The approach I take towards retaining the intuition that lies at the base of transitivity is to try and give a translation of a causal scenario of the event-view to a causal scenario on the powers view.

Take the causal scenario of the rock, the window, and the falling glass. It can be represented on the event view as such:

$$e_1 \longrightarrow e_2 \longrightarrow e_3$$

Where the arrows represent the causal relation of the event view, and e_n is an event. By transitivity we get that e_1 caused e_3.

$$e_1 \longrightarrow e_2 \longrightarrow e_3$$

Now how would this sequence of events look on the powers view? First of all, it must be said that on my view a rock does not literally have the power to break a window, since a rock is not a fundamental object and the power to break windows not a fundamental power. For the sake of argument we can either loosen the requirement of fundamentality, or just imagine that we replaced the events and objects form this example with fundamentals. In this way we can ignore the issue of fundamentality for the moment. Then, what is the cause of e_2 on the powers view. Of course it cannot be an event, since causes aren't events on my view. Rather, it must be some power that we will call p_1. And this power would manifest only if its manifestation conditions hold, so if e_2 happens, the manifestation conditions m_1 of power p_1 must hold. So we get something like this:

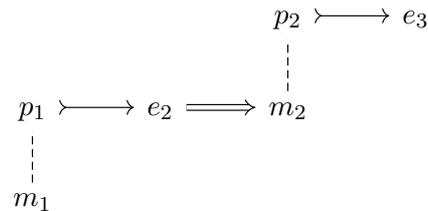$$p_1 \rightarrowtail\longrightarrow e_2$$
$$\vdots$$
$$m_1$$

Since the causal relation of the event view is something different from the production relation between powers and their manifestation, I use a different arrow between powers and their effects. The relation between the instantiated manifestation condition and the power is represented as a dotted line.

In Chapter 3, I speculated what the ontological category of manifestations might be. For now, I will be treating events as an appropriate category. As such, the breaking of the window may be seen as an effect of p_1. The natural candidate for what p_1 might be, is the power of the stone to break glass. And it seems right to say m_1 is some event or state of affairs that ensures the rock is thrown in the direction of the window with sufficient velocity.
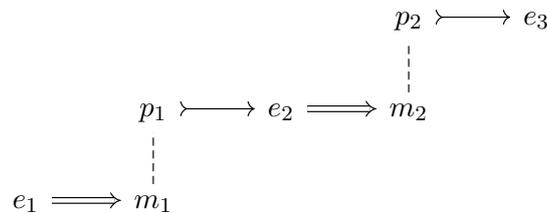
Now, what can be the cause of e_3. It must be some power p_2, given that its manifestation condition m_2 holds. So the cause of e_3 will be formally similar to the cause of e_2. The natural candidate for the power that causes e_3 is the power of the glass to make the sound of falling glass. The manifestation conditions could plausibly be

thought to be the falling of the glass on some hard surface. Let's say these manifestation conditions are satisfied by m_2.

Now what connects the breaking of the window to the sound that the glass makes? The answer is not hard: it must be that the event of window-breaking involves falling glass. The key to the connection is that the event e_2 implies the holding of the manifestation condition m_2. I use 'implies' here in a loose sense. For this relation between the event and the instantiated manifestation conditions, if will use yet another arrow, since it is not the relation of production, nor the conditional dependence the dotted line expresses.

$$
\begin{array}{ccc}
 & & p_2 \rightarrowtail\!\!\!\longrightarrow e_3 \\
 & & \vdots \\
p_1 \rightarrowtail\!\!\!\longrightarrow e_2 \Longrightarrow & m_2 \\
\vdots & & \\
m_1 & &
\end{array}
$$

However, now we see that e_1 is still not in the picture, so we cannot yet say what the relation is between e_1 and e_3. Of course, we can apply the same analysis to e_1 that we did for e_2.

$$
\begin{array}{ccc}
 & & p_2 \rightarrowtail\!\!\!\longrightarrow e_3 \\
 & & \vdots \\
p_1 \rightarrowtail\!\!\!\longrightarrow e_2 \Longrightarrow & m_2 \\
\vdots & & \\
e_1 \Longrightarrow m_1 & &
\end{array}
$$

This analysis of the event account in the vocabulary of the powers account shows that much needs to be in place for the event account to make sense. There need to be powerful objects involved in the events that are thought to stand in a causal relation. Also, there are many requirements of what could be called background conditions in the event account, that would be part of manifestation conditions on the powers account. For example, for the breaking of the window to cause the sound of falling glass, the window needs to be positioned above some hard surface. If the window were suspended above water, the falling glass wouldn't make the characteristic sound. What this shows is that, in order to say that one event is sufficient cause for another, without appealing to background conditions, these events have to be very rich entities. But maybe it can be argued that the same is true for the manifestation conditions of powers, if we are not restricted to fundamental powers.

To put it differently, in order to be allowed to draw the event-cause arrow $\rightarrow$ from e_1 to e_3, all the other arrows in the above diagram must be present. That is a strong requirement, which may be why transitivity sometimes intuitively fails on the event-account, even though transitive cases may be possible.

4.7   Irreflexivity

The question whether causation is reflexive is really the question whether anything can cause itself. On my view, like transitivity, this is prima facie impossible since causes and effects are not of the same kind. The received view is that causation is irreflexive, and I agree.

One type of happening where we might be tempted for a moment to say something is its own cause, is a spontaneous event. However, saying that it is its own cause doesn't

seem to me to have any advantages over saying it does not have a cause. Even better is to say it is caused by a power with no manifestation conditions, or manifestation conditions that obtain throughout this universe.

Claims of self-causation come up in religious contexts. It is a theological thesis of (some) monotheistic religions that God is self-caused. I think such claims are the proper domain of theology and not of the style of realist philosophy I try to follow in this thesis. However, even in this context I think the notion of being its own cause makes more sense if we take 'cause' here to refer to what Aristotle would have called a *formal cause*, not an efficient cause, which is the subject of this thesis.

4.8 Symmetry

On the received view the causal relation is anti-symmetrical. The motivation for that idea is usually that there is a connection between causation and time. If causes necessarily precede their effects, and 'preceding' is an anti-symmetrical relation, then causation must be an irreflexive relation. To put it another way, if causes precede effects, then symmetry violates the order of time.

A similar point can be made using the principle of special relativity that no thing can travel faster than the speed of light. If causation is something that is subject to this constraint, then causation cannot happen backwards in time or instantly at a distance.

An example of apparent violation of relativistic constraints on causation is the result of EPR-B experiments, so-called 'spooky action at a distance'. Quantum entangled particles can be made to exhibit certain behavior *immediately* at a distance. If this is causation, then effects can follow causes at a distance that is 'traversed' with a speed faster than light. But we must be careful with this argument. Quantum entangled state collapse might not be causation. And even if it is, that doesn't mean symmetry is possible. For going faster than light is not yet symmetry. There might be conceptual reasons other than adherence to relativity for thinking there can't be symmetry.

John Heil has written of causings as symmetrical. His example is the solution of salt in water. According to Heil, these two substances are reciprocal manifestation partners. The salt enables the water to become saline, and the water enables the salt to dissolve. "The salt and the water work in concert to yield a certain result"(Heil, 2012, p.118). Heil wants to discredit the idea that cause and effect need to have a temporal order: they may happen simultaneously. But simultaneity is not the same as the relational property of symmetry. In Heil's example, the result of the causing is the mutual manifestation of the powers of the salt and the water: he refers to this manifestation as "the salt's dissolving"(Heil, 2012, p.119). The picture that emerges is of two powerful substances that form each others manifestation conditions and together result in a manifestation.

$$
\begin{array}{c}
p_1 \\
\vdots \\
\rightarrowtail \longrightarrow m \\
\vdots \\
p_2
\end{array}
$$
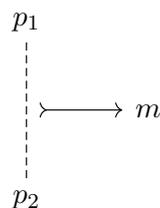
But here it seems to me it is not the causing-relation that is symmetrical. For the relation p_1 and p_2 bear to m is not symmetrical. The salt's dissolving does not relate to the powers as the powers relate to it. In other words, causation is not symmetrical.

What Heil seems to be getting at is that p_1 stands to p_2 as p_2 stands to p_1. It is the relation of *being a reciprocal manifestation partner* that is symmetrical.

Heil's other example is of two playing cards propped up against each other, so that they form a stable structure that stands upright. Again, the relation between the powers of the playing cards and the house of cards they form doesn't seem to be symmetrical. It is rather the relation that the one playing card bears to the other that is symmetrical.

Even though Heil's symmetry isn't literally symmetry of causation, I am sympathetic to the view. Causes may play out in such a way that they balance each other, or collectively produce some result they individually would not have produced. I think the playing card example can be translated to my way of talking about powers: the cards both have the power to exert a certain force (given gravity etc.). That the net sideways force is zero does not mean there is no causation.

The salt example also works on my view. The power of the water is a power to dissolve ionic substances. The solubility of the salt is the power to be dissolved in polar liquids. The reference to the manifestation partner is already contained in the power, for it is part of either the manifestation condition type of the power or the manifestation type.

4.9   Causation by absence

Can absences be causes? Some theorists think this must be possible.[14] This thesis is motivated by examples: not pulling the emergency break caused the train to hit the buffer stop, not watering the plants caused them to die, etc. I grant that these examples are acceptable causal explanations. However, explanatory practice is not always in concord with metaphysics. From the powers ontology presented in Chapters 2 and 3 it is clear that absences cannot be causes, since a power is something an object has, rather than has-not. When we do realist ontology, absences cannot be accepted as causes since they cannot be accepted as fundamental objects or properties. This is, I think, a fairly uncontroversial thesis.

There is one worry that remains: some things are absences or presences only by linguistic convention. For example, 'fasting' seems to denote an activity in which we engage, and thereby a presence rather than an absence. When we say 'I became malnourished because of the fast', at the surface we have a presence as a cause.[15] However, when we say 'I became malnourished because of a lack of nutrients', the presence is translated into an absence. The difference between the two statements does not seem to be a difference in the metaphysics of their referents, but rather a difference in the language used to express the causing.

If fasting is an absence in the linguistic guise of a presence, could there be more such examples? Could it be that something we think of as a presence is really an absence. I think this might be possible, but that it doesn't form an argument for causation by absence. In the case of fasting, we should say that the truthmaker for the causal explanation given in 'I became malnourished because of fasting' is not the presence of a fasting, nor the absence of nutrients, but rather some complex arrangement of powers in the body.

What if something we think is a fundamental and a presence, turns out to be an

---

[14]See, for example, Lewis (2004b). But Lewis agreees that the truthmakers for these causal claims can only be presences. Also see Heil (2012, p.126–130) for arguments against causation by absence.

[15]Bear in mind I use 'because' in a causal sense here. For discussion of the difference between cause and because see Lipton (2009).

absence? We can imagine (although it may be far-fetched) discovering that electrons are not really particles, but rather 'holes' in the e-field. In that case, electrons would not be presences, but absences of the e-field. Still, this gives us no reason to accept absences as causes. Our commitment to electrons as parts of our ontology goes only so far as our scientific knowledge goes. When we discover that electrons are really holes the e-field, we can simply conclude that our talk of electrons causing things was misguided, and that instead we should talk about the e-field having certain powers.

I am appealing here to empirical conclusions about absences and presences. I think it is the case that in language we treat many absences as presences, but metaphysically and physically we can distinguish the one from the other. In day-to-day causal explanation we may appeal to both, but in ontology we should stick to presences.

Conclusion

I have presented a view of causation as the bringing about of effects by powers. These powers have essences that determine in which circumstances they manifest, and what their manifestations are like. Others have given similar views; my original contribution lies in a specific combination of standpoints as well as in the way they are spelled out. I have combined the idea that causes are powers and manifestations are effects with the idea that powers are single-track dispositions essentially bound to certain types of manifestations. These ideas I have combined with the notion that manifestations are forces, and occurrent effects are only a secondary type of effect. To this potent mix of aristotelian dispositional essentialism I have added the idea that all powers are fundamental powers of fundamental objects. This coctail takes elements of the work of in particular Heil, Cartwright, and Bird, but differs on several points with each of their views.

What may be a stronger contribution to the literature on powers is so much se the position I take, but the way I have approached individuating powers, manifestations and manifestation conditions. In this thesis these categories are defined more precisely than usual in the literature. I think that the consequences of these precise specifications can be interesting even to those who hold a different variety of the powers view, or those who oppose powers altogether.

Another contribution I have tried to make in this thesis is to develop fully, by means of one example, (gravitational attraction) how a power is related to laws of nature, to its manifestation, and to its manifestation conditions. In other words, I have tried to provide a case-study of a power, thereby showing the general structure powers might have. The success of a powers-theory in dealing with this simple example provides reason to test this metaphysics against more complicated examples from physics in further research.

Perhaps the greatest advantage of any metaphysics that includes real powers is that it facilitates a connection between facts and modality. That powers can function as truthmakers for counterfactuals is a huge asset of dispositional essentialism. A full account of powers as truthmakers for natural or physical necessity is not given in this thesis, but would be an excellent subject for another project.

Through this investigation of powers I have become convinced that a metaphysics of powers is capable of providing an account of causation that is in harmony with philosophical method and intuition and with science. Furthermore, many of the downsides of the main contending theories of neo-Humean nature can be avoided with a powers metaphysics. Also it appears that a fundamental powers metaphysics, if done in the right way, can deal with most if not all of the problems associated with dispositionalism.

Several lines of argument explored in the previous chapters can be fertile grounds for further philosophy. For example, I think the idea of combination laws that tell us how forces lead to occurrent effects can be elaborated upon. If we can give a full account not only of powers and their manifestations, but also of combination laws and occurrent effects, the philosophical worth of a metaphysics of powers increases greatly. Another

area of further research could be how a metaphysics of causation that takes singular power-instances to be primary relates to universals and natural kinds. A project that I very much would like to engage in is to integrate the theory presented in this thesis with a more current account of fundamental objects from physics. I have the idea that a beautiful integration of metaphysics of causation and the Standard Model of particles could be accomplished, but that was beyond the scope of this thesis.

All in all these four chapters give what I believe to be a compelling image of how the world functions. Although there may be complications and counterarguments, I think the picture that emerged is elegant in that it takes only fundamental objects and their properties and builds out of it all movement and stability. This is how all causation is made up of powers, manifestations and their essences.

Bibliography

Anscombe, G. E. M. (1993). Causality and determination. In Tooley, M. and Sosa, E., editors, *Causation*, pages 88–104. Oxford University Press.

Armstrong, D. M. (1989). *Universals: an opiniated introduction*. Westview, Boulder.

Armstrong, D. M. (1993). The identification problem and the inference problem. *Philosophy and Phenomenological Research*, 53(2):421–422.

Armstrong, D. M. (1996). Dispositions as categorical states. In Crane, T., editor, *Dispositions: A debate*. Routledge.

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

Armstrong, D. M. (2004). *Truth and truthmakers*. Cambridge University Press, Cambridge.

Armstrong, D. M. (2010). *Sketch for a systematic metaphysics*. Clarendon Press, Oxford.

Beebee, H. (2009). Causation and observation. In Beebee, H., Hitchcock, C., and Menzies, P. C., editors, *The Oxford handbook of causation*. Oxford University Press.

Beebee, H., Hitchcock, C., and Menzies, P. C., editors (2009). *The Oxford handbook of causation*. Oxford University Press, Oxford.

Bird, A. (1998). Dispositions and antidotes. *Philosophical Quarterly*, 48(191):227–234.

Bird, A. (2007). *Nature's metaphysics: Laws and properties*. Clarendon Press, Oxford.

Black, R. (2000). Against quidditism. *Australasian Journal of Philosophy*, 78(1):87 – 104.

Bonevac, D., Dever, J., and Sosa, D. (2012). The counterexample fallacy. *Mind*, 120(480):1143–1158.

Campbell, K. (1990). *Abstract particulars*. Blackwell, Oxford.

Cartwright, N. (1983). *How the laws of physics lie*. Clarendon Press, Oxford.

Cartwright, N. and Pemberton, J. (2013). Aristotelian powers: Without them, what would modern science do? In Groff, R. and Greco, J., editors, *Powers and Capacities in Philosophy*. Routledge.

Choi, S. (2005). Do categorical ascriptions entail counterfactual conditionals? *The Philosophical Quarterly*, 55(220):495–503.

Choi, S. and Fara, M. (2014). Dispositions. http://plato.stanford.edu/archives/spr2014/entries/dispositions/.

Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, 78(2):67–90.

Collins, J., Hall, N., and Paul, L., editors (2004). *Causation and Counterfactuals*. MIT Press, Cambridge.

Crane, T., editor (1996). *Dispositions: a debate*. Routledge, London.

Craver, C. F. and Darden, L. (2013). *In search of mechanisms : discoveries across the life sciences*. University of Chicago Press, Chicago.

Creary, L. (1981). Causal explanation and the reality of natural component forces. *Pacific Philosophical Quarterly*, 62(2):148–157.

Davidson, D. (1980). *Essays on actions and events*. Clarendon Press, Oxford.

Dretske, F. I. (1977). Laws of nature. *Philosophy of Science*, 44(2):248–268.

Ducasse, C. J. (1926). On the nature and the observability of the causal relation. *The Journal of Philosophy*, 23(3):57–68.

Ellis, B. (2010). Causal powers and categorical properties. In Marmodoro, A., editor, *The Metaphysics of Powers: Their Grounding and Their Manifestations*. Routledge.

Fara, M. (2005). Dispositions and habituals. *NOÛS*, 39(1):43–82.

Fraassen, B. C. v. (1993). Armstrong, Cartwright, and Earman on laws and symmetry. *Philosophy and Phenomenological Research*, 53(2):431–444.

Groff, R. and Greco, J. (2013). *Powers and capacities in philosophy : The new Aristotelianism*. Routledge, New York.

Harré, R. and Madden, E. (1975). *Causal Powers: A Theory of Natural Necessity*. Rowman and Littlefield, Totowa.

Heil, J. (2003). *From an ontological point of view*. Clarendon Press, Oxford.

Heil, J. (2005). Dispositions. *Synthese*, 144(3):343–356.

Heil, J. (2012). *The Universe As We Find It*. Clarendon Press, Oxford.

Heil, J. (2015). Aristotelian supervenience. *Proceedings of the Aristotelian Society*, 115(1):41–56.

Hume, D. (2000/1748). *An enquiry concerning human understanding: A critical edition*. Clarendon Press, Oxford.

Johnston, M. (1992). How to speak of the colors. *Philosophical Studies*, 68(3):221–263.

Kim, J. (1993). *Supervenience and Mind*. Cambridge University Press, Cambridge.

Kistler, M. (2007). The causal efficacy of macroscopic dispositional properties. In Kistler, M. and Gnassounou, B., editors, *Dispositions and Causal Powers*. Ashgate.

Kistler, M. and Gnassounou, B. (2007). *Dispositions and Causal Powers*. Ashgate, Aldershot.

Leibniz, G. W. (1965/1714). *Monadology and other philosophical essays*. Macmillan Pub. Co., New York.

Lewis, D. (1968). Counterpart theory and quantified modal logic. *The Journal of Philosophy*, 65(5):113–126.

Lewis, D. (1973). Causation. *The Journal of Philosophy*, 70(17):556–567.

Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61(4):343–377.

Lewis, D. (1986a). *On the plurality of worlds*. Blackwell, Oxford.

Lewis, D. (1986b). *Philosophical Papers Volume II*. Oxford University Press, Oxford.

Lewis, D. (1997). Finkish dispositions. *The Philosophical Quarterly*, 47(187):143–158.

Lewis, D. (2004a). Causation as influence. In Collins, J., Hall, N., and Paul, L., editors, *Causation and Counterfactuals*. MIT Press.

Lewis, D. (2004b). Void and object. In Collins, J., Hall, N., and Paul, L. A., editors, *Causation and Counterfactuals*. MIT Press.

Lipton, P. (2009). Causation and explanation. In Beebee, H., Hitchcock, C., and Menzies, P. C., editors, *The Oxford handbook of causation*. Oxford University Press.

Manley, D. and Wasserman, R. (2007). A gradable approach to dispositions. *The Philosophical Quarterly*, 57(226):68–75.

Marmodoro, A., editor (2010). *The Metaphysics of Powers: Their Grounding and Their Manifestations*. Routledge, New York.

Martin, C. B. (1994). Dispositions and conditionals. *The Philosophical Quarterly*, 44(174):1–8.

Martin, C. B. (1997). On the need for properties: The road to Pythagoreanism and back. *Synthese*, 112(2):193–231.

Martin, C. B. and Heil, J. (1999). The ontological turn. *Midwest Studies In Philosophy*, 23(1):34–60.

McKitrick, J. (2003). The bare metaphysical possibility of bare dispositions. *Philosophy and Phenomenological Research*, 66(2):349–369.

McKitrick, J. (2013). How to activate a power. In Mumford, S. and Tugby, M., editors, *Metaphysics and Science*. Oxford University Press.

Mellor, D. H. (1974). In defense of dispositions. *Philosophical Review*, 83(2):157–181.

Michon, C. (2007). Opium's virtus dormitiva. In Bruno, G. and Max, K., editors, *Dispositions and Causal Powers*, pages 133–150. Ashgate.

Mumford, S. (1998). *Dispositions*. Oxford University Press, Oxford.

Mumford, S. and Anjum, R. L. (2011). *Getting causes from powers.* Oxford University Press, Oxford.

Mumford, S. and Tugby, M., editors (2013). *Metaphysics and Science.* Oxford University Press, Oxford.

Oddie, G. (1982). Armstrong on the eleatic principle and abstract entities. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 41(2):285–295.

Paul, L. (2009). Counterfactual theories. In Beebee, H., Hitchcock, C., and Menzies, P. C., editors, *The Oxford handbook of causation.* Oxford University Press.

Ryle, G. (1949). *The concept of mind.* Hutchinson, London.

Shoemaker, S. (1980). Causality and properties. In Van Inwagen, P. a., editor, *Time and Cause: Essays presented to Richard Taylor.* Reidel Pub. Co., Dordrecht.

Shope, R. K. (1978). The conditional fallacy in contemporary philosophy. *The Journal of Philosophy*, 75(8):397–413.

Sosa, E. and Tooley, M., editors (1993). *Causation.* Oxford University Press, Oxford.

Strawson, G. (1987). Realism and causation. *The Philosophical Quarterly*, 37(148):253–277.

Stroud, B. (1984). *The significance of philosophical scepticism.* Clarendon Press, Oxford.

Tooley, M. (1977). The nature of laws. *Canadian Journal of Philosophy*, 7(4):667–698.

Van Fraassen, B. C. (1989). *Laws and symmetry.* Oxford University Press, Oxford.

Van Inwagen, P., editor (1980). *Time and cause : essays presented to Richard Taylor.* Reidel, Dordrecht.

Yli-Vakkuri, J. (2010). Conditional and habitual analyses of disposition ascriptions. *The Philosophical Quarterly*, 60(240):624–630.