

On the dispensability of grounding:
Ground-breaking work on metaphysical explanation

JAMES NORTON

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy at the University of Sydney
April, 2017
FACULTY OF ARTS AND SOCIAL SCIENCES

ABSTRACT

Primitive, unanalysable *grounding* relations are considered by many to be indispensable constituents of the metaphysician's toolkit. Yet, as a primitive ontological posit, grounding must earn its keep by explaining features of the world not explained by other tools already at our disposal. Those who defend grounding contend that grounding is required to play two interconnected roles: accounting for widespread intuitions regarding what is ontologically prior to what, and forming the backbone of a theory of metaphysical explanation, in much the same way that causal relations have been thought to underpin theories of scientific explanation. This thesis undermines the need to posit grounding relations to perform either of these jobs. With regard to the first, it is argued that a pair of human psychological mechanisms—for which there is substantial empirical support—can provide a more theoretically virtuous explanation of why we have the intuitions that we do. With regard to the second, I begin by considering what we want from a theory of explanation, and go on to develop three attractive (yet grounding-free) theories of metaphysical explanation. I offer: i) a psychologistic theory that calls upon the aforementioned psychological mechanisms, as well as the modal relations of necessitation and supervenience, ii) a metaphysical variant of the deductive-nomological theory of scientific explanation, and iii) a metaphysical variant of the unificationist theory of scientific explanation. Furthermore, these theories draw upon mechanisms and relations (both logical and ontological) to which we are already committed. Thus, to posit grounding relations in order to explain our priority intuitions, or in order to develop a theory of metaphysical explanation, is ontologically profligate. I conclude that we should not posit relations of ground.

ACKNOWLEDGEMENTS

Firstly, I'd like to thank Kristie Miller. A cleverer or more dedicated supervisor one could not ask for. Likewise, I am grateful to Sam Baron, my associate supervisor, for piquing my interest in theories of explanation, and helping me to find the direction of the thesis. Thanks also to DBM for playing a kind of step-supervisory role (and planting in my head the seed of doubt about grounding) and John Maier for his associate supervision for the first year of my candidature.

Secondly, thanks to all those who've been involved with the weekly meetings Kristie runs. Discussions in these meetings have provided useful feedback on my research, and giving feedback on others' work has very much broadened my philosophical understanding. In particular, Michael Duncan and Andrew Latham have provided much needed critical feedback and camaraderie.

Thirdly, thanks to the Work In Progress group, organised by Dana Goswick—Ant Eagle, Al Wilson, Dana Goswick, Mike Raven—and to all the anonymous referees who have provided feedback on those thesis chapters that have been written up as journal articles. The useful comments that accompanied their rejections (and, eventually, acceptances) have vastly improved the quality of my arguments.

Fourthly, thanks to those who have attended, and provided feedback on, talks I have given on this material, and those who have given me feedback in the form of annual reviews. Likewise, I want to express pre-emptive gratitude to my examiners for taking the time to read and assess my thesis.

Finally, to my lovely wife Gina: thank you for nurturing me and keeping me sane during my PhD!

This thesis was partially completed under an Australian Postgraduate Award (APA) at the University of Sydney.

CONTENTS

Chapter 1: Introduction	1
1.1 The Project.....	1
1.2 Against Explanatorily Idle Ontology.....	3
1.3 The Master Argument.....	8
Chapter 2: The Explananda	21
2.1 Two Observations.....	21
2.1.1 Accidental and Non-accidental Modal Correlations.....	22
2.1.2 Priority Intuitions.....	25
2.2 Explaining versus Vindicating.....	30
2.3 Metaphysical Explanation.....	34
2.4 Desiderata for a Theory of Explanation.....	38
2.5 Scientific Explanation.....	42
2.5.1 Deductive-Nomological Theories.....	44
2.5.2 Causal Process Theories.....	49
2.5.3 Counterfactual Theories.....	53
2.6 The Plan Going Forward.....	54
Chapter 3: The Grounding-Based Explanation	57
3.1 Grounding.....	58
3.1.1 Relation or Sentential Operator?.....	59
3.1.2 Further Features of Grounding.....	63
3.2 The Grounding-Based Explanation.....	70
3.2.1 Explaining and Vindicating our Intuitions about Modal Correlations.....	71
3.2.2 Explaining and Vindicating the Priority Intuitions.....	72
3.2.3 Grounding-Based Theories of Metaphysical Explanation.....	73
3.3 Evaluating the Grounding-Based Explanation.....	76
3.4 Evaluating the Unfiltered Grounding-Based Theory.....	77
3.4.1 A Damning Analogy with Causal Process Theories.....	80
3.5 Evaluating the Filtered Grounding-Based Theory.....	88
3.6 Conclusion.....	95
Chapter 4: A Psychologistic Explanation of the Priority Intuitions	97
4.1 The Observations.....	98
4.2 General Schema of an Explanation.....	100
4.3 The Correlation and Causal Detection Mechanisms.....	103
4.4 Detecting Non-Diachronic Correlations.....	105
4.5 Co-opting the Causal Detection Mechanism.....	106
4.6 A Filter that Overgeneralises.....	111
4.7 Fooling the Co-opted Causal Detection Mechanism.....	114
4.8 Summing Up the Grounding-Free Explanation.....	119

Chapter 5: The Filtered Modal Relations Theory of Metaphysical Explanation	124
5.1 Grounding-based versus Grounding-free Theories	125
5.1.1 Why isn't this an Error Theory?	128
5.2 Motivating a Psychologistic Theory	130
5.3 The Devil is in the Details	136
5.4 Vagueness	140
5.5 Revisiting the Desiderata and the Theoretical Virtues	142
Chapter 6: Deductive-Nomological Theories of Metaphysical Explanation	145
6.1 A Classical DN Theory of Metaphysical Explanation	147
6.2 Metaphysical Laws as Metaphysically Necessary Generalisations	149
6.3 Material and Formal Entailment	154
6.4 Evaluating the DN Theories Against the Desiderata	156
6.4.1 Covering Cases	156
6.4.2 Irreflexivity	159
6.4.3 Asymmetry	160
6.4.4 Relevance & Understanding	166
6.4.5 Summary	170
6.5 deRossett's Theory is Not a DN Theory	171
6.6 Wilsch's DN Theory	172
6.6.1 The Constructional Conception of Metaphysical Laws	173
6.6.2 Grounding-Entailment	179
6.7 Conclusion	184
Chapter 7: Extended DN Theories	185
7.1 Strevens' Kairetic Theory	185
7.2 Kitcher's Unification Theory	187
7.2.1 Motivation and Examples in the History of Science	188
7.2.2 The Nuts and Bolts	191
7.3 Metaphysical Unification	197
7.4 Paucity and Stringency to the Rescue?	202
7.4.1 Asymmetry	202
7.4.2 Irrelevance	206
7.5 Conclusion	209
Chapter 8: Conclusion	210
BIBLIOGRAPHY	215
APPENDIX	225

Chapter 1: Introduction

1.1 The Project

There is a trend towards favouring science over philosophy in the investigation of the nature of the world. Indeed, some are convinced that there will be no work left for philosophy to do, as our increasingly sophisticated methods of empirical investigation will serve to give us all the understanding that we seek.¹ Doubtless, science is a powerful investigative tool. Empirical investigation in various sub-domains of science is clearly of great value, as is within-domain theorising within science. Yet, there are outstanding questions of how to elucidate the interconnections between these sub-domains to generate an overarching theory. We already see a move towards unification across the domains within the sciences, and indeed, when this cannot be achieved there is, unsurprisingly, often a concern that something, somewhere, has gone wrong. Witness, for example, the conversation that has resulted from the lack of fit between the macrophysical theory of general relativity, on the one hand, and the microphysical theory of quantum mechanics, on the other. One potential outstanding role for philosophy, then, is to aid the project of scientific unification by theorising about the connections between the scientific sub-domains. In particular, it is common to suppose that the entities posited by our best theories in each sub-domain can be organised into a hierarchical structure, with the most fundamental entities at the bottom (those posited by fundamental physics, say) moving up through chemistry, biology, psychology, economics and the social sciences.

Clearly enough, there are interesting relationships between microscopic and macroscopic phenomena. Physicists, chemists, biologists, psychologists, economists and social scientists of various stripes are all studying the same world. They are simply viewing it through different lenses, with different degrees of magnification. The focuses of their respective research projects (quantum particles, unemployment, chemical reactions, war, psychological states, etc.) are independently of great interest, but also of interest are questions of how these phenomena interact and perhaps explain one another.

The primary objective of this thesis is to demonstrate that the way some contemporary metaphysicians have been going about answering questions regarding the connections between the ontological posits of the various sciences is misguided. In

¹ In the first chapter of 'The Grand Design', for instance, Stephen Hawking bluntly announces that philosophy is dead.

particular, I argue that the pervasive assumption that the best way to understand the aforementioned hierarchical structure is via the characterisation and elucidation of primitive, unanalysable *grounding* relations, is ill-motivated.² Primitives like grounding are shrouded in mystery, and explicating mysteries with yet further mysteries merely serves to obfuscate, rather than elucidate, what is really going on. Getting the story about inter-level explanation (what I call *metaphysical explanation*) and the relationship between the ontological posits of the various sciences *right* is an interesting and exciting task. Yet, making progress towards this goal requires backing away from false starts, such as framing the investigation in terms of grounding relations. Admittedly, those metaphysicians interested in grounding might not see themselves as engaging in this unification project, so described. However, this is, at least incidentally, what they are doing, and consequently those of us who *are* interested in engaging with that project have a reason to evaluate the prospects of framing the investigation in terms of grounding. In any case, I hope to be forgiven for taking a primarily negative approach, here, for this is merely preparation for nailing down a better positive story. Moreover, the project is not entirely negative. As part of the argument against positing grounding relations, I offer alternative ways to do the same work, and thus point in the direction of some more profitable avenues of enquiry.

To give context to the project, I should perhaps note that this thesis forms part of a larger and more ambitious research plan to defend the view that there are no non-diachronic dependence relations³ and hence there is no ontological hierarchy. That which exists is ‘arranged’ (for want of a better word) flatly. However, the goals of the thesis are more focused and less ambitious. As stated, here I aim merely to argue against positing grounding relations.

² I will say little about grounding relations in this chapter, for an extended characterisation is the purpose of Chapter 3. However, to avoid confusion I should note here that I am taking grounding relations to be entities in the world, as supposed by, *inter alia*, Schaffer (2009), Cameron (2008), Audi (2012a), Raven (2012) and Trogdon (2013a). My criticisms do not apply to the sentential operator regimented in Fine (2012) and Correia (2005). Grounding, *qua* sentential operator, is ontologically neutral, and more closely resembles the notion of metaphysical explanation (see Chapter 3) than the notion of grounding with which I take issue.

³ In other words, there are, *in re*, no asymmetric, non-diachronic relations that run from the ‘more fundamental’ to the ‘less fundamental’ elements of our ontology. Note that this still allows for diachronic dependence relations like causation, and it allows for non-diachronic (a term I will explain in more detail in §2.1.2) non-symmetric modal relations such as necessitation and supervenience. Here, by ‘dependence relation’, then, I will intend to pick out only asymmetric relations (diachronic or non-diachronic). Of course, there are those who are happy to use the phrase ‘dependence relation’ to pick our non-symmetric instances of modal relations such as necessitation. My dispute with the latter parties is, of course, purely terminological.

The remainder of this introductory chapter is devoted to two goals. First, in §1.2, I will motivate a critical stance towards calling upon primitive entities where there is no need to do so. Generally acceptable principles of ontological parsimony tell against bloating one's ontology with new primitive entities in order to explain some phenomenon that can be explained by things one already has independent reason to believe in.⁴ Without this assumption, it is difficult for the arguments of the thesis to get off the ground (as it were). Second, in §1.3, I will briefly articulate the argumentative thread that runs through the thesis as a whole: the master argument. Then I will show how each of the remaining chapters serves to support this argument.

1.2 Against Explanatorily Idle Ontology

Imagine that you are walking through the streets of Mumbai when you come across a Sadhu serenely meditating, while levitating a foot above an ornate carpet. He is swathed in billowing orange robes, but while they do reach the ground they are hanging slackly, and are clearly not supporting his weight. *This*, you think to yourself, is a mystery. This is a phenomenon that needs to be explained. Yet, when you consider all the explanatory tools at your disposal—the entities you believe to exist, and which you have used to explain all the other phenomena to which you have been exposed—you still find yourself unable to explain this amazing spectacle.

Naturally curious, yet not wanting to disturb the serenity of the holy man, you ask a nearby shopkeeper how it is that gravity is being defied in this way. He smiles knowingly, and informs you that this Sadhu is known for his 'psychokinetic powers', whereby he uses his very thoughts to repel the ground and overcome the effects of gravity. Now, in all likelihood, an inquiring mind like yours or mine would probe further into the story that has been offered, but let us simply stipulate that you accept this explanation. There is a phenomenon that you cannot explain with your current ontological picture, and you recognise that it needs explaining, and so you admit into your ontology a new primitive entity: the psychokinetic powers of the Sadhu.

But just as you have taken on this new ontological commitment, a sudden gust of wind blasts down the street, knocking the Sadhu sideways and to the ground, and

⁴ Note that I am not making the (to my mind, implausible) claim that there is only one explanation for each phenomenon. My claim concerns the question of what *entities* we ought posit, given our explanatory goals.

revealing the small iron platform on which he was sitting. While others tend to the (unharmful, but nonetheless more grumpy than you would expect for one who appears so enlightened) Sadhu, you venture a quick glance under his carpet, where you find a powerful electromagnet, set up to repel the iron platform. As you already believe in the force of electromagnetism, you now have another available explanation for the Sadhu's levitation.

Now that you have available a better explanation for this phenomenon⁵, it would be unparsimonious to hold on to your belief in psychokinetic powers. Having electromagnetic forces in your ontology allows you to explain all sorts of phenomena, whereas psychokinetic powers were invoked purely for the sake of explaining this particular phenomenon. They do no explanatory work elsewhere. As you have many more reasons to believe in electromagnetic forces than in psychokinetic powers, you abandon your temporary commitment to psychokinetic powers and return to your relatively parsimonious ontological picture.

There are a couple of morals to be drawn from this anecdote. The primary moral is that we ought not (epistemically speaking) bloat our ontology by adding primitive entities⁶ which are not required in order to explain our experiences of the world. As the idea is sometimes phrased, we ought to be ontologically committed only to those entities that are indispensable to the best explanation of our observations. This is continuous with the Quinean (1948) project of establishing our ontological commitments by looking to the posits of our best (in Quine's case scientific) theories. If some entity is quantified over by our best theory, we should believe in it. If not, it doesn't make the ontological cut. One way to think about the Quinean project is in terms of inference to the best explanation: we ought to be ontologically committed to all and only those entities that explain (by being posited by our best theories) whatever data is in need of explanation. In the following section, we will consider how to tighten up this intuitive idea.

The secondary moral of the story of the Sadhu is that we are sometimes not well placed to see whether we have the explanatory resources required to account for some phenomenon. That is to say that there are sometimes competing theories that we are not

⁵ Better because more theoretically virtuous in the sense to be outlined in §1.3: the explanation in terms of the electromagnet has more predictive power, and is more unified with explanations of other phenomena, for example.

⁶ I am using the term 'entity' loosely here to encompass not only objects, but properties, facts, relations, etc. 'Entity', then, covers all existing things.

aware of. Even though you already believed in electromagnetic forces, and they were sufficient to explain your observations, it was not transparently the case that these forces could do this explanatory work, for the magnet was hidden. In other words, sometimes the explanatory power of our ontology is opaque to us, and this can lead us to posit new primitive entities without need. Nevertheless, once it is made clear that the entities we were already using to explain other phenomena can be used in this new case, there is a clear intuition that it would be a mistake to continue to believe in the newly invoked primitive.

Now, there is something misleading about my use of the term ‘already’, in that it is not the temporal order in which we posit various entities that gives them a more respectable ontological status, but rather, the explanatory work they do in the big picture. What is at stake is whether the primitive entity/entities can explain phenomena that could not otherwise be explained—whether they are indispensable in the sense outlined in §1.3. If electromagnetic forces can explain everything explained by psychokinetic powers, but not *vice versa*, this tells in favour of a belief in electromagnetic forces rather than a belief in psychokinetic powers.

Indeed, relative to various background ontological theories, invoking a new primitive may or may not be a rational move. Consider our pre-enlightenment ancestors noticing that their crops did not grow with consistent success. This annual variation was a phenomenon that required explaining, and one available explanation was a supernatural one: some years the Gods were happy, and the harvest was plentiful. Some years the Gods were less happy, and harvests were not so good. In the absence of a biological explanation taking into consideration hydration, soil nutrients, photosynthesis, etc., positing Gods to explain this variation is not obviously a poor rational manoeuvre. Yet, once the more predictively successful biological explanation becomes available, it is rational to scrap the supernatural explanation. Explaining seasonal variation in crop-growing success in terms of better-understood biological processes (which also serve to explain other phenomena) should lead us to abandon a belief in the relevant supernatural phenomenon (assuming there are no outstanding reasons for this belief; in this story the only reason to posit the Gods is to explain the variation).⁷

⁷ While I accept that there can be distinct explanations of the same phenomenon, when these explanations each posit different entities to do so, considerations of ontological parsimony should lead us to critically reconsider whether we want to hold on to both. Braddon-Mitchell and Jackson, for instance, remark that “Distinct explanations of the very same phenomenon typically *exclude* each other. When science came up

So, to summarise thus far, when we encounter a phenomenon that needs explaining, we should believe in the entities required for the truth of our best explanation. Only if we cannot explain the phenomenon in terms of entities we already have reason to believe in should we take seriously the option of invoking primitive entities to do so. Thus, the justifiability of invoking a primitive is relative to the ontology one currently endorses. Furthermore, it can sometimes be opaque to us that we already have the ontological resources required to explain the phenomenon in question.

Consider, for example, contemporary debunking arguments against objective moral realism.⁸ One (admittedly myopic) way to read the dialectic of these debates is as follows. We observe that we have strong and widely-shared intuitions about the truth of various ethical claims. We morally ought to keep our promises, for example, while we morally ought not torture puppies. One way to explain these converging intuitions is to posit primitive objective moral values as fundamental, mind-independent features of reality. With this ontological posit in hand, the explanation of our moral judgements is simple: we are tracking the moral values with which the world is imbued. Indeed, one might think that this posit is *required* to account for these observations.

Yet, with the principles of Darwinian evolutionary theory in hand, the sceptic about objective moral values can provide an alternative explanation. Roughly, the idea is that societies comprised of individuals who believe in binding moral obligations, which demand certain courses of action regardless of their potential inconvenience to oneself, are fitter societies. They are more cohesive, with altruism encouraged and freeloading discouraged. Not only do these societies do better than others, but those individuals within the society who most effectively internalise the demanding nature of the moral code of the group do better (on the assumption that we are good at detecting and punishing freeloaders). Those who try to get by without contributing to the group are granted lower status in the community, and are less likely to procreate. In this way, the

with the explanation of the operation of pumps in terms of air pressure this was taken to exclude the older explanation of their operation in terms of nature's abhorrence of a vacuum. Likewise, the explanation of lightning in terms of electrical discharges displaced the explanation in terms of Thor's thunderbolts, and the explanation of plant growth in terms of the chemistry of cell division displaced the explanation in terms of a vital spirit. In none of these cases did scientists say, 'How nice. We now have two explanations, where before we had only one.'" (2007:16).

⁸ Joyce (2006) endorses evolutionary debunking arguments, while FitzPatrick (2015) defends realism from debunking.

genes that lead one to believe strongly in the bindingness and objectivity of moral values are strongly selected for.⁹

Thus, says the sceptic, the observation of a widely shared commitment to objective mind-independent moral values can be explained by a theory that declines to posit anything answering to this description (Joyce, 2006). Indeed, while the availability of such an explanation may have been hitherto opaque to the realist, all we must believe in to account for these observations is the fitness-enhancing status of *believing in* objective moral values, and the evolutionary processes themselves (which we already call upon to account for other phenomena, such as complex and specialised organs, and the variety of life forms filling various ecological niches). Worse still for the moral realist, the sceptic notes that even if we assume for the sake of argument that there *are* objective moral values, the power of the evolutionary explanation of the moral judgements we in fact make brings into question whether these judgements would correspond to the moral values themselves. The idea that it might be a matter of coincidence, or luck, whether the values we espouse are the true values is a disturbing one for the objective moral realist. On the basis of these lines of thought, the sceptic concludes that to posit primitive moral values would be ontologically profligate, for such values are not part of the best explanation of the observation of our own moral attitudes. In other words, while the realist theory and the evolutionary theory can explain the same observations, parsimony tells us to prefer the evolutionary theory and thus not to posit primitive moral values.

Of course, I don't intend to weigh in, here, on the question of whether evolutionary debunking arguments against objective moral realism are successful. There are open questions regarding the veracity of the evolutionary explanation, and whether it can explain all that needs to be explained. However, I do take the following conditional claim to be true. If there is good evidence for the truth of the evolutionary story, and it can explain, *inter alia*, our moral intuitions, and our moral intuitions are the only reason we have to believe in primitive, mind-independent objective moral values (those entities not mentioned in the evolutionary story), then we ought not believe in such values.

These debates provide a useful analogue to the dialectic I present in this thesis. Instead of considering our shared moral intuitions, in the debate about metaphysical explanation it is observed that we have shared intuitions about what is ontologically prior to what (it is also observed that some modal correlations appear to be *accidental*, while

⁹ See, e.g., Street (2006).

others appear to be *non-accidental*, though a thorough discussion of this observation is beyond the scope of the thesis. See §2.1.1). Due to an apparent dearth of other resources with which to account for (and perhaps vindicate) these intuitions, many have seen fit to invoke primitive grounding relations, giving rise to a theory which undoubtedly can do this work. Yet, I argue, we have the resources to do this work without grounding, even if this is not transparently the case. Once this is made clear, the motivation for positing primitive grounding relations is entirely undercut, not unlike the way in which the motivation for positing objective moral values is undercut by the evolutionary explanation, and the motivation for positing psychokinetic powers is undercut when we see that the Sadhu is sitting on an electromagnet. This thesis, then, is devoted to showing how, unbeknownst to us, we can account for the phenomena which motivate positing grounding relations, using other resources to which we are already committed.

1.3 The Master Argument

The overarching argument of the thesis is a simple one. I argue that:

1. One ought (epistemically) to be ontologically committed to all and only those entities that are indispensable to the best explanation of our observations.
2. Grounding relations are not indispensable to the best explanation of our observations.
3. Therefore, we should not be ontologically committed to grounding relations.

I hope that the stories told in §1.2 have served to motivate the intuitive plausibility of (1), which is also known as the explanatory criterion for ontological commitment. I will refrain from mounting a lengthy defence of this premise, for I simply take it as a plausible starting point for this investigation. While I am not alone in finding (1) compelling,¹⁰ I recognise that it is not universally accepted, and that rejecting this assumption will undermine my argument against grounding.

However, I am taking the relevant notion of *observation*, here, to be fairly broad, such that accounting for our observations will include explaining our linguistic behaviours, our intuitions, philosophical judgements about cases, and so forth. Thus

¹⁰ See Harman (1977:6); Sayre-McCord (1988:441); Colyvan (2000, 2001).

understood, the principle is not ‘anti-metaphysical’, in such a way as to stack the deck against those who posit grounding. Prominently, Lewis (1986a) uses something like (1) to argue for concrete modal realism, that most inflationary of ontological views. Likewise, Colyvan (2001) uses (1) to argue for the existence of mathematical entities, another highly controversial ontological posit. Thus, the principle, as stated, is not unfriendly to those who would posit contentious entities, and although *some* defenders of grounding might take themselves to have independent reason to reject (1), those will not be reasons common to all advocates of grounding.

Nevertheless, there are two qualms one might have with (1). On the one hand, some metaphysicians will think that there are reasons—entirely unrelated to the explanation of our observations—to posit entities. Thus, they will reject the ‘only’ in premise (1). On the other hand, scientific anti-realists might deny that we should posit entities simply because they play an explanatory role in our best theory. Thus, they will reject the ‘all’ in premise (1). In the latter case, my arguments will still go through. Indeed, if scientific anti-realists are right, not only should we not posit grounding, we should be cautious to posit even those entities which do indispensable explanatory work. In the former case, my argument will not go through, for there may be a reason to posit grounding that is not undermined by the fact that grounding is dispensable to the best explanation of our observations. I have little to say in response to this, aside from a request for a story about the *other* reasons one might have for positing some entities.

Ultimately, I think that it is a principle very much like (1) that gets arguments *in favour* of grounding up and running in the first place. That is, the main reason to believe in grounding, as I see it, is that it is putatively indispensable to the explanation of certain observations.¹¹ Thus, someone who rejects (1) yet believes in grounding needs to provide an alternative argument in favour of grounding. Whether that can be done is not something I will consider here.

Before we delve into the competing explanations, I will briefly spell out what it is we are trying to explain. An in-depth exploration of these observations is the focus of Chapter 2. The twin observations that the theories discussed here are trying to account for are as follows. First, we observe widely convergent intuitions about which modal

¹¹ Kovacs (forthcoming), for example, claims that “The central motivation for grounding is that it is indispensable to a certain kind of explanation”. In Chapter 3, I note how some other putative motivations to posit grounding are only genuine motivations if grounding is required to do this explanatory work.

correlations are accidental and which are non-accidental. Second, we observe similarly convergent intuitions about what is ontologically prior to what. It is the second observation that will be the focus of the thesis. Before we go on, let's also stipulate that there is a general background theory—henceforth I will call this the *Background Theory*—to which these theories can be appended in order to explain all of our observations. The Background Theory is an overarching theory built from all the other sub-theories that explain our observations over a range of domains, and their associated ontological posits. I assume that the Background Theory can explain all of our observations, with the exception of those that are our focus here. This is, I hope, a harmless simplification.

The first theory we will consider is the Background Theory with the additional posit of grounding relations that obtain between some specified entities. Let's call this the *Basic Grounding Theory*. I want to call into question whether the Basic Grounding Theory can explain the relevant observations at all. Admittedly, given the Basic Grounding Theory, it is easy to see why *there are* non-accidental modal correlations (they're the correlations where there is a grounding relation between the correlated entities), and why some entities *are* ontologically prior to others (because said entities *ground* those they are prior to). However, the relevant observations are not that there *are* non-accidental correlations and that some entities *are* prior to others. It's not clear how we could, directly, observe such phenomena. The observations, instead, concern our collective *intuitions* about these phenomena.

For, our intuitions are all that we can observe in this domain, and such intuitions are defeasible; we could be mistaken in our intuitions. Moreover, it is far from obvious that positing grounding relations can account for the observed intuitions. While grounding relations are supposed to be part of the structure of the world, without some theory about how we could come to know what grounds what—some epistemology of grounding—the Basic Grounding Theory seems to make no predictions about what intuitions we will have. Thus, such a theory looks to be *empirically equivalent* to the Background Theory that does not posit grounding relations. Here, I am supposing that two theories, T and T*, are empirically equivalent iff, necessarily, any empirically testable prediction of one, is a prediction of the other. That is to say, there is no possible observational prediction that is a prediction of one theory but not the other.

As the Basic Grounding Theory does nothing to connect up the existence of grounding relations with our observed intuitions, it's hard to see how such a theory could

predict the set of intuitions that we observe (i.e. our priority intuitions, as evidenced by, *inter alia*, our linguistic behaviour). Hence it is unclear how positing grounding could contribute to an explanation of these intuitions, let alone be indispensable to this explanation. The natural way to parlay the Basic Grounding Theory into something that can do some explanatory work is to insist that we are somehow able to directly track grounding relations. I find the mystery of this unappealing, but I suspect that some defenders of grounding hold this view.

The second theory (and my preferred theory) contends that the ontology of the Background Theory we ‘already’ use to account for our other observations can be shown to also account for the intuitions that are our focus here. I call this the *Grounding-free Theory*. The Grounding-free Theory has the same ontology as the Background Theory, differing only in that it has a story to tell about how this ontology can account for our intuitions about modal correlations and our priority intuitions. The challenge of accounting for and vindicating our intuitions about modal correlations is a project I tackle elsewhere.¹² I won’t say any more about these intuitions, here, beyond recognising that these intuitions constitute further data that must be explained by the Grounding-free Theory. The conclusions of this thesis, then, can be seen as provisional, for they rely on the plausibility of this further research project. For now, there is more than enough work to be done accounting for the explanandum that is our primary focus: the priority intuitions.

I contend that the Grounding-free Theory can explain our priority intuitions via the appeal to evolved psychological mechanisms, and can vindicate them using apparatus provided by theories of scientific explanation (see Chapters 4, 5, 6 and 7). This explanation is empirically testable, and, indeed, there is strong empirical evidence for the mechanisms to which it appeals. They are the very same mechanisms we use to track causal dependencies. Furthermore, the Grounding-free Theory makes predictions about our priority intuitions that align with the intuitions we actually observe. For example, the theory predicts the observation that there are, in some cases, diverging priority intuitions. Even if one thought that Basic Grounding Theory could somehow *directly* account for our intuitions, that theory cannot account for this kind of divergence (if we were all directly tracking grounding relations, why would we disagree?). Thus, clearly enough, the Grounding-free Theory provides an empirically superior explanation of our priority

¹² In ‘How to wear the crazy trousers: defending flatland’ (Duncan, Miller and Norton, ms).

intuitions than the Basic Grounding Theory (if only by deigning to provide such an explanation).

However, I assume that not all contemporary defenders of grounding have been thinking that Basic Grounding Theory can explain the priority intuitions. Basic Grounding Theory, then, is something of a caricature. Instead of supposing that we somehow directly apprehend what grounds what, defenders of grounding can appeal to the psychological mechanisms spelled out by the Grounding-Free Theorist. Call this view *Sophisticated Grounding Theory*. The Background Theory will include, amongst its posits, the mechanisms appealed to by the Grounding-free Theory, as these are required for the best explanation of our judgements of causal dependence. Thus, the Sophisticated Grounding Theorist can call upon these mechanisms as a way of spelling out the requisite connection between grounding and our intuitions. In other words, the Sophisticated Grounding Theorist can piggyback upon the efforts of the Grounding-Free Theorist in order explain why we have the intuitions that we do. She will thereby be able to predict our intuitions (even the divergent ones) just as well as the Grounding-free Theorist. She will simply think that where intuitions diverge, at most one party is correct.

Of course, the Basic Grounding Theory is also committed to the existence of these mechanisms, as they do indispensable explanatory work elsewhere. Thus, the ontologies of the Sophisticated Grounding Theory and the Basic Grounding Theory are the same. Namely, the ontology of each of these views is that of the Background Theory (and the Grounding-free Theory), plus grounding relations. The difference between the views concerns what is alleged to explain our observations. For the Basic Grounding Theorist, the grounding relations directly explain our observations, whereas for the Sophisticated Grounding Theorist, the proximal explanation for these observations is the existence of these mechanisms. As such, the Sophisticated Grounding Theorist agrees with the Grounding-free Theorist about the proximal explanation for our observations. In contrast to the Grounding-free Theorist, however, she will contend that these mechanisms evolved, in part, in order to track grounding relations. Thus she will disagree with the Grounding-free Theorist not only about ontology (since she thinks there exist grounding relations, and the Grounding-free Theorist does not) but also about the distal explanation for our having the intuitions we do (since she thinks that we have those judgements because we have mechanisms that evolved to track grounding relations).

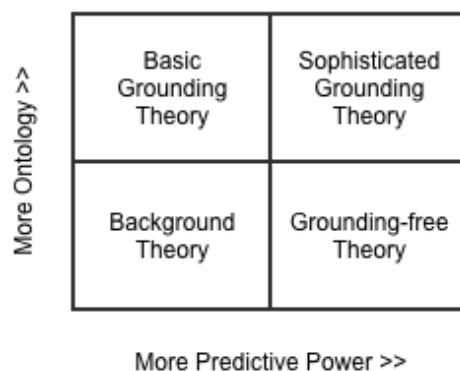


Figure 1.

In sum, there are three theories on offer: the Basic Grounding Theory, the Grounding-free Theory and the Sophisticated Grounding Theory (See Figure 1). The former is empirically inferior to the latter pair: it provides an inferior explanation of our observations. However, the latter two theories make the same predictions vis-à-vis our priority intuitions. Moreover, these theories make the same predictions in other domains, for they are based on the same Background Theory. However, being empirically equivalent does not equate to providing an equally good explanation of our observations. We could, for instance, add a causally inefficacious spaghetti monster to either of these theories, and while they would remain empirically equivalent, the spaghetti-monster-theory would clearly provide a worse explanation. That's because the spaghetti monster would be *dispensable* to the explanation.

Now we need to clarify the relevant notion of indispensability. This is of particular importance because there is more than one way to think about what it takes for an entity to be indispensable to the best explanation of our observations. Quine (1948), for example, supposed that an entity is indispensable to a theory just in case, when that theory is regimented into first-order logic, the entity in question is quantified over.¹³ However, in order to sidestep concerns about the various distinct ways a theory can be thus regimented, it will be more appropriate for our current purposes to proceed via the revised account of indispensability developed by Colyvan (1999) in the context of debates regarding the (in)dispensability of mathematical objects. What I will call 'Colyvan's Principle' suggests (roughly) that an entity, *E*, is dispensable to a theory, *T*, if

¹³ "A theory is committed to those and only those entities to which the bound variables of the theory must be capable of referring in order that the affirmations made in the theory be true." (Quine 1948: 33)

there is available a distinct theory, T^* , that is empirically equivalent to T yet makes no mention of E , and T^* exemplifies theoretical virtues to a greater extent than T .¹⁴

The purpose of Colyvan's Principle is to act as a tie-breaker between empirically equivalent theories such as the Sophisticated Grounding Theory and the Grounding-free Theory. Moreover, I take it that by using Colyvan's formulation of indispensability I am being charitable to my opponent. For, just as Colyvan's Principle is not unduly hostile to his own project of arguing for the indispensability of—and thereby the existence of—mathematical objects, it is also not unduly hostile to those who defend the indispensability of grounding relations. The principle is a nice way of formulating the intuitive thought at play when we reject the Sadhu's primitive 'psychokinetic powers' because of the availability of a more virtuous explanation of his levitation. Indeed, I contend that the popularity of belief in grounding is partially explicable by an apparent lack of an alternative explanation of our observations. Metaphysicians have posited the analogue of psychokinetic powers because this was, apparently, the only explanation on offer.¹⁵ What I hope to offer is the analogue of an explanation in terms of electromagnetism.

As these theories are empirically equivalent, if I can show that the Grounding-free theory is more virtuous than the Sophisticated Grounding Theory, I will thereby have shown the dispensability of grounding relations. On the other hand, if the Sophisticated Grounding Theory really is the more virtuous theory, this will justify positing grounding relations. While this comparison will be carried out in more detail in the following chapters, I am now in a position to briefly characterise the sense in which my Grounding-free Theory is more theoretically virtuous.

First, the Grounding-free Theory has the edge over its rivals in terms of ontological parsimony (henceforth I shall refer to this virtue simply as *parsimony*). This virtue reflects the value of a theory which declines to posit additional entities without need, and is therefore a cornerstone of my comparison between the Sophisticated Grounding Theory and the Grounding-free Theory. It is parsimony that tells us that the spaghetti-monster-theory is a bad one. Simply by positing the primitive ontology of

¹⁴ In what follows, I will often talk about whether grounding is *needed* to explain some observations. By this, I intend that grounding is needed, just in case any other explanation for the same observations is worse than that given by positing grounding. So, grounding is needed just in case it is indispensable to explaining the relevant observation.

¹⁵ See Baron (2015, 2016) for discussion of why we, as a philosophical community, should set a high bar for the justification of the invoking of a primitive.

grounding relations, the Sophisticated Grounding Theory runs afoul of parsimony, such that, *ceteris paribus*, it is a less virtuous theory.¹⁶ Thus parsimony ultimately tells against positing grounding.

Second, the Grounding-free Theory is more ideologically simple, in that it has fewer basic principles. I shall refer to this virtue as *elegance*. The Grounding-free Theory only uses ideology to which the Sophisticated Grounding Theory is already committed. However, the Sophisticated Grounding Theory appeals to ideology to which the Grounding-free Theory is not committed. That's because grounding is governed by a collection of principles (axioms of ground, if you will). For example, many think that grounding must be transitive, asymmetric, irreflexive, and well-founded (see Chapter 3 for a discussion of these features). As such, appending grounding relations to one's theory is not only an ontological burden that decreases parsimony, but also an ideological burden that decreases elegance.

Lastly, the Grounding-free Theory and the Sophisticated Grounding Theory both exemplify the virtue of unification, for the explanations they offer are continuous with explanations of other phenomena (assuming that the Sophisticated Grounding Theorist helps herself to a psychological explanation akin to the one I present in Chapter 4). In particular, they appeal to evolutionary processes (which explain a diverse range of observations) and to psychological mechanisms that already form the basis of some theories of scientific explanation. As the Sophisticated Grounding Theory uses the same psychological mechanisms as the Grounding-free Theory to explain our priority intuitions, these theories achieve unification to the same degree.

Moving on, I take it to be a desirable feature of an explanation that *vindicates* our priority intuitions (in the sense outlined in §2.2) that it also affords us a theory of *metaphysical explanation*. It might be thought that the only way to get such a theory up and running is to appeal to grounding relations, to play a role analogous to the role that causal relations play in a causal process theory of scientific explanation (see §2.5). This is not so. I develop (in Chapters 5, 6 and 7) three grounding-free theories of metaphysical explanation, to which the Grounding-free Theorist might appeal in order to achieve this

¹⁶ The Basic Grounding Theory would, of course, be arguably more parsimonious than the Grounding-free Theory were it not to be committed to the existence of the various mechanisms posited by the Grounding-free Theory. But, as noted earlier, such a theory is clearly a non-starter: even the Basic Grounding Theory must be committed to all the entities required by the best explanation of observations across all domains.

kind of vindication. One's prior views on explanation will dictate which is the most attractive of these theories (see §2.4).

Moreover, I will argue in Chapter 3 that grounding-based theories¹⁷ of metaphysical explanation are deeply flawed. For, these theories appeal to the obtaining of grounding relations to vindicate the priority intuitions, while also appealing to the priority intuitions to establish what grounds what. There is viciousness to this circularity, and I will argue that we should be sceptical that there is anything in the world that does *such* a good job of explaining that kind of intuition. Thus, to frame things in a Bayesian way, we should have very low prior credence that grounding relations (thus described) exist. So this apparent merit of the grounding-based theories of explanation should ultimately undermine our confidence that such theories are true, rather than bolstering said confidence.

Clearly enough, within the framework I have developed here, most of the heavy lifting is done via the comparison of the virtues of empirically equivalent theories. The goal then, in what follows, is to make a case that the Grounding-free Theory makes more accurate predictions than the Basic Grounding Theory and is more theoretically virtuous than the Sophisticated Grounding Theory. Let's return to the master argument. (3) follows from (1) and (2), and therefore the bulk of the thesis constitutes an extended defence of premise (2), which I shall henceforth refer to as EDG (the Explanatory Dispensability of Grounding).

However, the master argument assumes that grounding is *primitive*, and matters are less straightforward if one thinks that, if there is a relation of ground, it reduces to something else. For then one might be unmoved by my defence of EDG. One might concede that grounding relations are dispensable to the explanation of the observations outlined in the Chapter 2, yet insist that the reductive base of grounding is indispensable to the explanation of some other observations. For instance, perhaps positing impossible worlds is indispensable to explaining how it is that our mental states have certain representational content. In that case, if grounding can be reduced to the truth of certain counterpossibles, (à la Alastair Wilson, forthcoming) then one might insist that EDG is false. Of course, the reductionist about grounding would need to show that the reductive base of grounding is indeed indispensable to explaining some of our observations and

¹⁷ 'Grounding-based theories' is a term I use to capture both theories that posit grounding: namely the Sophisticated and Basic Grounding Theories.

that might prove difficult. So the master argument *might* go through even if one is a reductionist about grounding. But once again, that it not something I will defend here.

So, the aim is to show that there is a theory that is, empirically speaking, either better than or just as good as a theory that posits grounding. Moreover, if the theories are empirically equivalent, the Grounding-free Theory is more theoretically virtuous (EDG is true). Without further ado, here is an outline of the way in which the following chapters will defend this claim.

Chapter 2 will outline the twin observations that these competing theories are trying to explain—the explananda. I will briefly articulate one explanandum: the observation that we intuit that there are two distinct categories of modal correlations—the accidental and non-accidental. Explaining and vindicating these intuitions is a job I attempt elsewhere. The focus of the thesis is the other explanandum: that we have widely shared intuitions of ontological priority. We intuit, for example, that parts are prior to their wholes, mental states depend on physical states and an urelement is prior to its singleton set. I show how these intuitions are in need of explanation and, perhaps, vindication, and that this cannot be achieved merely by appeal to the traditional modal relations of necessitation and supervenience. Furthermore, I note that many will want this vindication to be accompanied by a theory of metaphysical explanation, which can systematically account for the truth or falsity of claims of the form ‘x *because* y’. I will say a little about what we might want from a theory of explanation (the *desiderata* for such a theory), and the resources that can be borrowed from theories of scientific explanation.

Chapter 3 will provide a thorough characterisation of a relation I believe not to exist: grounding. As a primitive, grounding is typically introduced by way of example, but in this case, the examples of priority intuitions introduced in Chapter 2 can serve to play this role. I will note the (few) relatively uncontroversial logical features of the relation. Unfortunately, its primitiveness also leaves the characterisation of grounding rather unconstrained, and thus most of the features considered are controversial. I will emphasise how disputes about these features tend to be framed in terms of explanatory intuitions. I go on to outline how the Basic Grounding Theory might explain and vindicate the observations noted in Chapter 2. I criticise the view for being less predictively powerful than the Grounding-free Theory. The Sophisticated Grounding Theory, on the other hand, is empirically on a par with the Grounding-free Theory: it makes the same predictions as the Grounding-free Theory. Nonetheless, I argue that we

should be suspicious of a view that uses our anthropocentric explanatory intuitions as evidence of what the mind-independent world is like. This suspicion is inherited by the supposedly ‘unfiltered’ grounding-based theory of metaphysical explanation. On the other hand, the filtered grounding-based theory (according to which, our priority intuitions fail to be perfectly aligned with the grounding structure of the world), suffers from an unfavourable comparison with the view I develop in Chapter 5, and creates an epistemological mystery regarding how we might find out about those grounding relations that obtain where we lack priority intuitions. Ultimately, by positing grounding, the grounding-based theories are less theoretically virtuous than the Grounding-free Theory. However, in order to properly make this comparison, I need to explicate my preferred theory. Thus, the scene will be set for my competing explanation of the priority intuitions.

Chapter 4 offers my grounding-free explanation of our observations. It will provide an explanation of why we have the priority intuitions that we do, framed in terms of the functioning of certain psychological mechanisms for which there is strong empirical evidence. To do this, I delve into the psychological literature regarding how humans seek out diachronic correlations and filter these correlations in order to determine what diachronically depends on what. In other words, I give an account of the mechanisms which even defenders of grounding should admit we use to discover causal dependencies. I argue that these very same mechanisms have been co-opted to filter non-diachronic correlations in search of non-symmetries. While often these mechanisms are effective (as when they allow us to discover a non-symmetric instance of a modal relation like supervenience or necessitation¹⁸), they also have a tendency to overgeneralise, and flag the presence of a non-symmetric relation in cases where the modal relations obtain symmetrically. Once we are aware of our own capacity to overgeneralise in this way, priority intuitions which one might have explained in terms of grounding relations can be accounted for, in a more parsimonious and elegant way, in terms of evolved human psychological mechanisms. Thus this chapter is directly a defence of EDG.

¹⁸ Following Russell (1903), symmetric relations are ones in which, for any x, y , if $x R y$, then $y R x$. Asymmetric relations are ones in which, for any x, y , if $x R y$, then it is not the case that $y R x$. And non-symmetric relations are ones in which, for some x, y , if $x R y$, then it is not the case that $y R x$. In what follows, I will call an instance of any non-symmetric relation in which $x R y$, and $y R x$, a symmetric instance of that relation, or, alternatively, I will say that the relation obtains symmetrically, and I will call an instance of any non-symmetric relation in which $x R y$ and it is not the case that $y R x$, a non-symmetric instance of the relation, or, alternatively, I will say that the relation obtains non-symmetrically.

With this explanation in hand of why we have the priority intuitions we do, the remaining chapters constitute various ways to extend the Grounding-free Theory so as to vindicate the priority intuitions. These chapters support EDG by showing that there is no need to posit grounding in order to achieve this vindication. **Chapter 5**, for example, draws upon the psychological mechanisms identified in Chapter 4 to develop a psychologicistic account of metaphysical explanation: the filtered modal relations theory. This view will be attractive to those who like to think about explanation as a primarily epistemic notion (as opposed to thinking of explanation in the world). This is because it is a direct consequence of the view that from different psychological perspectives, different metaphysical explanations are true. Several psychologicistic accounts are offered, but I contend that the most attractive version of this view is (what I call) Dispositional Community Relativism. Roughly, the view is that the metaphysical explanation ‘x *because* y’ is true (relative to a centred world) iff i) the community in which the individual at the centre of the world is embedded is disposed to have the mental state according to which x *because* y and ii) x necessitates y or y supervenes on x. This theory constitutes a more parsimonious and elegant way to vindicate the truth and falsity of putative metaphysical explanations than does positing grounding relations, as we already believe in these psychological states and modal relations.

Chapter 6 explores the prospects of developing a metaphysical variant of Hempel and Oppenheim’s (1948) deductive-nomological (DN) account of scientific explanation. I show how we can develop a powerful theory of metaphysical explanation simply by positing metaphysical laws and using the apparatus of classical logic. The power of this theory depends on precisely how we characterise the metaphysical laws. In other words, if we use a very lightweight conception whereby metaphysical laws are merely descriptions of exceptionless generalisations about modal space, the theory struggles with the problems of symmetry and irrelevance that plagued traditional DN theories. Thus, we can either accept the truth of some counterintuitive explanations (and thereby vindicate fewer of our intuitions), do some heavier lifting with a more robust notion of a metaphysical law (though I am not sanguine about the prospects of doing so non-anthropocentrically), or once again draw upon some psychologicistic machinery to get a less objective theory with a more palatable semantics.

Chapter 7 considers the metaphysical analogues of some of the proposed adaptations of the DN theory. After concluding that Strevens’ (2008) Kairetic theory is

of little use in this context, I go on to show that Kitcher's (1981, 1989) unification theory can go a step further than the plain DN theory, in that our theory of metaphysical explanation need not appeal to a robust conception of metaphysical laws or import psychologistic machinery. However, to do so we must help ourselves to the assumption that a certain class of truths are not in need of metaphysical explanation. Nonetheless, for those seeking a non-anthropocentric vindication of the priority intuitions, the unification view is the most attractive of the options I consider here. Following Kitcher, we can develop an objective theory of metaphysical explanation using only logical machinery to which we are already committed.

Finally, **Chapter 8** will conclude the thesis by showing how earlier chapters have argued for the truth of EDG, and thus for the soundness of the master argument. I contend that the evolutionary explanation provided in Chapter 4 entirely undermines the supposed indispensability of grounding to account for our intuitions of ontological priority. Moreover, even those who seek to vindicate said intuitions can do so via a theory of metaphysical explanation that makes no appeal to grounding, regardless of the features one demands from a theory of explanation.

Chapter 2: The Explananda

In Chapter 1, I explicated a certain view on what it takes for some entity to be indispensable to our best explanation of our observations. Before we go on to articulate and appraise the relevant competing explanations, however, we must clearly identify the observations we are trying to explain: the explananda. Achieving this clarity is the role of this chapter. This will set the scene for the remainder of the thesis.

The plan for the chapter is as follows. §2.1 will identify the explananda: a class of intuitions about which modal correlations are *accidental* and which are *non-accidental*, and a class of non-diachronic *priority intuitions* (intuitions about what is ontologically prior to what). §2.2 articulates the distinction between *explaining* and *vindicating* these intuitions. §2.3 characterises the notion of *metaphysical explanation*, for some readers will desire that the priority intuitions be vindicated in such a way that we also end up with a useful theory of metaphysical explanation. To give us the tools to compare such theories, §2.4 identifies a class of *desiderata* that we might want from a theory of explanation. This lays the groundwork for a comparison of the competing accounts in the following chapters. §2.5 will provide a brief overview of the theories of scientific explanation put forward during the last century. As well as providing a very helpful leg-up when it comes to developing our own theories of metaphysical explanation, this will help illuminate a tension between the desiderata. Finally, §2.6 gestures at how my argument will proceed in the following chapters.

2.1 Two Observations

The twin observations that require explaining—the explananda—are as follows. On the one hand, there exists an array of widely shared, converging intuitions that modal correlations come in two distinct categories: some are accidental, while others are non-accidental. On the other hand, there exists an array of equally widely shared, equally converging intuitions regarding what is ontologically prior to what. Defenders of grounding take these intuitions to be evidence for the existence of grounding relations, as we shall see in Chapter 3. In Chapter 4 I will offer a better explanation. For now, the plan is characterise these observations (this data) in a way that is neutral between the competing explanations.

To preface my characterisation of the twin observations, I should emphasise that I am not mounting any kind of *defence* of the claim that we share these intuitions—either those regarding modal correlations or those regarding priority. Rather, I am taking it as raw data that these intuitions are shared amongst contemporary metaphysicians. Furthermore, these intuitions are thought to be representative of the intuitions of the broader philosophical community, and indeed of non-philosophers insofar as they are familiar with the relevant concepts. All of this I am charitably taking at face value, and thus I grant to the defenders of grounding that there are intuitions here that are in need of explanation. If it turns out that these intuitions are less widespread or convergent than I am granting here, then there is nothing to be explained (better: there is something entirely different to be explained). However, this discovery would do the work of undercutting the motivation for positing grounding all by itself, rendering my arguments superfluous.

2.1.1 Accidental and Non-accidental Modal Correlations

The first explanandum is the observation that we have powerful intuitions according to which we want to distinguish between *accidental modal correlation* and, *non-accidental modal correlation*. Here, I am taking modal correlations to simply be regularities across modal space. So, for example, the fact that every car-owner—in every possible world—is a vehicle-owner, reflects a modal correlation between car-owners and vehicle-owners. Thus, the notion of a modal correlation is tightly intertwined with what I call the *traditional modal relations*: necessitation and supervenience.¹⁹ For, I am conceiving of these relations as ontologically lightweight, in the sense that they merely describe modal correlations. Thus, to say that A necessitates B is merely to say that there is a particular kind of modal correlation between A and B: every world where A exists is a world where B exists. Likewise, to say that A-properties supervene on B-properties is merely to say that there can be no change in the A-properties without a change in the B-properties. Understood this way, there is no difference between talk of traditional modal relations and talk of modal correlations.

Here is a reminder of the terminology introduced in Chapter 1. Symmetric relations are ones in which, for any x, y , if $x R y$, then $y R x$. Asymmetric relations are

¹⁹ For ease of readability, I will sometimes omit the ‘traditional’.

ones in which, for any x, y , if $x R y$, then it is not the case that $y R x$. And non-symmetric relations are ones in which, for some x, y , if $x R y$, then it is not the case that $y R x$. I will call an instance of any non-symmetric relation in which $x R y$, and $y R x$, a symmetric instance of that relation, or, alternatively, I will say that the relation obtains symmetrically, and I will call an instance of any non-symmetric relation in which $x R y$ and it is not the case that $y R x$, a non-symmetric instance of the relation, or, alternatively, I will say that the relation obtains non-symmetrically.

With this terminology in hand, we can note that the traditional modal relations are non-symmetric: they have both symmetric and non-symmetric instances. So, for example, that Alex is a car-owner necessitates that Alex is a vehicle-owner, yet that Alex is a vehicle-owner fails to necessitate that Alex is a car-owner, for there are worlds where Alex owns a motorcycle, but no cars. This is a non-symmetric instance of necessitation. In contrast, (assuming that numbers exist necessarily), that 2 exists necessitates that 3 exists, and that 3 exists necessitates that 2 exists. This is, therefore, a symmetrical instance of necessitation. Moreover, the traditional modal relations are reflexive. For any A , A necessitates A and A -properties supervene on A -properties. It is easy to see why: trivially, any world where A exists is a world where A exists, and there can be no change in A -properties without a change in A -properties.²⁰

The explanandum regarding accidental and non-accidental correlations, then, concerns our intuitive categorisation of the modal relations into those that are accidental and those that are not. For example, we notice that the existence of the number 2 symmetrically necessitates the existence of the singleton set $\{2\}$: wherever we find one, we find the other. We now notice that the existence of $\{2\}$ symmetrically necessitates the existence of the universal *blueness* (assuming that universals exist necessarily). Indeed, we notice that for any two necessary existents, these things are modally correlated. Yet we are inclined to think that in the former case there is a special connection between the existence of 2, and the existence of $\{2\}$. This is no accidental modal correlation. By contrast, we are inclined to say that the same is not true of the modal correlation between 2 and the universal *blueness*. This we intuit to be a mere accident, a correlation

²⁰ To pre-emptively head off any confusion, according to my terminology here modal relations are simply those that describe patterns of modal correlation. Thus grounding is not a modal relation, for it comes apart from modal correlation in various ways (see Chapter 3). Given that the obtaining of grounding relations is typically thought to entail certain patterns of modal correlation, some might object to my labelling of grounding as a non-modal relation. However, nothing of substance hangs on this distinction, and I am in need of a term to distinguish grounding from (what on this view we might think of as the *other*) modal relations.

that is due to the necessary status of these entities, rather than any interesting connection between them.

Likewise, that tigers exist necessitates both that animals exist and that *blueness* exists. Yet it strikes us as no accident that the existence of tigers necessitates the existence of animals, and entirely accidental that the existence of tigers necessitates the existence of *blueness*. Thus, there is something to be explained, here. All else equal, it is desirable to have some account of *why* we intuit some modal correlations to be accidental connections while others are non-accidental.

Some might be inclined to say that what is in need of explanation is not a modal correlation at all. What is in need of explanation is a *non-modal* connection between 2 and {2}—namely that the latter owes its existence to the former—and what is *not* in need of explanation is why both 2 and {2} exist of necessity. I agree that it is not the latter which is in need of explanation. However, neither do I think the relevant explanandum concerns why or how {2} owes its existence to 2. Rather, I take the explanandum to be our intuition that *the fact that 2 is modally correlated with {2} is no accident*: there is some important connection between the two, in virtue of which that correlation holds. I call this non-accidental modal correlation, though some might think of this as a non-modal connection. Nothing substantive hangs on this use of terminology.

Those who believe that grounding relations exist have resources at hand with which they can explain this phenomenon: non-accidental modal correlations are those where one entity *grounds* the other, whereas accidental modal correlations are absent any kind of grounding relationship. Moreover, in some way or another we can track what grounds what. It is less obvious that the Grounding-free Theorist can explain these intuitions. While I am confident that she can do so, explicating my preferred explanation is well beyond the scope of this thesis. I hope to be forgiven for flagging, but not dealing with, this problem, but that is the focus of another project. As I mentioned in Chapter 1, the conclusions of the thesis are provisional on the success of that project.

Thus, I now narrow my focus to the priority intuitions.

2.1.2 Priority Intuitions

As well as having intuitions about the accidental (or otherwise) nature of the modal correlations, many people judge the following claims to be true:

- A. The flower is red *because*²¹ the flower is maroon
- B. The bicycle exists *because* of the existence and arrangement of the wheels, spokes, handlebars, etc
- C. <a man exists>²² is true *because* Pythagoras exists
- D. {Pythagoras} exists *because* Pythagoras exists
- E. <Pythagoras exists> is true *because* Pythagoras exists
- F. God loves X *because* X is good²³
- G. [Pythagoras exists]²⁴ obtains *because* Pythagoras exists
- H. $2+2=4$ *because* 2 exists and 4 exists

whilst judging the following to be false:

- a) The flower is maroon *because* the flower is red
- b) The arrangement of wheels, spokes, handlebars, etc exists *because* the bicycle exists
- c) Pythagoras exists *because* <a man exists> is true
- d) Pythagoras exists *because* {Pythagoras} exists
- e) Pythagoras exists *because* <Pythagoras exists> is true
- f) X is good *because* God loves X
- g) Pythagoras exists *because* [Pythagoras exists] obtains
- h) 2 exists and 4 exists *because* $2+2=4$

Let's call the judgements that claims like (A) through (H) are true, while (a) through (h) are false, *priority intuitions*. The observation that we have these intuitions constitutes the

²¹ I use 'because' as a neutral way of expressing these claims (i.e. a way that does not commit one to thinking there are grounding relations). Those who think that we need to posit a relation of grounds to explain these (and other) cases, will rearrange the relevant sub-sentential phrases and read 'because' as 'grounds'.

²² I use <P> to indicate the proposition that P.

²³ Debates about the truth of (F) date back (at least) to the scenes described in Plato's Euthyphro dialogue (the dialogue reportedly took place in 399BCE. Translation published 2002). The truth of (F) is contentious, with some believing (f): X is good *because* God loves X. Importantly, those disposed to endorse (F) will deny (f), and *vice versa*. I will account for this divergence in Chapter 4. It is less clear that defenders of grounding can do so.

²⁴ I will use [square brackets] to indicate facts, which I take to be structured entities comprised of objects, properties and relations.

second explanandum. This list of priority intuitions is, of course, far from exhaustive.²⁵ Yet it serves to give a flavour of the kind of intuitions that need to be explained.

Notice that the priority intuitions have many commonalities. Firstly, they concern *non-diachronic* connections, as there are no cases where one relatum exists at t_1 and the other exists at t_2 . Intuitions regarding *diachronic* priority and dependence are another kettle of fish, and are not taken to be evidence for the existence of grounding relations, but evidence for the existence of causal relations, which I assume to be diachronic.²⁶ Some of the listed cases are *synchronic*, such that each relatum exists at the same time. This is so for cases (A) and (B) (and perhaps (G)). The bicycle and its parts exist simultaneously, and thus the intuition of priority between the bicycle and its parts is an intuition of synchronic priority. Likewise, the flower manifests the property of being red and the property of being maroon at the same time. Other cases are not best described as synchronic, for at least one relatum is an abstract object, and is therefore not best thought of as existing at a time. For instance, it is odd to attribute a temporal location to a set, or the truth of a proposition, or God's attitudes. As such, I think of these cases as *atemporal* rather than synchronic.²⁷ Thus, I am using the term 'non-diachronic' to capture

²⁵ The following is a longer—yet clearly still not exhaustive—list of the kind of claims one sees in this literature. Some of these claims are more plausible than others. As such, I am not endorsing the below list—its role is merely illustrative:

- The champagne glass is fragile because of the crystalline bonds between its component molecules
- Action X is wrong because it will cause a great amount of suffering and no happiness
- James is in a state of pain because his C-fibres are firing
- Dispositional properties are *posterior to* categorical properties
- Mental properties are instantiated *in virtue of* physical properties
- Singleton sets are *grounded* in their urelements
- <the rose is red> is true *because* the rose is red
- Wholes are metaphysically *explained by* their parts
- Smiles ontologically *depend on* mouths
- Sets ontologically *depend on* their members
- Events or states of affairs *ontologically depend on* their participants
- Chemical substances *ontologically depend on* their molecular/atomic constituents
- Tropes *ontologically depend on* their bearers
- Aristotelian universals (e.g. redness) *ontologically depend on* their bearers (e.g. objects that are red)
- Holes *ontologically depend on* their hosts
- Boundaries *ontologically depend on* their hosts
- The fact that this action is wrong obtains *in virtue of* the fact that it was done with the sole intention of causing harm
- The fact that this particle is accelerating obtains *in virtue of* the fact that it is being acted upon by some net positive force
- Endurantists believe that an entity's existence is *ontologically prior to* the existence of its temporal parts, perdurantists believe that the temporal parts are *prior to* the whole

²⁶ Some, such as Wilson (forthcoming) have argued that the relation the priority intuitions are tracking is a kind of metaphysical causation.

²⁷ One might have background views about these cases that make it less odd to think of these cases as synchronic. For example, if sets are identical to their members then they occupy a temporal location. Nothing in what follows hangs on a distinction between intuitions of synchronic and atemporal priority.

both synchronic and atemporal cases, and rule out intuitions about diachronic causal dependence.

Secondly, these judgements involve a kind of asymmetry and irreflexivity. When I talk of the *asymmetry* of the priority intuitions, what I mean by this is that we do not observe a pair of priority intuitions such that we intuit both that ‘x *because* y’ is true, and that ‘y *because* x’ is true. We judge that (A) is true and (a) false, and so on for the other cases. Likewise, when I say that there is an *irreflexivity* of priority intuitions, what I mean by this is that we never observe the intuition that for some x, ‘x *because* x’ is true.

Thirdly, when we intuit the truth of ‘x *because* y’, we seem to be intuiting the truth of a certain kind of explanation; what I call *metaphysical explanation*. Let’s use some fresh examples to highlight this (these cases are drawn from the very same pool of cases that generated (A) through (H)). Consider:²⁸

1. Action X is wrong *because* it will cause a great amount of suffering and no happiness
2. The composite object James exists *because* of the existence and arrangement of the simples that compose him
3. James is in a state of pain *because* his C-fibres are firing
4. {Socrates} the singleton set exists *because* its urelement (Socrates, the man) exists
5. The proposition <A friendly Airedale terrier exists> is true *because* of the existence of Jasper, the friendly Airedale
6. The champagne glass is fragile *because* of the crystalline bonds between its component molecules

These propositions are naturally thought of as expressing explanations, not least due to the tell-tale term ‘because’, which tends to be a reliable indicator of the presence of an explanation. Moreover, some theorists of explanation think that explanations are answers to why-questions (van Fraassen, 1980), and each of the above can be broken down into a why-question and associated answer. For example, (1) can be broken down into: ‘why is

²⁸ I foresee the complaint that (1) through (6) are less universally accepted than (A) through (H). Though, for example, (1) is more plausible than (F) above (for those of us who are not theists), given the perennial and seemingly irresolvable disagreements between deontologists, consequentialists, virtue ethicists, etc., one might insist that there is no convergence on claims like (1). Thus, I should reiterate that my assumption of the convergence on the truth of these claims is friendly to my opponent, and that I have the resources to account for diverging intuitions (Chapter 4) where my opponent arguably does not (Chapter 3).

action X wrong?’ and ‘because it will cause a great amount of suffering and no happiness’. The same interpretation can be applied, *mutatis mutandis*, to (2) through (6).

The cases can also be given a contrastive treatment, which is a further indication of the presence of explanation (van Fraassen, 1980; Schaffer, 2005). Consider example (2):

The composite object James exists (rather than some other composite object existing) *because* some simples are arranged James-wise (rather than some other way).

Or:

The composite object James exists (rather than fails to exist) *because* some simples are arranged James-wise (rather than those simples not existing).

In the case of example (3):

James is in a state of pain (rather than pleasure) *because* his C-fibres (rather than his dopamine pathways) are firing.

Or:

James is in a state of pain (rather than Sam being in a state of pain) *because* James’ C-fibres (rather than Sam’s C-fibres) are firing.

In the scientific context, the use of contrast classes allows us to focus upon a particular aspect of an explanation. That this technique is of similar use here is indicative that we are, indeed, dealing with a kind of explanation.

Furthermore, debates about the more contentious claims above can be very naturally framed in terms of what explains what. Ethicists, for example, argue at length about precisely what explains why an action is wrong or right (though, non-naturalists aside, they tend to agree that it is some non-moral feature of the world which explains the wrongness or rightness of the action). Whilst (1) above reflects a broadly utilitarian view of first-order ethics, a deontologist could easily substitute her explanation of the wrong-making features of the action. Perhaps X is wrong because X is a murder. These debates are disagreements about what explains the moral properties of actions.

Consider, now, example (2), drawn from the domain of mereology. Mereological nihilists will deny that James, the composite object, exists. However, those universalists

and restrictivists who do accept his existence think that it needs an explanation: James' existence is explained by the existence and arrangement of the parts from which he is composed, and the same applies to any composite object.²⁹

Moving into the philosophy of mind, consider (3). Dualists, of course, won't agree that a mental state of being in pain requires explanation solely in terms of a brain state. Yet a certain breed of non-reductive physicalist will suppose that there are mental states that are distinct from physical states, such that the former explains the latter. There *is* a mental state as of pain, but it is not unexplained. There is only pain because of the physical goings-on in James' brain.³⁰ Case (4) provides us with an example from set theory. On the *iterative* conception, a set is built from its elements, and thus the existence of a set is explained by the fact that its element(s) exist. The existence of a set is not inexplicable.³¹

Example (5) is drawn from the truthmaking literature. Truthmaker theorists contend that 'truth depends upon being', in the sense that whenever some proposition is true, there is some worldly entity that makes it true (Armstrong, 2004).³² These theorists also contend that the truth of the proposition is *explained* by the existence of that entity (see McFetridge, 1990; Liggins, 2005). Finally, (6) is an example of a dispositional property being explained by a categorical property. There is debate about the direction of explanation in such cases. For example, dispositionalists seek to explain categorical properties in terms of dispositional ones.³³ Regardless, both parties to these disputes agree that there are explanatory relationships between these kinds of properties.

In sum, a *prima facie* case can be made that the observed intuitions have characteristics associated with explanation, and, certainly, the theorists who engage with these kinds of claims treat them as explanations. Of course, this leaves it open that there

²⁹ Notably, if priority monism (as defended by Schaffer, 2010) is true, then, while there is an explanatory relationship here, it runs the opposite way to the direction presupposed above: it is the existence of James that explains the existence of his parts.

³⁰ If one subscribes to the mind-brain identity theory then there is either no explanation to be had here, or this is an unusual case of reflexive explanation. See Jenkins (2011) and §3.5.

³¹ The exception to this rule is the empty set, which has no element to explain its existence. The empty set is best thought of as fundamental/unexplained.

³² This idea has roots going back to Aristotle, who claimed that "[I]f there is a man, the statement whereby we say that there is a man is true, and reciprocally—since if the statement whereby we say that there is a man is true, there is a man. And whereas the true statement is in no way the cause of the actual thing's existence, the actual thing does seem in some way the cause of the statement's being true: it is because the actual thing exists or does not that the statement is called true or false." (Aristotle, 1984:22, quoted in Schaffer, 2016).

³³ See, e.g., Mumford (1998) for discussion.

are good reasons to suppose that these are not genuine cases of explanation at all, and hence not cases of metaphysical explanation. I will not take a stand on this. I merely seek to point out that some will think that the priority intuitions are intuitions about what explains what, and such folk will expect particular features from an explanation of the priority intuitions (see §2.4). In what follows, I will take seriously the idea that the priority intuitions are intuitions about metaphysical explanation. This is charitable to my opponent, for if there are no metaphysical explanations, then there is no argument to be made that grounding relations are indispensable to account for their truth.

2.2 Explaining versus Vindicating

With the explananda thus presented, we can now consider what we might want from a putative explanation. Of course, simply put, what is required is a story about *why* we have the intuitions that we do. But such stories come in two varieties: what I shall call *vindicating* and *non-vindicating* explanations. The distinction, here, is between an explanation of why we have certain intuitions and make certain claims, that also accounts for the *truth* of certain propositions that those claims express, and an explanation of why we have these intuitions and make certain claims, that not only fails to account for the truth of those claims, but, by contrast, suggests that the claims are false. The former vindicates (or justifies) the intuitions, whilst the latter is a kind of *explaining away*, or in some cases, a kind of debunking. One who seeks vindication of a certain class of intuitions will consider it a constraint on what will count as a good explanation that it vindicates the relevant intuitions. Others may be content with having their intuitions explained away.

Consider Andrew, a man who believes that there is an alien in the room with him. We might discover that there really is an alien in the room, and that the alien is visible. Then we can explain why Andrew believes that there is an alien in the room: he is having a veridical visual experience of seeing it. In addition, our explanation serves to vindicate Andrew's belief, for if he believes the alien to be present because he can see it (where 'seeing' is read as a success term), then his belief is true. On the other hand, suppose that we discover that Andrew has imbibed a hefty dose of a hallucinogenic substance. This also serves to explain why Andrew believes that there is an alien in the room: he is hallucinating. This explanation, however, doesn't serve to vindicate Andrew's belief. It

explains why he believes what he believes without showing his beliefs to be true. On the contrary, it suggests that his belief is false.

Naturally, this distinction applies to putative explanations of the priority intuitions. While both kinds of explanation involve telling a story about why we have these intuitions, a vindicating explanation will also show that the propositions we intuit to be true are, in fact, true. I will use the phrase ‘vindicating our priority intuitions’ as a shorthand way of saying that an explanation is such that the set of propositions that we jointly intuit to be true, are (for the most part, at least) true. While my opponent can both explain and vindicate the priority intuitions by positing grounding relations (and piggybacking on my psychological story; see Chapter 3), the competing explanation I offer in terms of evolved psychological processes (Chapter 4) is naturally read as explaining, but not vindicating, these intuitions.

But we are getting ahead of ourselves. Before we move on to these more sophisticated explanations, I will briefly consider a natural proposal to explain, and indeed vindicate, the priority intuitions in terms of the obtaining (or otherwise) of the traditional modal relations. If this project succeeds, then the arguments presented in this thesis are redundant. So, let me show why it doesn’t.

Consider that for each priority intuition, appropriately schematised in the form ‘*x because y*’, the referent of ‘*y*’ necessitates the referent of ‘*x*’.³⁴ For instance, borrowing the examples from §2.1.2, if we look to (A) we see that every possible maroon flower is also a red flower. Being maroon necessitates being red, but not *vice versa*, since there are red flowers that are not maroon. Thus, necessitation holds non-symmetrically in this case. Furthermore, the failure of symmetry in this case seems to account for our judgement about priority. We judge the direction of explanation to be aligned with the ‘direction’ of necessitation.³⁵ Similar thoughts apply to (B). At every world in which those parts exist, and are arranged in the right way, a bicycle exists. However, there are worlds in which bicycles exist, while those parts fail to exist (or are arranged differently). Bicycles, like the property of being red, are multiply realisable. Once again, there is a modal relation—necessitation—that holds non-symmetrically. Likewise, *mutatis mutandis*, for (C).

³⁴ Though see the arguments from contingentists about grounding: Leuenberger (2014), Chudnoff (ms) and Skiles (2015).

³⁵ The direction of necessitation goes from A to B, iff A necessitates B, but B does not necessitate A.

Having noted that the priority intuitions in cases (A) through (C) are aligned with, and thus explicable in terms of, non-symmetric instances of necessitation, the idea naturally presents itself to explain and vindicate the priority intuitions in terms of the obtaining (or otherwise) of the modal relations. If we intuit that ‘x *because* y’ is true just in case y non-symmetrically necessitates x, then there is no need to appeal to something like grounding in order to explain and vindicate our intuitions. Yet, as indirectly noted by many defenders of grounding in the process of motivating the need to posit a new relation—namely ground—our priority intuitions come apart from the obtaining of the modal relations, and thus this explanation is not adequate (see, e.g. Schaffer, 2009). However, it is instructive to see why it fails. In brief, it fails because, while y necessitates x whenever we intuit that ‘x *because* y’, there are many cases where y necessitates x but we do not intuit that ‘x *because* y’. An explanation purely in terms of the modal relations fails to account for the difference between these cases. I spell this out in more detail below.

The explanation of the priority intuitions in terms of the modal relations would be a good one only if we intuit ‘x *because* y’ to be true if and only if y necessitates x. However, our intuitions often come apart from the obtaining of the modal relations. Despite the fact that the direction of necessitation in cases (A) through (C) is aligned with our priority intuitions, in cases (D) through (H) the modal relations obtain symmetrically; there is necessitation running ‘in both directions’. So, for example, the truth of <Pythagoras exists> necessitates that Pythagoras exists, and *vice versa*. Likewise, the existence of {Pythagoras} necessitates that Pythagoras exists, and *vice versa*. Yet, as noted, we intuit that (D) and (E) are true while (d) and (e) are false. The difference in intuition between the two cases cannot be accounted for in terms of the traditional modal relations. Similar concerns apply to the remaining cases.

Furthermore, thinking beyond cases (A) through (H), we find (sometimes even non-symmetric) instances of traditional modal relations that obtain even when there appears to be no priority intuition between the entities thus related. For instance, necessary existents supervene both on any other necessary existent, and, indeed, on any contingent existents. The truth of impossibility claims does likewise. Any necessary truth necessitates any other necessary truth. So, for example (on the assumption that both numbers and universals exist necessarily), the existence of *blueness* is (symmetrically) necessitated by the existence of the number 2, and indeed non-symmetrically necessitated by the existence of an arbitrarily chosen ladybug.

Yet we do not intuit, in many of these cases, that one relatum is prior to the other, despite the obtaining of a modal relation. Thus, the obtaining of the traditional modal relations fails to explain these intuitions, too. The nail in the coffin of the project of explaining priority intuitions purely in terms of traditional modal relations stems from the reflexivity of these relations. For instance, that the flower is maroon straightforwardly necessitates that the flower is maroon; but it is intuitively false that the flower is maroon *because* the flower is maroon.³⁶

Ultimately, the project of accounting for the priority intuitions purely in terms of the obtaining or otherwise of the modal relations is a failure. However, I think that the modal relations have a role to play in the best explanation of this observation, and we will revisit this idea in later chapters.

Note that the failed explanation in terms of the modal relations would, if successful, also have served to vindicate the priority intuitions. For, if those intuitions were such that they aligned themselves neatly with the patterns of modal correlation described by the modal relations, the modal relations would have constituted plausible truthmakers for claims of the form ‘x *because* y’. For similar reasons, the Sophisticated Grounding Theory provides an explanation that also serves to vindicate these intuitions.

However, in Chapter 4 I show how the Grounding-free Theory can provide an explanation of the priority intuitions in terms of the overgeneralisation of a pair of evolved psychological mechanisms. This explanation, in itself, does not do double-duty as a way of vindicating our intuitions. Indeed, I contend that my explanation will, for many readers, *undermine* the motivation to vindicate the priority intuitions. For, vindication of our intuitive judgements is not always a good thing, and we should embrace humility when faced with strong theoretical reasons to suspect that our intuitions lack veracity. The history of science is a story punctuated by intuitions that were explained, but failed to be vindicated, as we came to know more about the world. As such, for those with a disposition to give up on the veracity of their intuitions in the face of this kind of psychological explanation, Chapter 4 may constitute the end of the investigation.

³⁶ One might, legitimately, stipulate that the relevant instances of the traditional modal relations are those where the relata are distinct, and thereby rule out the reflexive instances. However, the aforementioned problems still remain.

However, I am agnostic regarding whether, when faced with the psychological cum evolutionary explanation I offer, we should give up on vindicating our intuitions. Indeed, it is something of a stretch for me (or, indeed, anyone who thinks there is a substantial epistemic element to what explains what) to insist that the claims of the form ‘x *because* y’ listed in §2.1.2, that seem very much like true explanations are, in fact, false (or, at the very least, are not true).

Some readers, for example, will think that, even if we come to our beliefs in the way I describe, this does little to undermine their truth. Such readers will still seek a story of what makes true, claims of the form ‘x *because* y’, if not grounding relations. By analogy, consider that coming to believe that one’s moral beliefs are the result of evolutionary processes need not directly lead one to a moral error theory—though it may encourage one to widen the search for moral truthmakers so as to potentially include things like the attitudes of the community in which one is embedded. I present a view roughly analogous to this in Chapter 5. Importantly, I will demonstrate that the ability of grounding-based theories to not only explain, but also vindicate, the priority intuitions cannot be parlayed into an advantage for these theories.

In sum, I don’t plan to weigh in on whether or not we should be trying to vindicate the priority intuitions. If there is no need for this vindication, then there is no indispensable need for grounding to do this work. If, however, there *is* a need for vindication, I contend that this can be done (and done well) without grounding. Indeed, my preferred explanation of the priority judgements is compatible with several ways of vindicating these judgements.

2.3 Metaphysical Explanation

For those who think that the priority intuitions are intuitions about explanation, a putative advantage of the grounding-based route to vindicating these claims is the ability to provide a systematic story about how these judgements fit together. In other words, if we vindicate the truth of the priority intuitions by appealing to grounding, we are granted a powerful tool in the form of a ready-made theory of *metaphysical explanation* (there are in fact two incarnations of such a theory; see Chapter 3). For, the grounding-based theorist can use her framework to identify other cases of metaphysical explanation—ones where we perhaps didn’t already have a priority intuition—and she can unify them all together

under the banner of grounding. The grounding-based theories, then, do not merely vindicate the intuitions, but does so in a way that has a certain theoretical utility.

Now, not everyone will see the fact that positing grounding grants one a ready-made theory of metaphysical explanation as an advantage. Some will take issue with the very notion of metaphysical explanation, and some will think that, while there are metaphysical explanations, they are nothing like the phenomenon I describe here. I have no qualms with either of these views. If, for whatever reason, one rejects the notion of metaphysical explanation characterised below, then one will not seek to vindicate the priority intuitions in a way that also generates a theory of metaphysical explanation. As such, one will have one fewer reasons to endorse a grounding-based theory.

The notion of metaphysical explanation that I will characterise below is one that *can* be accounted for in terms of grounding—it would be grossly uncharitable of me to characterise the notion some other way. For, if the notion considered here were sufficiently dissimilar to the notion that my opponent is working with, she could legitimately accuse me of changing the subject. So, the notion spelled out below will resemble the grounding-based theorist’s notion—though, as I shall argue, it need not rest on her preferred ontology. For, as I will demonstrate, the Grounding-free Theory also has the resources to vindicate the priority intuitions in a way that grants us a theory of metaphysical explanation—if that is indeed the kind of vindication that we’re after. In sum, the current characterisation is merely intended to be neutral between the theories discussed in this thesis (in that it is neutral between ontologically heavy-duty and lightweight views of explanation), not neutral between all potential ways of thinking about metaphysical explanation.

Hence, much of the below characterisation of metaphysical explanation will echo the treatment of that notion in the grounding literature. Indeed, one way to positively characterise the relevant notion is by ostension: metaphysical explanations are *the things that grounding is supposed to elucidate*, and *the things that are putatively made true by the existence of grounding relations*. Indeed—as argued in Chapter 3—the ontologically lightweight ‘sentential operator’ view of grounding closely resembles this notion of metaphysical explanation.

Ultimately, the plan is to argue that there are grounding-free ways of developing a theory of metaphysical explanation that are at least as good as the ready-made theories

we get from positing grounding. Thus, the grounding-based approach to vindicating the priority judgements gains no dialectical advantage by way of a theory of metaphysical explanation. Furthermore, the grounding-free ways of developing a theory of metaphysical explanation are embedded in a more parsimonious (as they don't require positing grounding relations) and elegant overall theory. So, my preferred explanation of the priority judgements is compatible with vindicating these judgements in a way that is consistent with a theory of metaphysical explanation.

To be clear: the structure of my argument is conditional: *even if* one seeks vindication of the priority intuitions that is compatible with a theory of metaphysical explanation of the kind my opponent desires (I am agnostic about whether we ought to seek this kind of vindication), this can be done just as well without positing grounding. In what follows, I will say a little more to neutrally characterise the metaphysical explanations. This will help to spell out the target of investigation, and thereby help us see what we might want from a theory of that thing.

Naturally, the metaphysical explanations inherit the features of the priority intuitions that we noted in §2.1.2. Thus, we can begin by noting that they are non-diachronic. Defenders of grounding contend that this is because metaphysical explanations are in the business of tracking the generation of the world. However, rather than tracking its diachronic generation, from one moment to the next, instead the thought is that metaphysical explanations track its non-diachronic generation, from the more fundamental to the less fundamental. In other words, metaphysical explanations are typically thought to move 'up' the axis that Wilsch (2015, 2016) calls the 'axis of fundamentality', with the more fundamental providing explanations for the less fundamental. Yet, the metaphysical explanations need not be understood in this way, which seems to beg the question in favour of ontic, grounding-based theories. Indeed, Wilsch characterises the idea of the 'axis of fundamentality' in terms of grounding.

Instead, we might understand the relevant notion of fundamentality in terms of metaphysical explanation, such that the term 'fundamental' in this context describes something that does not itself require a metaphysical explanation. It is explanatorily basic; primitive; unexplained.³⁷ This leaves open, of course, that what is unexplained, and

³⁷ There are, of course, other ways of spelling out what is fundamental that need not be tightly linked to metaphysical explanation. Perhaps the fundamental is the minimal supervenience base for everything else, or that which is quantified over in our microphysical theories. Thus, one who denies the truth of the

thereby fundamental, from one perspective might not match what is fundamental from another perspective.

With this notion of fundamentality in hand we can express the related notion of *derivativeness*. That which is not fundamental is derivative. Thus the derivative requires metaphysical explanation, and is ultimately explained by the fundamental. The explanandum of any metaphysical explanation must be derivative, as it is explained in terms of something else. The explanans of a metaphysical explanation may or may not be fundamental, depending upon whether it is further explained by something else.

The metaphysical explanations inherit several other features from the priority intuitions. For example, they have a particularly broad modal scope. Just as we observed that there is a traditional modal relation accompanying each priority intuition (though the inverse is not the case; §2.2) metaphysical explanations are such that every possible world in which the explanans is true is a world at which the explanandum is true also (though, as also noted, this is contentious: see Leuenberger, 2014; Skiles, 2015 and Chudnoff, ms). Thus, Trogon argues that it is in the nature of explanation to show how the explanandum “couldn’t have been otherwise” given the truth of its explanans (2013b:footnote 3) and deRossett insists that “If it is possible for the *explanans* to obtain while the *explanandum* does not, then the *explanans* does not make the *explanandum* obtain, and so the explanation fails.” (2013:15).

Likewise, metaphysical explanations mirror the priority intuitions in being irreflexive (nothing metaphysically explains itself) and asymmetric (if A metaphysically explains B, B does not metaphysically explain A).³⁸ This reflects our preferences regarding explanation more generally. We don’t find a putative explanation of a phenomenon in terms of itself to be at all illuminating (consider the banality of the exchange: “Why X?” ... “Because X”). Similar thoughts apply to the circular explanations that arise from symmetrical explanations. So, the metaphysical explanations are irreflexive and asymmetric.

While this is not explicitly captured by the priority intuitions, the metaphysical explanations also appear to admit of a distinction between whole and partial explanation

propositions expressed by the priority intuitions, even *sans* a theory of metaphysical explanation, can still meaningfully talk about fundamentality.

³⁸ See the discussion of Jenkins (2011) in Chapter 3 regarding *quasi-irreflexivity*. Likewise, Barnes (forthcoming) can be read as a defence of the symmetry of some metaphysical explanations.

(following Fine, 2012). This is useful to note here, as not all features of the former are shared by the latter. For example, while a whole metaphysical explanans should necessitate its explanandum, not so for a partial explanation. Recall that, on a roughly utilitarian picture, the causing of a great amount of suffering and no happiness is the whole metaphysical explanation of why action X is wrong. Moreover, the part of this explanans which states that action X will cause a great amount of suffering is a plausible partial explanation of its wrongness. In the mereological case, while the existence and arrangement of all of the simples composing James might constitute a whole explanation for his existence, his existence is partially explained by the existence and arrangement of the atoms that compose his head. Of note is that neither of these partial explanantia necessitate their respective explananda. In the ensuing discussion of theories of metaphysical explanation, our target will be whole explanations (unless explicitly mentioned otherwise) and the features thereof.

That concludes the neutral characterisation of the metaphysical explanations, and the explication of how one might consider it a virtue of an explanation of the priority intuitions if it also lends itself to a theory of metaphysical explanation. However, we currently lack the tools with which to compare such theories. Thus, we now move on to consider some desirable features of a theory of explanation more generally.

2.4 Desiderata for a Theory of Explanation

In order to be in a position to evaluate the relative merits of the grounding-based and grounding-free theories of metaphysical explanation, we need some independent metric against which to measure these theories. As such, here I will briefly identify some desiderata for a theory of explanation, more generally. These I have gleaned, for the most part, from the ways in which theories of scientific explanation have been criticised. So, if some theory has come under attack for failing to exhibit feature X, I take it that, for some at least, X is a desideratum. I will say a lot more about how these desiderata interact with the literature on scientific explanation in §2.5.

It will quickly become apparent that some of the following desiderata are in tension with one another. My diagnosis of this—following Bromberger (1965), Lewis (1986b) and Bird (2005)—is that there are (at least) two senses in which the word ‘explanation’ is used, which Bird (2005) calls the *subjectivist* and *objectivist* senses, but I

prefer to call the *epistemic* and *ontic* senses. In the subjectivist/epistemic sense, an explanation is something offered to an agent in order to enhance her understanding of some phenomenon. In the objectivist/ontic sense, one phenomenon can explain another without any agent ever realising that this is the case, for these explanations are there in the world, such that they obtain regardless of agents' interests or understanding.

As Bird (2005) argues, it is plausible that these senses are interrelated. For, when we offer epistemic explanations we are trying to tap into part of the objective explanatory structure of the world so as to illuminate some phenomenon for some agent. Nevertheless, that explanation is commonly thought of in these distinct ways is sufficient to explain why the desiderata pull in different directions. For, while most desiderata seem motivated by thinking of explanation in the epistemic sense, there is one that is motivated by thinking of explanation in the ontic sense.

The following desiderata are motivated by epistemic intuitions about explanation.

COVERING CASES: A theory of explanation in some domain should capture the explanations actually used in that domain. If there are explanations that are not covered, something is missing.

For example, Salmon's (1984) causal process theory is criticised for being putatively unable to cover actual explanations in terms of absences (see §2.5.2; Beebe, 2004 for discussion). Thus, I take it that covering cases of actual metaphysical explanation is a desideratum for a theory of metaphysical explanation.

IRREFLEXIVITY: Nothing explains itself. While there may be unexplained, or unexplainable phenomena, to cite some event as its own explanation is not to explain it at all.

ASYMMETRY: Explanation is an asymmetric affair, such that if A explains B, it is not the case that B also explains A.³⁹

The motivation for irreflexivity and asymmetry is that we, as agents, find self-explanations and circular explanations to be unhelpful in illuminating why some phenomenon is the way that it is. For instance, Hempel and Oppenheim's (1948) deductive-nomological theory of scientific explanation is criticised (by, for example,

³⁹ Note that asymmetry entails irreflexivity, as a self-explanation is a case where both A explains B, and B explains A—it's just that $A = B$.

Salmon, 1984) for being unable to rule out symmetrical explanations (§2.5.1). Thus, I take it that asymmetry and irreflexivity are desiderata for a theory of explanation.

UNDERSTANDING: The explanans should shed light upon the explanandum, making the latter more intelligible. This is the most explicitly epistemic element of explanation, whereby the phenomenon in question should be illuminated by the explanans.

RELEVANCE: The entirety of the explanans should be relevant to the explanandum. A perfectly good explanation can be ruined by the addition of extraneous, irrelevant information.⁴⁰

These candidate desiderata once again reflect an inclination to think of explanations as the kinds of things that have a certain psychological effect upon agents. So, for example, the deductive-nomological theory is often criticised for allowing that the event of a storm can be explained in terms of barometer readings, which we intuit to be irrelevant (see §2.5.1).

The final candidate desideratum is strongly suggested by the ontic understanding of explanation.

OBJECTIVITY: What explains what is not an interest-relative or subjective matter. Some parts of the world explain other parts, and this is simply a matter of how the world is.

It is a desire for objectivity that motivates theories of explanation framed only in terms of laws of nature or causal processes—things that exist in a mind-independent way—rather than theories that appeal also to contingent psychological features of agents.

With these desiderata in hand, there are three points I want to emphasise. Firstly, I take the above list to be defeasible, and I am not trying to convince the reader to endorse any desiderata they do not already find intuitively appealing. As these are merely candidate desiderata, I am happy to allow that each of us will place different weight upon each desideratum, perhaps placing no weight on some, while maintaining that others are non-negotiable. Again, my intention here is not to argue that any desideratum is *required*, nor indeed that the reader should have a particular attitude to any of them. I merely seek

⁴⁰ The question of whether there is a mind-independent, non-anthropocentric notion of relevance is a difficult one that I won't try to resolve, here.

to identify some features which are often deemed good-making features of a theory of explanation. Ultimately, which theories of explanation one finds attractive will depend upon how one weights the desiderata.

Second, the *reason* I can stay so agnostic about whether these candidate desiderata are genuine desiderata is that there is no view on the desiderata which would lead one to prefer a grounding-based theory of metaphysical explanation over the alternatives I offer in Chapters 5, 6 and 7. If one held the view that *all* of the above desiderata must be satisfied, one might initially be attracted to the unfiltered grounding-based theory—indeed, in the literature on grounding and metaphysical explanation, it *does* seem that defenders of grounding want every desideratum to be met. This is a thought to which we shall return, for, as I shall argue in Chapter 3, the unfiltered grounding-based theory of metaphysical explanation seeks to satisfy all of the desiderata in an ambitious, but ultimately implausible, way. Thus, I hope to later show that given any plausible view on the desiderata, there is a grounding-free theory of metaphysical explanation that is preferable to the grounding-based alternative. Of note is that it is exceedingly difficult to simultaneously satisfy both the epistemic and ontic desiderata: those seeking to satisfy all of the above desiderata are asking a lot, as will emerge in the below discussion of theories of scientific explanation. Perhaps these theorists of scientific explanation realise that it is a fool's errand to attempt to satisfy desiderata derived from differing conceptions of explanation, and thus seek to strike a balance between them.

Thirdly, I want to point out that these desiderata are aligned with features of the priority intuitions. For example, there are no reflexive or symmetrical priority intuitions. Likewise, in each case of 'x *because* y', y seems relevant to, and to increase our understanding of, x. Significantly, however, there is nothing in the priority intuitions that explicitly implies objectivity (though one might have the further intuition that it is a mind-independent truth that 'x *because* y', as reported by defenders of the unfiltered grounding-based theory).

Via consideration of these desiderata, I will show, in Chapters 5, 6 and 7, that the Grounding-free Theory has the resources to vindicate our priority intuitions in a way that also provides the theoretical utility of a strong theory of metaphysical explanation. Recall that, as long as I can show that my preferred theories of metaphysical explanation are at least as good as the grounding-based theories, I can show that my account is better overall, since it is more parsimonious not to posit grounding relations.

Before moving on, however, we conclude this chapter with a detour via some theories of *scientific* explanation. The purpose of this detour is not to argue in favour of any one of the following theories. Instead, it will play the following three roles. Firstly, it will identify a useful way to categorise theories of explanation—one to which we shall often return in the following chapters. Secondly, it will identify some prototypes of various kinds of theory of explanation, to be used as templates for theories of metaphysical explanation in the following chapters. Thirdly, it will illuminate the tension between objectivity and understanding, and thus illuminate the tension between the epistemic desiderata, on the one hand, and the ontic desideratum, on the other.

2.5 Scientific Explanation

As I will discuss in Chapter 3, a cursory read of the literature on grounding suggests that a grounding-based theory of metaphysical explanation is the only viable option, given that accounting for such explanations in terms of the modal relations fails (§2.2). The received wisdom is that, insofar as we want to take metaphysical explanations seriously, we must posit purpose-built grounding relations. This is false. A quick review of the literature on scientific explanation reveals that there are several plausible avenues through which to develop a theory of explanation. Of particular note is that there are several powerful and useful theories of scientific explanation which have the potential to be co-opted, with minor tweaks, to apply in the metaphysical case. Moreover, these co-opted theories do not require positing grounding. Here I will briefly explore the terrain populated by theories of scientific explanation.⁴¹

As a primary goal of this section is to get clear on some resources that we can call upon in the development of grounding-free theories of metaphysical explanation, I will not only characterise the theories but evaluate them against the desiderata identified in §2.4. This will serve several purposes. First, it will serve to flag potential challenges to the metaphysical variants of these views. As I lack the space to comprehensively perform this evaluation, substantial discussion of controversies regarding whether a particular view of scientific explanation can fulfil a particular desideratum will not make an

⁴¹ As a preliminary, note that the views considered here are not concerned with giving an account of ‘explanation’ which does justice to all of the diverse ways the word is used. In the present context, ‘explanation’ is being used continuously with its use in the literature in the philosophy of science, which tries to capture a notion of explanation of particular interest to scientists and philosophers.

appearance. Rather, where there is controversy I will reference some relevant work. Ultimately, establishing whether the scientific variants of these views fulfil the desiderata is not required to achieve the goals of the thesis. Far more substantial are the equivalent questions about the metaphysical variants.

Second, it will serve as a useful exercise in the application of the desiderata to some theories of explanation, to give the reader a feel for how the comparison of the theories of metaphysical explanation will proceed. And thirdly, comparing the theories against the desiderata will reveal the aforementioned tension between the ontic desideratum and the epistemic desiderata.

Before getting into the nuts and bolts of the theories, I want to emphasise a common thread regarding the way in which they cleave the class of genuine explanations from those putative explanations which fail to ‘make the cut’. It will emerge that this cleaving tends to take the form of a two-step process. Thus, a comprehensive theory of explanation will often play two roles.

In the first instance, such theories identify a determination or dependence relation *in re*—a candidate explanatory relation. Explanations are thought to ‘track’ such relations. Examples include the nomic determination relations defended by Hempel & Oppenheim’s (1948) deductive-nomological (DN) theory (§2.5.1), the spatially continuous causal processes appealed to by Salmon’s (1984) causal process theory (§2.5.2) and the counterfactual dependencies (§2.5.3) espoused by Lewis (1986b) and Woodward (2003). In what follows we will examine the candidate determination relations identified by each of these theories, considering the intuitive and counterintuitive consequences in each case.

In the second instance, a theory of explanation might provide some way of distinguishing, amongst the instances of the candidate relation identified, those which are explanatory from those which are not. Think of this second stage as a kind of ‘filtration’ of the candidate relation. This stage is optional, in the sense that one might think that *every* instance of a particular candidate determination relation is explanatory. I call such views ‘unfiltered’.

However, in response to the charge that unfiltered views cast the explanatory net too wide, some method of filtering the candidate relations is often deemed desirable. One method of filtration is to identify some further objective feature which distinguishes

the explanatory from the non-explanatory instances of the candidate relation. This is the route two notable developments of the DN theory have taken (namely, Strevens' (2008) Kairetic theory and Kitcher's (1981, 1989) unification theory). Alternatively, the filter might appeal to psychologistic constraints, picking out only those instances of the candidate relation which illuminate, or increase *our* understanding of the explanans. Psychologistic filtering is naturally paired with the view that the truth or falsity of explanatory claims is a relative matter. The most extreme view of this kind is the pragmatic theory defended by van Fraassen (1980), which, uniquely, does away with the first stage of identifying candidate relations.

This way of thinking about the theories of explanation ties in nicely with Bird's (2005) conciliatory proposal. The views that do not appeal to anything psychologistic can be thought of as a kind of ontic, *in re* explanation, and will be objective. In contrast, the psychologistically filtered views correspond with the epistemic sense of explanation, and are oriented towards those putative explanations which increase our understanding.

The sketch of these theories, which follows, will helpfully illuminate the space of possibilities for a theory of scientific explanation. First, something will be said of the nature of the various candidate relations. Secondly, we will look at how these candidates might be filtered or left unfiltered. In doing so we provide a model for the roles a theory of explanation might play, and thus set the scene for exploring potential adaptations into theories of metaphysical explanation.

2.5.1 Deductive-Nomological Theories

The DN theory, first proposed by Hempel and Oppenheim in 1948, kickstarted the modern debate about the nature of explanations in science. Subsequent theories have, to a large degree, been developed as a response to the recalcitrant counterexamples to the DN theory. An introduction to the theory is aided by noting the ideological background of its creators. While it is a natural thought that to explain some phenomenon is to identify its causes, Hempel and Oppenheim's positivistic leanings made them highly sceptical that we could come to know about dubious ontological posits such as relations of causation. Thus they excised any mention of causation from their theory, instead presenting explanations as *arguments* from explanans to explanandum. These arguments are deductively valid (hence 'deductive') and include a

law of nature (hence ‘nomological’), usually coupled with a description of some state of affairs. The state of affairs together with the laws entails the explanandum, and each premise must be required for this derivation.

In this way the explanatory burden is shifted away from an ontology of causal relations and onto an ontology of laws of nature, a notion which proponents of the view found more continuous with scientific discourse.⁴² As such, this theory holds that explanations track relations of *nomio determination*. It is the notions of logical structure and laws of nature doing the work in cleaving the potentially explanatory relations from the rest. We will consider four incarnations of the DN theory here: the unfiltered view, and three methods of filtering—the psychologistic DN theory, the Kairetic theory, and the unification theory.

It will be no surprise that the DN views have been heavily criticised for a lack of an adequate account of laws of nature (see, e.g., Cartwright, 1983. Even Hempel himself voiced the complaint that characterising the laws has proved “highly recalcitrant” (1965:338)). Hempel and Oppenheim characterised the laws of nature as ‘non-accidental generalisations’. This is a notion we have an intuitive grasp on, but is notoriously difficult to nail down. Hempel (1965) distinguishes the generalisation that ‘all members of the Greensbury School Board for 1964 are bald’ from the generalisation that ‘all gases expand when heated under constant pressure’. The latter strikes us as lawful while the former does not. As our intuitions are clear in cases such as this, Hempel ultimately takes a ‘we know ‘em when we see ‘em’ approach. While in many cases this may be true, the lack of a precise account of the laws of nature remains a problem for DN theories, one to which we will return in Chapter 6.

The DN theories do well with covering cases. Many actual scientific explanations explicitly feature the DN form, and for those that do not (for example, those which lack laws), it is open to the defender of the DN theory to claim that underlying the proposed explanation is a DN derivation. As long as there is a DN explanation to be had with the same events in the explanans and explanandum, the theory has the case covered. It also does well with irreflexivity, as presumably there is no law of nature connecting an event with itself.

⁴² One might demur on the grounds that talk of causal processes is as much a staple of scientific discourse as talk of laws of nature. Nothing in what follows hangs on whether Hempel and Oppenheim were *right* to think that laws are less metaphysically suspect.

However, DN theories struggle with asymmetry. The problem derives from the apparently symmetrical characteristics of the laws of physics (see e.g., van Fraassen, 1989). These lawful symmetries allow many DN derivations where the explanans and explanandum are inter-derivable from the same law. For example, the current arrangement of our solar system is equally derivable from how it was arranged yesterday (in combination with the laws) and how it will be arranged tomorrow (in combination with the laws). Yet many feel the intuition that the latter ought not count as an explanation.

Furthermore, there is a problem with relevance. The minimal characterisation of the laws of nature allows seemingly irrelevant information to be included in the explanans, yet the argument still be deemed an explanation. To use another classic example, barometer readings are reliably (and non-accidentally) correlated with storms. Thus one can construct a DN argument whereby the barometer's reading explains the occurrence of the storm (via a 'barometer-storm law'), despite a strong intuition that both the reading and the storm were explained by the change in air pressure.⁴³

Finally, the DN theory allows for explanations that are beyond human comprehension and thereby fail to increase our understanding. For, there is surely an immensely complex DN derivation from the initial state of affairs of the universe, together with the laws of nature, to any current state of affairs. For example, the fact that the curtains in my study are black can be derived in this way.

The different sub-categories of DN theory deal with these concerns in different ways. The unfiltered DN theorist, for example, will bite all the above bullets. She will maintain the objectivity of what explains what, at the cost of abandoning asymmetry, relevance and understanding. This move is part of a more general strategy of shrugging off objections that the candidate relations outstrip the intuitive explanations by appealing to ontic intuitions about explanation. If explanation is an objective feature of the world, it is plausibly one to which our intuitions provide a poor guide.

At the other end of the spectrum, the psychologistic DN theorist will say that only those DN derivations that fulfil the epistemic desiderata should count as explanations.

⁴³ This kind of irrelevance is not of the 'tacking on idle wheels' kind, but rather a case where the entire explanation appears to be idle. Irrelevancies of the former kind are avoided by appealing to the principle of *minimality*, whereby every premise in a DN derivation must be required for the validity of the argument. More will be said about minimality in Chapter 6.

Thus, she will jettison the objectivity of what explains what in order to have a theory that excludes symmetrical and irrelevant derivations, as well as those that fail to increase our understanding. This kind of position is motivated by arguments such as van Fraassen's (1980) that what explains what is determined contextually, by the interests and knowledge of the question-asker. Thus the view is explicitly anthropocentric, and the truth of explanatory claims is relative. However, such a position needn't appeal to the contexts of *individuals*, like van Fraassen's theory. Presumably there is some unity in the kinds of things that humans find explanatory. So we can think of the psychologistic filtration here as representing human interests and understanding. There is evidence that Hempel himself was sympathetic to such a view, as he responded to the symmetry counterexamples (e.g., the famous pole and shadow case) by relegating the choice between the two directions of derivation to 'pragmatics' (1965).

Each of the remaining pair of DN views tries to ride the line between the above approaches, maintaining the objectivity of what explains what without allowing too many counterintuitive explanations. Strevens' (2008) Kairetic account, for example, appeals to further in-the-world considerations to filter the candidate relations provided by the DN theory in a non-psychologistic way. The Kairetic view has it that in addition to satisfying the DN argument form, a putative explanation must also represent an underlying asymmetric relation of determination in the world (including, amongst others, causal relations). The relata of this relation must be represented by the non-lawful premises and conclusion of the argument. The DN derivations which fail to track these relations are not explanations.

As such, the Kairetic account (in the scientific context) is a hybrid of the DN theory and the causal process theory (see §2.5.2), insofar as both ingredients are required. The result of this hybridisation is a powerful theory, but one with twice the baggage. Not only does the Kairetic theory need to give an account of the laws of nature, it needs to give an account of the asymmetric determination relations which must exist *in re* alongside every explanatory DN derivation. In this way, the 'dubious ontology' which Hempel and Oppenheim were trying to do away with has been readmitted, and therefore the Kairetic theorist finds herself needing to give an account of both laws of nature *and* causal relations.⁴⁴

⁴⁴ Strevens (2008) explicitly denies the latter point, for he takes his theory to be agnostic between all theories of causation.

However, if one is happy with the commitments of the view, it achieves asymmetry in virtue of the asymmetric nature of causal relations. However, the solution is not so simple with regards to relevance, as there are counterexamples to both the DN and causal process theories which will still be problems for the Kairetic theory (e.g. the man who takes birth control pills discussed in §2.5.2). Finally, the Kairetic theory still allows those DN derivations which fail to increase understanding. For example, those derivations from the initial state of the universe to some current state of affairs will still have an underlying causal relation. Nonetheless, the Kairetic theory remains a powerful and attractive theory.

Finally, Kitcher (1981, 1989) advocates an alternative way of filtering the DN derivations without appealing to psychologicistic considerations. According to the unification view, the only DN derivations that count as explanations are those which belong to the set of arguments that *best unifies* all the explananda which need explaining. As such, whether a certain derivation is explanatory cannot be assessed in isolation. Instead, entire corpora of potential explanations are compared against each other to see which overall system is the best. Kitcher uses a lot of complex machinery to make precise this notion of the most unified system. We will explore the machinery in more detail in Chapter 7, but (in brief) the best system is the one which uses the *fewest* argument patterns to generate the biggest conclusion set. These patterns must also be *stringent*, such that derivations which instantiate the same argument pattern are genuinely similar.

The resulting theory implies that even if there is a DN derivation in the offing for some phenomenon, if this derivation instantiates a new argument pattern, and the phenomenon is already derivable by an instance of a pre-existing argument pattern, then this putative new explanation doesn't count as an explanation. So, for example, the unificationist can avoid temporally symmetric explanations by noting that if we already use past-to-future explanations, adding future-to-past explanations decreases the unification of our system. However, this idea of 'already using' explanations from past-to-future is misguided (as we discussed in §1.2). A systematisation which makes use of either the past-to-future or future-to-past pattern will be equally unified, for the machinery provided by the theory can only rule out that they exist side-by-side. Thus, arguably, asymmetry is satisfied, but it remains underdetermined whether past-to-future

or future-to-past explanations are correct. Thus, Kitcher may have to follow Hempel's lead in relegating the decision to pragmatics.⁴⁵

The constraints of unification also prevent irrelevant premises sneaking into the DN arguments, satisfying relevance. If we can already account for some phenomenon using a derivation pattern, a putative explanation of that phenomenon using some other pattern with idle premises will be ruled out. Again, this will be presented in much more detail in Chapter 7. However, nothing in the view rules out incredibly complex DN derivations which fail to increase our understanding.

So, there are four DN-based views of scientific explanation, which perform differently according to the desiderata. Which view does best overall will depend, of course, on how one weights the desiderata. Notably, for those who emphasise the epistemic nature of explanation, the psychologistic DN view can fulfil *all* of the other desiderata by giving up on objectivity. Of the objective theories, the unification view does best in securing the epistemic desiderata, but (arguably) still falls short.

2.5.2 Causal Process Theories

It is a natural thought that the counterintuitive instances of DN explanations arise precisely because, in doing away with the notion of causation, significant causal considerations have not been respected. Thus, while perhaps the current arrangement of the solar system is nomically determined by both the future arrangement and past arrangement, only the past arrangement *caused* the present arrangement. In this way our intuitions against temporally symmetrical explanations can be preserved. Similarly, what is going on in the case of the barometer and the storm is easily explicable via an appeal to causal notions. The barometer reading does not cause, and thus does not explain, the storm. Rather, the storm and the reading are both caused (and explained) by the change in air pressure.

This reaction to the DN counterexamples has motivated those with less positivistic tendencies to posit theories of explanation which make explicit appeal to causal relatedness. Wesley Salmon (1984), for example, argues that many scientific explanations track ontologically upstanding, worldly relations: namely the causal relations. For Salmon,

⁴⁵ See Barnes (1992) for a criticism of the ability of the unificationist picture to give us asymmetry.

the cause—and thus explanation—of an event is cashed out in terms of spatiotemporally continuous causal processes which “transmit causal influence from one part of spacetime to another” (Salmon, 1984:297) and causal interactions themselves, which occur when multiple causal processes intersect in spacetime and impact one another in a lasting fashion. The lasting impacts of these intersections are known as ‘marks’. Thus the explanation of an event will involve a story about the marks which led to the event.

Causal processes are distinguished from ‘pseudo-processes’ insofar as genuine causal processes have the ability to transmit marks. For example, if I am driving down the road and my car runs into a pole, the pole will be ‘modified’ in a lasting fashion: it is now dented. Meanwhile, the car’s shadow, making its way down the road next to me, might ‘hit’ any number of objects. However, when the shadow moves on the object has not been modified, and this inability to leave marks identifies it as a pseudo-process. This characterisation makes clear that there is a counterfactual element to the notion of a causal process. A lonely electron travelling through the outer reaches of space without ever encountering another object counts as a causal process because *if it had* encountered something it would have marked it.

A more modern incarnation of this kind of view is found in Dowe’s (2000) conserved quantity theory. Dowe appeals to the notion of ‘conserved quantities’ found in physics, such as momentum. In this way the nature of the marks required for a causal interaction can be made clearer: when the cue ball transfers some of its momentum to the eight ball, a causal interaction has taken place.

I’ll remain neutral between these views, simply assuming that some satisfactory story about the causal processes themselves is available. Nevertheless, there are two distinct incarnations of the causal process theory, understood as a theory of scientific explanation, that require our consideration, here. The first is the unfiltered theory, which has it that any putative explanation which tracks a causal relation counts as an explanation. The second is the psychologically filtered causal process theory, which demands that the putative explanations also have the right sort of epistemic character. That is, which causal relations are explanatory is relative to a psychological orientation.

The causal process theories both do well with covering cases. Even in the cases where, as it happens, scientists tend to explain some phenomenon by reference to a natural law, as long as there are causal processes running from the explanans to

explanandum the causal process theorist can say she has the case covered.⁴⁶ Assuming that nothing causes itself, the views also do well with irreflexivity. Furthermore, the asymmetry of most accounts of causation is inherited by the attendant theory of scientific explanation.

However, the unfiltered causal process theory struggles with relevance. Even if the barometer case is dealt with, it does less well with another famous counterexample to the DN view, ironically posed by Salmon himself (1971). Consider that it may be a law that no man who takes birth control pills will fall pregnant: it's a generalisation, and is surely no accident! Thus there is an argument from the fact that a man called Jones takes birth control pills, via this law, to the conclusion that he does not fall pregnant. The problem, of course, is that the birth control pills seem irrelevant.

However, (following Hitchcock, 1995) consider that there are spatiotemporally continuous causal processes running from the taking of the pills through to their digestion and the absorption of their ingredients, to the eventual non-pregnant state of Jones. Whatever concerns we might have about explaining the lack of pregnancy in terms of the birth control pills on the DN view are also concerns for the causal process view, so long as there is no way of distinguishing the relevant causal processes from the irrelevant ones. Indeed, on this view any event in the past light cone of the explanandum from which we can trace a causal process counts as part of the explanans.

Similar considerations apply to the understanding desideratum. For example, the initial state of the universe exerted causal influence that is still shaping events today, and yet including these ancient influences in an explanation does little to increase our understanding of why current events are as they are. Ultimately, however, the defender of the unfiltered causal process theory will likely espouse an ontic understanding of explanation, and will thus see these apparent shortcomings as a feature, rather than a flaw (or, at least, she will dismiss them as insignificant). The complete explanation will, in Salmon's terms, have situated the explanandum within a causal nexus, and include all of the causal information leading up to its occurrence. While this story will do nothing to *highlight* some process we intuit to be in need of highlighting, it is a complete story. The inability of our finite minds to comprehend some elements of the explanantia, and our tendency to focus on some elements thereof, is completely beside the point. Of course,

⁴⁶ Arguably, causal process theories struggle to handle cases of absence causation (though see Beebe, 2004).

the view does very well according to objectivity. Thus, if one prefers an ontic view of explanation to an epistemic one the view remains attractive.

However, if one does not share the conviction that explanation is a worldly phenomenon, divorced from our interests and comprehension, is it natural to give up on objectivity and apply some kind of psychologistic filtration to the causal process theory. Thus, while explanations will still cite causal processes, not all such processes will underlie an explanation. The addition of such pragmatic considerations allows us to attend to different causal processes in different contexts, perhaps using machinery similar to van Fraassen's. This helps us achieve relevance. Jonathan Schaffer (2005), for example, advocates a contrastive treatment of causation. Consider again the case of Jones. We might ask: why is Jones not pregnant, as opposed to being pregnant? Simplifying greatly, the answer is not that Jones took birth control pills, as opposed to not taking the pills, but that he is male rather than female. As Jones is male, the causal influence of the birth controls does not favour the former over the latter, and so despite the presence of a causal process this does not explain his lack of pregnancy.

This view also has the power to rule out explanations that are incomprehensible, and fail to increase our understanding. In most contexts, for example, causal processes which began a long time before the explanandum event will not help us comprehend why the event occurred. However, this will vary according to context. For example, imagine that we are interested in why a player was injured in a football game. We gain the most understanding by focussing on the moments just before the injury happened: exactly how the tackle occurred, the way in which the ankle twisted, etc. Perhaps we also gain some understanding by looking a little earlier: the warm-up before the game, the recent training schedule. If we are interested in why *that* player was injured, while another player involved in a similar tackle, with the same training schedule etc was not, we might look back a little further, even to genetic influences on how their bodies developed. It seems that there is no context, however, in which the initial conditions of the universe can shed light on why the injury occurred. In contrast, if we are interested in the question of why the universe is currently expanding, it seems that an understanding of the causal processes emanating from the early events of the universe *does* become relevant, and *does* serve to increase our understanding. For those interested in the epistemic dimension of explanation, this kind of contextual sensitivity is a considerable advantage.

To summarise the available causal process theories, on the one hand we have the unfiltered view, which maintains objectivity and is friendly to those who think of explanation as an ontic phenomenon. The view struggles with relevance and illumination, but this is presumably of little concern to its proponents. On the other hand we have the psychologically filtered view, which allows us to zone in on the causal processes we actually find explanatory in a particular context, at the cost of abandoning objectivity. It is noteworthy that there is no view here which has the best of both worlds, such that it is an objective matter what explains what, yet the class of explanations perfectly coincides with our epistemic explanatory intuitions.

2.5.3 Counterfactual Theories

As I lack space in this thesis to consider the prospects of counterfactual theories of metaphysical explanation,⁴⁷ I will steer clear of a lengthy explication of these theories in our survey of views of scientific explanation. Nonetheless, it is another natural thought that the DN and causal theories fail to rule out explanatory irrelevancies such as Jones' unusual birth control pill consumption because they fail to respect the notion of 'making a difference'. There may be both a lawful deduction and a causal process running from the pill consumption to the lack of pregnancy, but we intuit these as non-explanatory and irrelevant because Jones—as a man—would not have fallen pregnant *whether or not he took the pills*. In other words, the pills made no difference to whether or not Jones fell pregnant. In response to the thought that *this* is the missing ingredient, theories of explanation have been developed that revolve around difference-making. Thus these theories are, roughly, counterfactual in nature. As a gloss, the thought is that if, had the putative explanans been different, the explanandum would also have been different, the explanation succeeds. If, however, changing the nature of the putative explanans makes no difference to the nature of the explanandum, there is no explanatory relation between the two. Distinct ways of precisifying this idea can be found in Lewis (1986b) and Woodward (2003).

To cut a long story short, these theories also admit of unfiltered and psychologically filtered variants. As there is counterfactual dependence in each putative case of scientific explanation, counterfactual theories do well with covering cases. It

⁴⁷ Such as Wilson's (forthcoming) and Baron's (ms) counterpossible theories of metaphysical explanation.

follows from this that the filtered counterfactual theory, while failing to be objective, will secure all of the epistemic desiderata. The unfiltered counterfactual theory will, due to the symmetries in the physical laws, allow symmetrical explanations.⁴⁸ Moreover, all contemporary states of affairs counterfactually depend on the big bang, and thus the theory allows for explanations that fail to increase our understanding. However, unlike the other unfiltered theories discussed here, the unfiltered counterfactual theory does well in ruling out irrelevant explanations. Of note is a repeat of the pattern whereby the unfiltered theory retains objectivity yet struggles with asymmetry and understanding, while the filtered theory achieves the epistemic desiderata by giving up on objectivity.

To summarise, the purpose of this section has been threefold. Firstly, we have identified a useful way to categorise theories of explanation in terms of the candidate relation that ‘backs’ explanation and an optional *filtration* of the instances of that relation. This filtration can be done psychologically, or otherwise. Secondly, we have considered some prototypes of various styles of theory of explanation, which will serve as templates for the theories of metaphysical explanation developed in the following chapters. Thirdly, we have uncovered a tension between these desiderata. In particular, it is notable that objectivity does not sit nicely with understanding, relevance and asymmetry. This mirrors the fact that if a view provides the objectivity desired by those who like to think of explanation as an ontic phenomenon, it is less likely to secure the more epistemic desiderata. The unification view comes the closest to satisfying the demands of both of these camps, but still fails on understanding. As such, there is no view that is objective while also identifying only relations we find explanatory in the epistemic sense.

2.6 The Plan Going Forward

In sum, the Sophisticated Grounding Theory and the Grounding-free Theory are competing attempts to explain our priority intuitions. Each will provide a story about why we have the intuitions that we do. Furthermore, each has the resources to vindicate these intuitions (should we seek vindication), and in a way consistent with a theory of metaphysical explanation (should we want it). Furthermore, these theories of metaphysical explanation can be evaluated in terms of the listed desiderata.

⁴⁸ Assuming that Elga (2000) is right that there is no Lewisian ‘asymmetry of overdetermination’.

In Chapter 3, I characterise grounding, and show how the existence of grounding relations can help account for the priority intuitions. The idea is a simple one: we have these intuitions—and what we intuit is true—because there are, *in re*, grounding relations which correspond with our intuitions, and which we are successfully tracking. However, this explanation is less theoretically virtuous than the explanation I offer in Chapter 4, and thus grounding relations are not indispensably required to do this explanatory work.

Nevertheless, it might be argued that the explanation offered by the Sophisticated Grounding Theory is the only game in town, for only it can *vindicate* the priority intuitions via one of a pair of ready-made theories of metaphysical explanation: the filtered and unfiltered grounding-based theories. I respond by noting that, while these theories do well according to the desiderata, there are strong reasons to dislike them. For, there is a substantial disanalogy between, on the one hand, filtered and unfiltered theories of scientific explanation, and on the other, filtered and unfiltered grounding-based theories of metaphysical explanation. In the former case, there is agreement on the nature of the candidate relation that our explanations track: what it's like and where it obtains. For instance, filtered and unfiltered causal process theorists agree about the causal structure of the world. They merely disagree on whether a good theory of explanation should psychologically filter the causal relations.

In stark contrast, the unfiltered and filtered grounding-based theories *disagree* about what grounding relations obtain, yet for the most part *agree* about what metaphysically explains what. The disagreement between these views, then, concerns whether there are instances of grounding that have no corresponding metaphysical explanation. I will argue that this is because the supposedly 'unfiltered' grounding-based theory has, in fact, been illicitly 'pre-filtered', for epistemic considerations have already been appealed to in establishing where in the world grounding relations are instantiated. On the other hand, I will argue that the filtered grounding-based theory is strictly worse than the theory I offer in Chapter 5, as it does equally well according to the desiderata but is less parsimonious.

Chapter 4 will provide a competing explanation of the priority intuitions, framed in terms of the overgeneralisation of a pair of evolved psychological mechanisms. I argue that this explanation is more virtuous than one in terms of grounding. If we seek explanation but not vindication, the story provided here is the whole story. In other words, by itself this explanation lends itself to *explaining away*, rather than vindicating our intuitions. It is parsimonious, and enjoys empirical support.

However, in Chapter 5, I show how this explanation can be framed as a way of vindicating our intuitions, and can serve as the basis for a psychologistic theory of metaphysical explanation. Such a theory has it that metaphysical explanations track the traditional modal relations, but are psychologically filtered according to what the majority of individuals in a community find explanatory. For those friendly to an epistemic understanding of explanation, this account will both vindicate the priority intuitions and provide a theory of metaphysical explanation that does well according to the epistemic desiderata.

Subsequently, Chapters 6 and 7 will develop metaphysical variants of the DN theories. How one weights the desiderata will determine which of the theories offered in Chapters 5, 6 and 7 is the most attractive. However, as noted in Chapter 1, these theories are more virtuous than their grounding-based competitors.

Chapter 3: The Grounding-Based Explanation

Recall the master argument of the thesis:

1. One ought (epistemically) to be ontologically committed to all and only those entities that are indispensable to the best explanation of our observations.
2. Grounding relations are not indispensable to the best explanation of our observations.
3. Therefore, we should not be ontologically committed to grounding relations.

In Chapter 2, I described the twin observations alluded to in premise (1) and (2) of the master argument: our intuitions about which modal correlations are accidental and non-accidental, and our priority intuitions. We are well placed, now, to introduce grounding relations, and show how positing these relations can account for these explananda. While I have made clear that my focus is on the priority intuitions, I will take a moment during this chapter to articulate the grounding-based explanation of the intuitions about modal correlations. It is only by comparing the grounding-based explanation⁴⁹ of the priority intuitions to my preferred explanation (offered in the following chapters) that I can demonstrate the truth of (2): that positing grounding is no part of the best explanation of these observations. As I will show, while the putative existence of grounding relations can form part of an explanation of these observations, this grounding-based explanation is less parsimonious and elegant than the explanation I offer in Chapter 4.

Furthermore, once grounding is posited to account for our priority intuitions, a pair of rival theories of metaphysical explanation naturally present themselves: the unfiltered grounding-based theory and the filtered grounding-based theory. The unfiltered theory has it that every grounding relation backs an explanation while the filtered theory admits *more* instances of grounding—unlike the filtered theories considered in §2.5—and then imposes psychologistic constraints on which instances of grounding back explanations. While these theories end up identifying roughly the same class of metaphysical explanations, they do slightly differently according to the desiderata introduced in §2.4, and lend themselves to slightly different characterisations of grounding. For, according to the unfiltered view, the features of grounding mimic those

⁴⁹ When I refer to *the* grounding-based explanation, I have in mind the explanation offered by the Sophisticated Grounding Theory, which, unlike the Basic Grounding Theory, is a genuine competitor to my Grounding-free Theory.

of metaphysical explanation, whereas if one filters the grounding relations to find the explanatory instances, one can have a more permissive stance regarding the features of grounding.

For those who are tempted by the unfiltered grounding-based theory, I will show that there are reasons to be suspicious of the fact that grounding is characterised in such a way as to allow a putatively objective theory of explanation that also has all the merits of an epistemic theory. I will argue that the view has been illicitly *pre-filtered*, for epistemic considerations play a substantial role in the characterisation of grounding itself. The filtered grounding-based theory fares no better, for it faces some awkward questions about the epistemology of grounding. Furthermore, this view suffers from an unfavourable comparison with the filtered modal relations theory I offer in Chapter 5, for it fulfils precisely the same desiderata, yet posits additional ontology in the form of grounding relations.

The plan for the chapter is as follows. §3.1 will characterise grounding, and narrow our focus to the ‘relation’ view (as opposed to the ‘sentential operator’ view). I will provide an overview of some debates about the features of grounding, emphasising that such debates tend to be framed in terms of what putative metaphysical explanations we, in fact, find to be explanatory. With this characterisation in hand, §3.2 will present the grounding-based explanation of the explananda, show how it vindicates our intuitions, and describe the two grounding-based theories of metaphysical explanation gestured at above. Section §3.3 will examine whether the grounding-based explanation is theoretically virtuous, and §3.4 and §3.5 will criticise the unfiltered grounding-based theory and the filtered grounding-based theory, respectively. §3.6 will sum up and point the way forward.

3.1 Grounding

In recent years, it has become very popular to frame metaphysical questions in terms of grounding: as questions of what *grounds* what. This has been accompanied by a slew of papers trying to get clear on the precise nature of grounding. Here I do my best to present an orthodox account of grounding. However, grounding is a moving target, with many competing conceptions on offer.

Recent discussion of the notion of grounding was initiated by seminal papers from Fine (2012), Schaffer (2009), and Rosen (2010), though—as we will see in §3.1.1—these authors are talking about quite different things, for Fine’s theory of grounding is more closely aligned with what we are here calling metaphysical explanation. While many aspects of grounding are controversial, I will use the term ‘grounding’ to refer to a primitive,⁵⁰ constitutive determination relation, typically thought to be asymmetric, irreflexive and transitive⁵¹ (See, e.g., Schaffer, 2009, 2010; Raven, 2012, 2015; Rosen, 2010; Audi, 2012a, 2012b; Cameron, 2008) and thus which gives rise to a partial ordering structure. Further, it is often deemed to be non-monotonic and well-founded (see Cameron, 2008 and Orilia, 2009 for discussion). More in-depth coverage of the debates about the contentious features will be delayed until §3.1.2, and how these match up to the desiderata for a theory of metaphysical explanation will be considered in §3.4 and §3.5.

Grounding is, putatively, the relation by which the fundamental existents ‘generate’ the derivative existents. As it is a primitive, unanalysable relation, grounding is typically introduced by way of examples. However, we already have our examples in hand, for the list of priority intuitions introduced in §2.1.2 can do double-duty as examples of grounding. Such is the tightness of the connection between grounding and the priority intuitions. Before saying more, it will be of use to clarify the target of our discussion as the ‘relation’—as opposed to the ‘sentential operator’—view of grounding.

3.1.1 Relation or Sentential Operator?

There are two putatively competing camps amongst those theorists interested in grounding: those who think of grounding as a sentential operator, and those who think of grounding in an ontologically robust way, as a relation that exists in the world, and that obtains between worldly entities. Here I will argue that there is, in fact, no competition between these views. Those in the former camp are in the business of providing a rigorous formalisation of the notion of metaphysical explanation introduced

⁵⁰ There are several projects that seek to provide a reductive account of grounding. For instance, Wilson (forthcoming) seeks to do so in terms of the truth of certain counterpossible claims, while Tallant (2015) seeks to do so in terms of some combination of Lowe’s (2010) relations of Rigid Existential Dependence and Identity Dependence. Grounding, so reduced, is not my target here.

⁵¹ Grounding is transitive just in case whenever A grounds B and B grounds C, A grounds C. I have already introduced the notions of irreflexivity and asymmetry.

in §2.4, whereas those in the latter camp are arguing for the existence of a primitive relation which such explanations track. Insofar as there is tension between these two projects, it is insubstantial tension about which notion should be called ‘grounding’. As the notion of grounding with which I take issue is the relation view, I will use ‘grounding’ to pick out that relation. The sentential operator view is quite compatible with a grounding-free explanation of the priority intuitions, and I will say little about it beyond this subsection.

To see how the dispute between these camps came about, note that claims about what grounds what, what is prior to what, and indeed, what metaphysically explains what, come in a variety of linguistic forms. While our illustrative examples of priority intuitions were uniformly expressed using ‘*because*’ to help make clear why some think they are intuitions about explanation, the more extensive list of these intuitions (in footnote 25) includes locutions of ‘grounds’, ‘because’, ‘in virtue of’ and others besides. While on the one hand we can surely translate between these locutions, on the other hand they have different syntactic implications. As Trogdon (2013a:5) points out, “On the syntactical surface level, the verb ‘ground’ is a relational predicate, ‘because’ is a sentential connective, and ‘in virtue of’ is a sentence-forming operator that requires a sentence as its first argument and a singular term as its second.”

One camp of grounding theorists takes the relational predicate interpretation to be the canonical form of grounding claims. These include Schaffer (2009), Cameron (2008) Audi (2012a), Rodriguez-Pereyra (2005), Raven (2012) and Trogdon (ms-a). The other camp views the sentential operator interpretation as canonical. These include Fine (2012), Correia (2005), Dasgupta (2014) and Litland (2013). Yet, this is not merely a disagreement about how best to formulate the sentences that express grounding claims. It is a disagreement about what we are talking about when we talk of grounding. Those who defend the relation view are ontologically committed to both the relation of grounding and the relata it relates. Those who defend the sentential operator view think of grounding as a non-truth-functional sentential connective, which takes arguments/sentences on either side. They need take no ontological stand.

Fine, for example, thinks that “the notion of ground should be expressed by means of a sentential operator, connecting the sentences that state the ground to the sentence that states what is grounded” (2012:12). The view is not, as stressed by Trogdon (2013a),

the view that grounding is a *relation between sentences*, but the explicit absence of a position on whether there is a grounding relation at all. A putative advantage of the view is its ontological neutrality. The thought is that one can accept, for example, that ‘{Socrates} exists *because* Socrates exists’ is true,⁵² without ontological commitment to philosophers, sets or the grounding relation. Of course, it remains open to a defender of the operator view to believe in philosophers and sets and some relation obtaining between the two. Indeed, the relation Fine (1994) calls ‘ontological dependence’ very much resembles the relation view of grounding. But, importantly, no such commitment is implied by a commitment to the truth of the grounding claim, and sentential operator theories of grounding are solely focussed on the elucidation of logical relationships between grounding claims.

Trogon (2013a) notes that there is a parallel dispute in the truthmaking literature. Rodriguez-Pereyra (2005) has argued that the correct form of truthmaking claims is that of a relation between the truth of a proposition and an entity, its truthmaker. Consider ‘<the rose is red> is true in virtue of the redness of the rose’. This expression of truthmaking implies that there is a thing which is the redness of the rose, which is being reified in order to stand in a relation of truthmaking to the proposition. On the other side, Hornsby (2005) argues that the correct form of truthmaking claims resembles ‘<the rose is red> is true because the rose is red’, where ‘because’ is a sentential operator and there is no implicit commitment to the redness of the rose existing, and standing in a worldly relation to the truth of the proposition. The former understanding seems more in the spirit of Armstrong’s (2004) truthmaker theory, where the existence of things—in his case states of affairs—making propositions true is the name of the game. However, those wary of buying into additional ontology may think that the operator understanding is sufficient to explain the truth of the proposition.⁵³

Returning to grounding, the relation and sentential operator theories connect to metaphysical explanation in very different ways. According to the relation view, grounding is a candidate determination relation that our metaphysical explanations might (and, if it exists, hopefully do) track, in the sense described in §2.5. I will evaluate the resultant theories in §3.4 and §3.5. In contrast, the sentential operator view resembles the notion of metaphysical explanation itself (§2.3). Thus, a natural view to have is that

⁵² Sentential operator theorists also express such claims using the phrase ‘*in virtue of*’ (see Fine, 2012).

⁵³ See Correia (2011) and Schnieder (2006).

grounding claims (in the sentential operator sense) are made true by the existence of grounding relations. With this in mind, the supposed tension between the views evaporates.

Rodriguez-Pereyra (2005) appeals to Ruben's (1990) ontic view of explanation according to which some worldly relation is required to underwrite metaphysical explanations in order to push those in the operator camp over into the relation camp. On the one hand, this resembles a mere semantic quibble: Rodriguez-Pereyra would prefer to use the term 'grounding' to talk about worldly relations. On the other hand, this argument brings us to where I believe the real action is. In other words, the debate should revolve around whether there are grounding relations, *in re*, that back our metaphysical explanations. The sentential operator view, then, is not the target of my criticism.

In sum, I think it is a mistake to conceive of the sentential operator and relation views as competing theories of the same thing, as this implies that if we go with one, we must scrap the other. I think that both the sentential operator and relation views are good theories of *something*. Thus I endorse Schaffer's (2016) proposal of *reconciliation*. That is, while the relation theorists are trying to give an account of the ontology that underpins metaphysical explanations, the operator theorists are attempting to formalise the logic of the metaphysical explanations themselves. Thus 'grounding', as expressed by the operator theorists, corresponds with 'metaphysical explanation' as I am using the term. This sits well with Dasgupta's claim that in his use, "the term, 'ground' is an *explanatory* notion: to say that X grounds Y just is to say that X explains Y, in a particular sense of 'explains'." (2014:3).

It should be noted that Schaffer himself rejects *reconciliation*, as he believes that there are no distinctively metaphysical explanations for the operator theorists to be theorising about—just explanations, backed by different relations (2016). While Schaffer's argument may hit home against the sentential operator theory (and against the notion of metaphysical explanation), it does little to impact the dialectic in this thesis. Perhaps the explanations I call metaphysical explanations are not separable from the remaining explanations (most notably, for instance, the scientific explanations). It is still an open question how to account for the priority intuitions. Furthermore, if the goal is to vindicate these intuitions in a way consistent with a general theory of explanation—as

opposed to a particular theory of metaphysical explanation—so much the better for my preferred views, which are continuous with (indeed, plausibly extensions of) theories used in the context of scientific explanation.

So, I see the efforts of Fine (2012) and Correia (2005) as attempts to provide a logical system which coherently regiments our intuitions about metaphysical explanations (though one might be sceptical that our priority intuitions exhibit some of the features these authors build into their theories: transitivity, for example. See §3.4). This regimentation is compatible with a grounding-free explanation of the priority intuitions, and indeed a grounding-free account of metaphysical explanation. That such a systematisation of these explanations can be developed bodes well for the intelligibility of the notion of metaphysical explanation. As argued by Raven (2012), a successful regimentation of a theory goes some way towards vindicating the theory itself.

Going forward, I will assume there is a well-defined class of metaphysical explanations. When it comes to divvying up the terminology, I will use ‘grounding’ to indicate the worldly relation that, according to some, backs these metaphysical explanations. With our focus now solely on *grounding relations*, we now move on to say a little more about these relations.

3.1.2 Further Features of Grounding

So, our focus is on grounding relations which (putatively) exist, are irreducible, and capture a certain kind of ontological dependence which gives rise to a hierarchical ontological structure—the “great chain of being” (Schaffer, 2009:376). Grounding relations always run ‘upwards’ on Wilsch’s (2015, 2016) axis of fundamentality. Thus, as we noted in §2.3, the terms ‘fundamental’ and ‘derivative’ can be cashed out in terms of grounding: fundamental existents are those which are not grounded, while derivative entities are those grounded in one or more fundamental entities (Schaffer, 2009:373-4). Furthermore, grounding is thought to admit of the partial/whole distinction, as well as the immediate/mediate distinction. Fine (2012; §2.3), in particular, has carefully distinguished whole, partial, mediate and immediate grounding.

While we have established that the target of our discussion is grounding understood as a relation, rather than a sentential operator, there is a further question

about *which* entities constitute its relata. Schaffer (2009, 2010) and Cameron (2008) defend a permissive stance such that diverse entities including objects, true propositions and facts can enter into grounding relations. Others restrict the relata of grounding to *facts*. For Audi (2012a), facts are like Armstrongian states of affairs. For Rosen (2010) and Trogdon (ms-b), facts are like true Russellian propositions.

Raven conceives of facts as “the truths of true representations, as opposed to the true representations themselves. Facts, so conceived, are states of the world, have structure, have constituents, and distinct facts can obtain in the same circumstances.” (2012:3). Though these authors describe their ‘facts’ in particular ways, the central idea is the same: facts are worldly entities, akin to Armstrongian states of affairs. They are structured pieces of ontology comprising objects, properties and relations.

We need not take a stand on this question, here. For, while they describe different metaphysical pictures, the permissive and fact-only views seem to offer the very same potential explanation and vindication of our priority intuitions. Moreover, it is simple enough to translate between object-talk (or truth-talk) and fact-talk. For instance, Schaffer and Cameron might allow that ‘Socrates grounds {Socrates}’, but those who restrict the relata to facts can capture this very same dependence with ‘[Socrates exists] grounds [{Socrates} exists]’.⁵⁴ So, I remain agnostic between the permissive and fact-only views.

Another in-house dispute about which I shall remain agnostic concerns the *adicity* of the grounding relation. Some have argued that grounding is a one-one relation, others that it is many-one, and still others that it is many-many. Dasgupta (2014), for example, argues that the relata of grounding relations are irreducibly plural. Fine (2012) argues that grounding is variably polyadic, and, as I see it, this permissiveness is justified in the face of such a variety of grounding claims. As with the debates about relata, I assume that nothing substantive hangs on these differences. Given that the very same claims of the form ‘*x because y*’ can be explained and vindicated given these various views on adicity, for the purposes of this thesis we can remain neutral.

Moving on, let’s explore some arguments about contentious features of grounding, with a particular focus on the way such arguments tend to proceed. For—and it is

⁵⁴ Notably, Audi (2012a) denies this possibility. Audi-style facts cannot have existence as a component because they cannot include quantifiers, and he denies that existence is a property. However, for those with a more liberal account of facts, such translations seem reasonable.

important to flag this early, to allow the reader to see this for herself—the reasons given for imbuing a theory of grounding with this feature or that are, for the most part, direct appeals to the priority intuitions, and the nature of explanation. Such is the intimacy of the connection between grounding and intuitions of priority. We will return to this thought when we evaluate the grounding-based theories of metaphysical explanation. For example, we have noted that grounding is thought to have the formal features of asymmetry, irreflexivity and transitivity (Schaffer, 2009; Audi, 2012a; Rosen, 2010; Cameron, 2008; Trogdon, 2013a; Raven, 2012) and thus has a similar structure to the proper parthood relation, generating a partial ordering over its relata.

First, consider irreflexivity. If grounding is irreflexive, nothing can stand in the grounding relation to itself. Audi (2012a) and Raven (2012) appeal to the nature of explanation to justify the irreflexivity of grounding. They contend that, due to strong intuitions that nothing can explain itself, nothing can ground itself. As grounding is a primitive posit of that which is being tracked by our metaphysical explanations it seems bound to have this feature.

There are some interesting putative counterexamples to irreflexivity. For instance, Paseau (2010) notes that while it is orthodoxy amongst defenders of grounding that the existence of a singleton set is grounded by the existence of its member, some people think that the set is *identical* to its member. If one holds both these views, this constitutes a reflexive instance of grounding. There are a few responses available here. The first, and most popular, is to deny that sets are identical to their members. If one holds on to the identity of the set and its member, one could either accept that there is reflexive grounding (and thereby reflexive explanation; though see the discussion of *quasi-irreflexivity* in §3.5) or deny that there is an explanatory relationship and a grounding relation that obtains between them (indeed, ‘them’ is a misleading term in this context, as the view is that there is only one thing, that may or may not stand in the grounding relation to itself). Notably, this latter route detracts from the ability of grounding to explain the priority intuition in question.

Another putative counterexample to irreflexivity comes from Fine (2010). He points out that it is a fact that everything exists. Furthermore, the fact that everything exists is grounded by every fact that obtains. Thus the fact that everything exists is a partial ground for itself. Another incarnation of this puzzle is that the truth of the

proposition that <every proposition is true or not true> appears to be partially grounded by its own truth. This is a tricky case for defenders of grounding, for, on the one hand, there is a generic intuition that there are no reflexive explanations, yet on the other hand, the fact that everything exists (better, perhaps: that everything that exists, exists), does seem to help to explain its own existence. Ultimately, the overwhelming majority of grounding theorists rule in favour of irreflexivity, because, they think, it would be odd to have some fact explain its own existence.

We now move on to look at asymmetry, the feature which denies that a pair of entities can mutually ground one another (Schaffer, 2009; Cameron, 2008; Trogdon, 2013a). In keeping with the theme of explanatory justification of features of grounding, Audi (2012a) and Raven (2012, 2015) point out that the prospect of truths which mutually explain one another does not sit well with an intuitive view of explanation, and thus should not be a characteristic of grounding. Just as circular explanations are unacceptable, so are cycles of ground. Furthermore, recall that within the grounding structure every entity is either fundamental or grounded in the fundamental. Symmetrical grounding would allow a pair of entities to fall into neither of these categories, as they are both grounded (and thus not fundamental) yet not grounded in the fundamental (as they are grounded in one another).

Similarly, grounding is thought to be transitive (Schaffer, 2009; Cameron, 2008; Audi, 2012a; Raven, 2012).⁵⁵ While I shall discuss transitivity in §3.4 and §3.5, it is of note in the present context that without transitivity, ground would not generate a partial order, for it would be possible for A to ground/explain B, and B to ground/explain C, and C to ground/explain A *without* violating irreflexivity or asymmetry. Fine (2010) argues for transitivity on the basis that “completely satisfactory” explanations do not involve such cycles.

Moving beyond partial ordering, grounding is also thought to be constrained by relevance. This feature, too, is motivated on explanatory grounds. In the grounding literature, the debate about relevance is framed in terms of *non-monotonicity*. Many advocates of grounding claim that the relation is non-monotonic, in the sense that if A grounds B, it does not follow that the plurality of A and some arbitrary C will also ground B⁵⁶ (Trogdon, 2013a; Dasgupta, 2014; Audi, 2012a; Raven, 2012; Rosen, 2010).

⁵⁵ Schaffer (2010) has argued against the transitivity of *partial* grounding.

⁵⁶ This claim can be appropriately translated according to one’s preferred view on the adicity of grounding. 66

The C in question can only be added to A if it is *relevant* to the obtaining of B. Thus grounds must be relevant to what they ground, where the kind of relevance in question is supposed to be something upon which we have an intuitive grasp.

This idea can be further clarified by noting that, in contrast, the modal relation of necessitation is monotonic, and not constrained by relevance. To see this, consider that if A necessitates B then it follows that (A & C) together necessitate B, for any C. For example, if [there is a red cube] necessitates [there is a cube] then [there is a red cube and Saturn is a planet] also necessitates [there is a cube], despite the irrelevance of Saturn to cubes.

Trogon (ms-a), Dasgupta (2014) and Raven (2012) appeal to the link between grounding and explanation to defend this kind relevance constraint upon the instances of grounding. In Dasgupta's phrase, such constraints are "one of the central features used to distinguish ground from metaphysical necessitation and logical consequence." (2014:5).⁵⁷ Thus, in rough terms, one cannot 'tack on extra stuff' to the more fundamental side of a grounding relation while preserving the truth of the grounding claim.

To be clear, non-monotonicity is *not* the claim that if A grounds B, then for *no* C will the plurality (A & C) ground B. This is especially clear in the case of partial grounding. [My left thumb exists] partially grounds [my left hand exists]. While the plurality [my left thumb exists and my right thumb exists] *does not* ground [my left hand exists]—due to the irrelevance of my right thumb to my left hand's existence—the plurality [my left thumb exists and left index finger exists] *does* partially ground [my left hand exists]. More generally, when A and C are both partial grounds of B, (A & C) will ground B.

Similar considerations apply to whole grounding. Perhaps [a red soccer ball exists] (wholly) grounds [a red thing exists]. The plurality [a red soccer ball exists and a red cricket ball exists] will still wholly ground [a red thing exists], as the red cricket ball is relevant to the existence of red things. In contrast, [a red soccer ball exists and a blue soccer ball exists] does not ground [a red thing exists], for the blue soccer ball is irrelevant.

⁵⁷ Constraints of relevance will only differentiate grounding from classical consequence; it is less obvious that they will differentiate grounding from relevant consequence.

Consider, now, the modal implications of grounding. The orthodox position is necessitarianism, the view that if A grounds B then A necessitates B. Like most features of grounding, defence of the necessitarian intuition tends to be phrased in explanatory terms. For instance, Trogdon (2013b:footnote 3) argues that it is in the nature of explanation to show how the explanandum “couldn’t have been otherwise” given the truth of its explanans, and deRossett asserts that “If it is possible for the *explanandum* to obtain while the *explanans* does not, then the *explanandum* does not make the *explanans* obtain, and so the explanation fails.” (2013:15). Others simply insist that grounding exhibits this feature. For instance, Fine (2012:1) states that “a necessary connection must hold between the relata if the [grounding] relation is to obtain,” and Correia (2005:61) writes that “Necessarily, if the fact that A is grounded in some given facts, then it is impossible that the latter facts all exist but fail to ground the fact that A.” Notably, even contingentists, (see Leuenberger, 2014; Skiles, 2015 and Chudnoff, ms) who demur from endorsing necessitarianism, argue for their preferred view in explanatory terms. For, they are naturally read as thinking there are some cases of metaphysical explanation—and therefore grounding—where the explanans fails to necessitate the explanandum.

Moving on from necessitarianism, we now consider the question of whether chains of grounding are *well-founded*, an interesting putative feature of grounding that is once again defended, in part, on the basis of explanatory considerations. This term has been used in two different senses in the grounding literature, so it is useful to specify that I will here follow the majority in holding that if a grounding chain is well-founded, then it eventually reaches an ungrounded entity. It ‘hits the ground’ such that every link in the chain is, in the end, grounded in the fundamental, ungrounded entities. Cameron (2008) uses the term to refer to the weaker constraint that for any *subset* of a grounding chain, there is an element which is the *most fundamental of the subset*. This constraint rules out cycles of ground, but is compatible with infinite chains of grounding which fail to ‘bottom out’, viewed as a whole. I take it that the well-foundedness of grounding in this *latter* sense is uncontroversial. Thus, ‘well-foundedness’ in this thesis refers to the stronger claim.

If grounding is well-founded, there are both fundamental entities and fundamental explanantia. So, why think that grounding is well-founded? Fine (2010) has argued for the well-foundedness of grounding on explanatory grounds. He insists that there is “a plausible demand on ground or explanation that we are unable to evade. For given a

truth that stands in need of explanation, one naturally supposes that it should have a ‘completely satisfactory’ explanation, one that does not involve cycles and terminates in truths that do not stand in need of explanation.”(p9). It is interesting to note that the parallel of well-foundedness in the context of causation and causal explanation would be the view that there must be some first cause, itself uncaused. Yet, this is not obviously an intuitive restriction on causal explanation. At the least, there remain unresolved arguments in this area.⁵⁸

With these features—and, more importantly, the way in which these features are debated—identified, I will now briefly consider some putative motivations for the recent surge of interest in grounding. For instance, defenders of grounding are quick to claim that their notion of ontological structure is not a new-fangled posit resulting from the eccentricities of 21st century metaphysicians.⁵⁹ The notion finds precursors in the Aristotelian ontological picture. While this is not exactly a *motivation* to posit grounding, the idea is that we, like Aristotle, need this kind of notion to say the things we want to say. In other words, something like grounding has been implicit in philosophical theories for a long time, and grounding is merely the explicit posit of that thing. To this, I say that it is *intuitions of priority*, and *metaphysical explanations* that have been implicit in our theorising, and that if the arguments of this thesis are successful, the fact that we have these intuitions and trade in these explanations fails to motivate positing grounding.

Another proposed motivation to posit grounding is the attractiveness of the hierarchical ontology it provides. Such an ontological picture is argued to be strictly superior to a flat, unstructured ‘Quinean’ ontology (the Quinean picture is akin to the ‘flatlander’ picture I mentioned in Chapter 1). This is because the ‘list of existing entities’ view preferred by the Quinean can be ‘read off’ the structured picture, while the inverse is not the case (Schaffer, 2009). In other words, while it is simple enough to *flatten* a hierarchical ontology and derive the non-hierarchical alternative, it is not so simple to build the hierarchical structure without the grounding relations in place. Here, the thought seems to be that the hierarchical view is better because its associated picture contains more information.

⁵⁸ For instance, Craig and Sinclair (2009) have argued that if time had no beginning then we would have to have traversed an infinite temporal distance to get where we are, an argument which has attracted multiple responses (see, *inter alia*, Puryear, 2014).

⁵⁹ Though it is notable that in relatively recent work, Rosen claims that the idioms of grounding are “not part of anyone’s official vocabulary” (2010:109). As Kovacs (forthcoming) notes, this is very much no longer the case.

However, organising ontology into a hierarchy of this kind is only attractive insofar as we have good reasons to think that the world truly is hierarchically structured. Furthermore, as I see it, the only reason to think the world has this structure is that we observe our own priority intuitions.⁶⁰ Thus this kind of argument ultimately reduces to the argument that grounding is required to explain the priority intuitions. For, only if grounding is the route via which we should explain our priority intuitions is the ontological hierarchy desirable. The fact that the flat picture can be easily derived from the hierarchical one, but not *vice versa*, is beside the point. The connections described by the hierarchy may well be interesting ones, but such connections can be elucidated by grounding-free theories of metaphysical explanation, such as those considered in this thesis (though they may well be treated with less ontological seriousness).

Ultimately, the only substantial argument in favour of positing grounding is its putative indispensability in accounting for the priority intuitions and generating a theory of metaphysical explanation (and in accounting for the distinction between accidental and non-accidental modal correlations). This is often presented as a simple argument from the fact that we trade in metaphysical explanations to the need for grounding to back such explanations. However, as I will demonstrate (in Chapters 5, 6 and 7), there are alternative methods of accounting for metaphysical explanations. There are a number of putative relations we could be tracking. Thus the simple argument from the fact that we posit metaphysical explanations, to the belief that grounding exists, is weakened. Those who defend grounding need to show that it is *the best* way to account for metaphysical explanations, not merely that it can do so. It is to the grounding-based theories of metaphysical explanation that we turn in §3.2.3.

3.2 The Grounding-Based Explanation

In the context of the Sophisticated Grounding Theory, grounding relations can help explain the twin explananda, and can do so in a way that vindicates our intuitions. I won't argue against this. Rather, I will argue, later, that there are better ways to do this

⁶⁰ It might be objected that there are reasons to believe that there is a hierarchical ontological structure that do not stem from the priority intuitions. Perhaps, for instance, the phenomenon of emergence in science motivates the hierarchical view. Whether this is so will depend on how one thinks of emergence. If emergence is an epistemic phenomenon (see Chalmers, 2008), then, while it might motivate a hierarchy of some kind, it clearly does not motivate an *ontological* hierarchy. If emergence is ontological, such that the emergent phenomena are new *entities*, then perhaps emergence provides a motivation that is distinct from the priority intuitions.

work. Furthermore, the vindication allowed by positing grounding lends itself naturally to one of two grounding-based theories of metaphysical explanation. Once again, I argue that there are better ways to do this work. Before our more substantial discussion of the priority intuitions and the theories of explanation, I will briefly show how positing grounding relations can help explain (and vindicate) our intuitions about accidental and non-accidental modal correlations.

3.2.1 Explaining and Vindicating our Intuitions about Modal Correlations

Recall that the first explanandum concerns our intuitive categorisation of modal correlations into those that are accidental and those that are non-accidental. For example, we intuit that it is non-accidental that the existence of 2 necessitates the existence of {2}, while we intuit that the necessitation of {2} by the existence of my head is accidental. That we are so strongly inclined to make this classification is in need of an explanation.

If grounding relations exist, the explanation of these intuitions is simple. The reason that we intuit that the modal correlation between 2 and {2} is no accident, while the modal correlation between my head and {2} is an accident, is that the existence of 2 *grounds* the existence of {2}, whereas there is no grounding relation between the existence of my head and the existence of {2}, and we are somehow tracking whether or not these grounding relations obtain. More generally, we intuit that some correlations are non-accidental because there is something more than correlation at play, and we are, via some psychological mechanisms, able to detect this. There is, in addition to correlation, a relation of ground such that one thing *generates* another. Indeed, it is in virtue of this *generation* that there is such a correlation. Thus the distinction here mimics the distinction between accidental diachronic correlations and those diachronic correlations that are underpinned by a causal relation.

Furthermore, this kind of explanation also serves to vindicate our intuitions. For, if we are tracking the difference between those correlations where there are grounding relations and those where there are not, there is an ontological basis for the distinction we are making, and our intuitions are rightly said to be vindicated. In other words, if there exist grounding relations, and these relations obtain only in those cases of modal correlation which we intuit to be non-accidental, the propositions expressing these

intuitions will be true. In this way, the explanation also provides vindication.

As has been made clear, discussion of this explanandum is not my aim in this thesis, so I shall put it to the side.

3.2.2 Explaining and Vindicating the Priority Intuitions

The second explanandum, and the one with which I shall be concerned, is the presence of our priority intuitions. The explanatory story afforded by appealing to grounding relations is that the world is imbued with a hierarchical structure, limned by the obtaining of grounding relations. Such relations exist *in re*. In some way or another, we are able to track the grounding structure of the world and thereby able to discern what grounds what. This tracking manifests itself in the form of the priority intuitions. Thus, each intuition that ‘x *because* y’ is, assuming we intuit the truth, due to the obtaining of a grounding relation. Specifically, we have this intuition because y grounds x.⁶¹ As such, this intuition is a *veridical* one that is not only explained, but also vindicated by, the corresponding grounding relation. In sum, the grounding-based story has it that ‘x *because* y’ is true only if y grounds x.⁶² Moreover, our intuitions that ‘x *because* y’ tend to coincide with ‘x *because* y’ being true, because we are good grounding trackers. Thus our intuitions are vindicated.

In Chapter 1, I took defenders of the Basic Grounding Theory to task for a lack of an epistemological story about how we come to successfully track grounding relations. For, merely supposing that there is grounding in the world does little to explain how we might come to have any intuitions about what grounds what. In this way, there is an interesting parallel with the debunking arguments about moral realism (§1.2). However, this criticism only hits home for the Basic Grounding Theory. I suspect that something resembling the story I provide in Chapter 4, spun appropriately, could do at least some of this work on behalf of those who endorse the Sophisticated Grounding Theory. Nevertheless, for those who endorse the filtered grounding-based theory of metaphysical explanation, this epistemological problem returns with renewed force, as we shall see in §3.5. My primary argument, however, is that we lack a reason to posit grounding

⁶¹ Or, perhaps, if the intuition is about partial explanation, y is part of a plurality which grounds x. Recall that we are assuming, unless otherwise specified, that we are dealing with whole explanations.

⁶² I express this in conditional rather than biconditional form so as not to beg the question against so-called ‘separatists’ (§3.2.3)

relations, not that, once posited, we should suspect that we lack the means to correctly track them. I consider the explanation offered by the Sophisticated Grounding Theory in the light of the theoretical virtues in §3.3. Now, I shall show how this explanation of the priority intuitions naturally leads to one of a pair of grounding-based theories of metaphysical explanation.

3.2.3 Grounding-Based Theories of Metaphysical Explanation

As well as explaining and vindicating our priority intuitions, the Sophisticated Grounding Theory opens up the possibility of two theories of metaphysical explanation. These can be thought of as the metaphysical analogues of the unfiltered and filtered causal process theories (though, both here and in §3.4.1, I note significant disanalogies).⁶³ These theories have it that the candidate relation that backs metaphysical explanations is grounding. Moreover, as with the other candidate relations, we must discern whether every instance of grounding has an associated metaphysical explanation, or whether there should be some filtration of the class of grounding relations in order to zone in on the explanatory subclass.

Unusually, in the case of the grounding-based theories, there is disagreement between the filtered and unfiltered views regarding which instances of grounding there are in the world. Recall that causal process theorists agree about which instances of causation there are in the world but disagree about whether to filter this class of relations. That is, filtered and unfiltered causal process theorists disagree about whether every instance of the causal relation is explanatory or not. In contrast, filtered and unfiltered grounding-based theorists generally agree about what metaphysical explanations there are—it's just that the unfiltered theorists think that these are perfectly aligned with the instances of grounding relations that obtain, while the filtered theorists think that there is *more* grounding in the world, and thus have a need to filter this larger class of instances. As will become apparent, my contention is that the unfiltered theorists are, in fact, doing the same psychologistic filtering that filtered theorists are doing, but at an earlier stage in the theorising. More specifically, the filtering is being done in the characterisation of grounding as a candidate relation. The characterisation proffered is more 'strict', thus ruling out as instances of grounding, some of the instances

⁶³ It has been argued by Wilson (forthcoming, ms) that grounding is simply a type of causation, and by Schaffer (2016) that grounding should be given the same formal treatment as causation.

countenanced by the filtered grounding-based view. Thus, while the view is sold as being unfiltered, it is better thought of as an illicitly ‘pre-filtered’ view (§3.4.1).

The idea that grounding is intimately tied up with metaphysical explanation is not something that I am foisting onto my opponent. Indeed, as we saw in §3.1.3, debates about the formal features of grounding, and about what grounds what, are often framed in terms of what could plausibly metaphysically explain what. Furthermore, several authors (e.g., Audi, 2012a; Raven, 2015; Rodriguez-Pereyra, 2005) have explicitly argued that we *need* a primitive grounding relation to provide ontological underpinnings for metaphysical explanations. To use the terminology introduced in §2.5, we need grounding as the candidate relation we track with our metaphysical explanations.

For example, Rodriguez-Pereyra claims that “Explanation is not and does not account for grounding – on the contrary, grounding is what makes possible and ‘grounds’ explanation” (2005:28). Trogdon (2013a) adds that “we treat grounding in this context as carrying explanation in the sense that, if one thing is grounded in another, the latter explains the former”, while Schaffer (2010:31) asserts that “Grounding connects the more fundamental basis to the less fundamental result, and thereby backs a certain form of explanation. Thus the right sort of physical system can support a biological organism such as a cat, and one sort of answer to the question of *why there is a cat afoot* involves the underlying physical system”.

However, as noted, there are two ways to think of the link between grounding and metaphysical explanation; two grounding-based theories of metaphysical explanation. The *unfiltered grounding-based theory* of metaphysical explanation has it that ‘x *because* y’ is true just in case y grounds x. As such, what metaphysically explains what is simply a matter of where the grounding relations obtain. This is the more natural view, particularly for those who think of explanation as an ontic phenomenon, and thereby identify the question of what metaphysically explains what with the question of what grounds what. Fine (2012), appropriately translated into the terminology we are using here, can be understood as endorsing this view. Others include Dasgupta (2014), Litland (2015), and Raven (2012, 2015). For obvious reasons, Raven (2015) dubs this position *unionism*.

As noted, such a view is roughly analogous to the unfiltered causal process theory (§2.5.2). Just as some causal process theorists emphasise the ontic aspect of explanation,

and think that what scientifically explains what is purely a matter of which causal relations obtain, these theorists contend that what metaphysically explains what is purely a matter of which grounding relations obtain. Recall, however, that in holding on to the objectivity of explanation, the unfiltered causal theorist opened herself up to criticism about some epistemically counterintuitive cases. Specifically, she had to give up on the more epistemic desiderata of relevance and understanding. In contrast, the unfiltered grounding-based theorist contends that, while grounding is a mind-independent feature of the world, grounding relations only obtain in cases where these desiderata are fulfilled (with the arguable exception of understanding. See §3.4).

In this way, the unfiltered grounding-based theorist thinks she can have her cake and eat it too. In response to this, I argue below that such theorists place themselves in an odd position whereby they putatively eschew the epistemic elements of explanation, yet appeal to the priority intuitions to help establish the features of grounding—intuitions which, surely, have a psychological basis. In other words, I take the unfiltered grounding-based theorist to task for appealing to what *we* (contingently) find explanatory in the way she establishes where the putatively mind-independent grounding relations exist, *in re*.

On the other hand, there is the (psychologically) filtered grounding-based theory that Raven (2015) dubs, once again for obvious reasons, *separatism*. This view is roughly analogous to a psychologically filtered causal process theory. Thus, while grounding is the candidate relation that we track with our metaphysical explanations, not all instances of grounding have an associated explanation. Only those instances of grounding that have the right epistemic features underlie metaphysical explanations. Audi, for instance, leaves it open that something beyond grounding *might* be required: “It is only a necessary condition of an explanation’s holding between two facts that a relation of determination hold between them. More might be required to fill out a sufficient condition. There may, for all I have said, be pragmatic or epistemic requirements on what counts as an explanation.” (2012a:21).

Likewise, Rodriguez-Pereyra contends that “it does not even follow that every relation of grounding must be susceptible of being tracked or represented by an epistemic explanation.” (2015:20). And Trogon remarks that “if one fact is grounded in another, *there are conditions in which* citing the latter suffices to explain the former” (2013a,

my italics). Arguably, Schaffer's (2016) proposal to model grounding with directed acyclic graphs (as is currently in vogue in the literature on causation; see *inter alia* Woodward, 2003) is a form of separatism. For, on one reading of that view, while the total set of graphs will give the totality of the grounding facts, a particular subset of those graphs will serve to identify the grounding relations that count as explanatory from a given perspective.

With the grounding-based explanation—and its associated theories of metaphysical explanation—now in hand, I am now well placed to present my criticisms.

3.3 Evaluating the Grounding-Based Explanation

Here, I will evaluate the grounding-based explanation in terms of the theoretical virtues. As noted in §1.3, these virtues can serve as a tie-breaker between the empirically equivalent Sophisticated Grounding Theory and the Grounding-free Theory. Of course, the following discussion of theoretical virtues underdetermines whether the Sophisticated Grounding Theory is a good one. For, in order to properly evaluate the indispensability (or otherwise) of grounding, we must *compare* the theory to empirically equivalent theories that lack grounding relations. As we have not yet articulated the competing theory, we are not yet well placed to do this.

Still, there is value in seeing how the Sophisticated Grounding Theory measures up. Consider first the virtue of elegance. It is manifestly clear in the above discussion that there is substantial ideological complexity to be found in the principles that govern grounding. This tells against the Sophisticated Grounding Theory, unless the Grounding-free Theory is saddled with its own unique ideological complexity. However, I contend (in Chapter 4) that the defender of grounding ought to believe in the very mechanisms to which the grounding-free explanation appeals, in order to account for our intuitions about *causal* dependence. Thus, when we come to compare the Sophisticated Grounding Theory with its competitor, I will argue that the Sophisticated Grounding Theory is less elegant, for it is also committed to the complexities of its rival.

Furthermore, the Sophisticated Grounding Theory is not parsimonious. For, it is contrary to parsimony to call upon primitive entities to do this explanatory work. Better: it is contrary to parsimony to call upon primitive entities to do this explanatory work

when the work can be done without them. Once again, I am pre-empting the comparative evaluation which will take place after I develop the competing theory.

Finally, while the Basic Grounding Theory fails to fit in with a unified story about why, more generally, we make the observations that we do, the Sophisticated Grounding Theory, by piggybacking on the psychological story I offer in Chapter 4, is just as unified as my alternative explanation. For, both explanations call upon evolutionary pressures and psychological mechanisms which serve to explain many phenomena of interest.

Now, the defender of grounding might insist that, despite lacking theoretical virtues, the grounding-based theories *uniquely* vindicate the priority intuitions in a way that allows us to develop a theory of metaphysical explanation that has the features we want. Thus, insofar as we are in need of such a theory, we are in need of grounding: it is the only game in town. The following two sections are dedicated to showing that not all is well with the grounding-based theories of metaphysical explanation. Chapters 5, 6 and 7 are dedicated to the development of grounding-free theories.

3.4 Evaluating the Unfiltered Grounding-Based Theory

We will now proceed to evaluate the unfiltered grounding-based theory against the desiderata for a theory of explanation (as identified in §2.4). This will serve to play several roles. First, it will emerge that this theory does very well according to the desiderata. It gives us (arguably) everything we might want from a theory of explanation, from the ontic desideratum of objectivity, right through to the epistemic desiderata. However, I invite the reader to recall the *kinds* of arguments provided for or against thinking of grounding as having a certain feature. Of note is that such arguments often appeal to explanatory intuitions. So, for example, it is argued that *grounding* must be asymmetric because we don't find circular *explanations* to be illuminating. Yet what grounds what is supposed to be a mind-independent matter of ontology, not sensitive to anthropocentric preferences. In light of this, I ultimately argue that we should be suspicious of a relation that is, supposedly, a feature of the world, and yet has its instances delimited by what is explanatory *for us*. As I have said, far from being *unfiltered*, the view is, in fact, *illicitly pre-filtered*.

So, let us begin our evaluation. Recall that the view currently under consideration has it that metaphysical explanations reliably track grounding relations. Wherever there is grounding, there is metaphysical explanation, and wherever there is metaphysical explanation there is grounding. So, for instance, an early time-slice of Schaffer (ms of ‘Grounding as the Primitive Concept of Metaphysical Structure’, quoted in Audi 2012a⁶⁴) contends that “this notion of metaphysical explanation is just grounding by another name. It is not a distinct concept. What is revealed is rather the integration of the grounding relation within the family of explanatory concepts.” Likewise, Raven (2015:325) claims that “ground is often taken on its own terms as providing a *bona fide*, distinctively metaphysical kind of explanation”.

To begin, we should note that the grounding relation⁶⁵ does a splendid job of covering the intuitive cases of metaphysical explanation, and thus does well with the desideratum of covering cases. This should come as no surprise. Grounding is a primitive relation, posited as present in every case in which we have priority intuitions. Moreover, as grounding is an ontological posit that exists *in re*, the view does well according to objectivity.

Furthermore, the unfiltered grounding-based theory does well according to irreflexivity, asymmetry and relevance. What is significant, with regards to these desiderata, is *why* the unfiltered grounding-based theory succeeds in fulfilling them. For, as we saw in §3.1.3, defenders of grounding are quite happy to appeal to the nature of explanation, and to explanatory intuitions, in the characterisation of grounding itself. So, for example, Raven (2015:322) contends that “the explanatory and metaphysical aspects of ground can provide guidance on what its relata are as well as the explanatory logic and metaphysical characteristics it has”.

As we saw, grounding is deemed to be irreflexive and asymmetric because nothing explains itself, and cycles of explanation are undesirable. No surprise, then, that the unfiltered grounding-based theory does well against these epistemic desiderata. The concern, here, to which we will return, is captured nicely by Trogon (ms-a) in his defence of the filtered grounding-based theory (§3.5). He notes that “While we may think that there are asymmetric structure-imposing relations, we don’t think that explanation is

⁶⁴ To my knowledge, this passage was never published in any of Schaffer’s work.

⁶⁵ As the view under evaluation supposes that there is a one-to-one correspondence between grounding relations and metaphysical explanations, I shall, for ease of exposition, frame the below as a discussion of whether grounding itself fulfils the desiderata.

asymmetric because we think that these relations are asymmetric. (If anything it goes the other way around—I return to this idea shortly.)” In short, there is something suspicious about using explanatory intuitions to decide the features of the *in re* relations our explanations reliably track.

Trogon is also concerned about this kind of defence of relevance constraints on grounding. He asks us to “consider the intuition that something is an explanation only if any part of that explanation is *relevant* to what it explains. The notion of relevance is itself epistemic, and we don’t think that something is a portion of the world’s structure only if it has a particular epistemic property.” I think that Trogon is right that we should be concerned about the notion of relevance at play in the discussion of the non-monotonicity of grounding. In particular, given that what grounds what is supposed to be purely a matter of what the world is like, the relevance in question cannot be relativised to the psychological orientation of any individual, or group of individuals. It must be an objective fact of the matter which facts are relevant to others.

Admittedly, it is not *only* explanatory considerations that are appealed to in the characterisation of grounding. There is also the notion of *structure*. Grounding must be irreflexive, asymmetric and transitive because it limns a hierarchical partial ordering structure over that which exists. Yet, as I noted in §3.1.2, the motivation for thinking that the world has this kind of structure stems directly from the priority intuitions. Thus it is not clear that this is really an independent motivation.

Of note is that the putative transitivity of grounding, while required to generate the aforementioned structure, is arguably in tension with the desideratum that metaphysical explanations ought to increase our understanding. For, consider that there might be a chain of grounding relations running from a particular distribution of microphysical particles to the fact that your shares in Grounding, inc. are worth \$1. It follows from this, by transitivity, that the distribution of particles grounds—and, according to the unfiltered view, *explains*—the value of your shares. For another example, consider that a collection of particles arranged a certain way might ground that a composite lump exists, which in turn grounds that a statue exists, which in turn grounds that the singleton set {Statue} exists, which is the grounds of the truth of the proposition <{Statue} exists>. Thus, by transitivity, the existence and arrangement of the particles grounds the truth of the proposition <{Statue} exists>.

Cases such as these bring into question whether the unfiltered grounding-based theory can, in fact, satisfy the understanding desideratum. For myself, I am unsure whether there is an increase in understanding, here: if I squint, I can see how the respective higher-level facts might be illuminated by the obtaining of the lower-level facts. Thus, nothing in what follows will hang on this criticism, as it's not clear whether it really hits its mark. However, those who think these cases don't meet understanding have further reason to dislike the unfiltered grounding-based theory, since there is one desideratum it fails to meet. Moreover, concerns about these cases might, along with the concerns I raise below, motivate defenders of grounding to prefer the filtered grounding-based theory (§3.5).

Now, however, I want draw out exactly what is suspicious about using explanatory intuitions to decide the features of the *in re* relations our explanations track. For, ultimately, I think that the way in which grounding has been tailor-made to deliver such a strong theory of metaphysical explanation is not a virtue, but a vice.

3.4.1 A Damning Analogy with Causal Process Theories

In this section I present explicitly the objection to the unfiltered grounding-based theory of explanation hinted at in the previous section. While there is no doubt that the unfiltered grounding-based theory of explanation is a powerful one that does well according to the desiderata, as I have repeatedly hinted, we should be suspicious of that power. For, discussion of the features of grounding is aptly summarised by the following quote from Raven (2015:327).

Ground's *explanatory* aspect seems to impose on it a distinctive logic, including: (i) *irreflexivity*: just as nothing explains itself, so too nothing grounds itself; (ii) *asymmetry*: just as cyclical explanations are prohibited, so too are cycles of ground; (iii) *transitivity (cut)*: just as explanations chain, so too ground chains; (iv) *well-foundedness*: if explanations must begin, then so too any grounded fact must ultimately be grounded in facts which themselves are ungrounded; and (v) *non-*

monotonicity: just as explanation needn't survive arbitrary additional premises, so too ground needn't survive arbitrary additional grounds.⁶⁶

Yet, as Raven himself notes (326),

...these metaphysical and explanatory aspects seem to be in tension. On the one hand, it is supposed that metaphysics concerns phenomena in the world itself independently of the explanatory interests and goals of inquirers like us. On the other hand, it is supposed that explanations are sensitive to the explanatory interests and goals of inquirers like us.

As I will argue, this tension is serious.

According to the unfiltered grounding-based theory of explanation, grounding is custom-made to be the relation tracked by metaphysical explanation. The reason it does a splendid job of according with the various desiderata is that explanatory considerations are appealed to in constraining instances of grounding (see §3.1.2; §3.4). But our explanatory intuitions are a strange thing to appeal to when characterising a supposedly mind-independent feature of the world. Trogon (ms-a), for example, shares my intuition that if the unfiltered view is true “then our intuitions about the nature of explanation don't have a substantive role to play in our theorizing about grounding, as the [unfiltered theorist] claims that the relevant sense of explanation in this context is one such that explanations aren't epistemically constrained.” Trogon is here emphasising that the reason to prefer the unfiltered view stems from ontic intuitions about the objectivity of explanation. If explanation is an ontic phenomenon, it is odd to see appeals to epistemic considerations when characterising the candidate explanatory relation. And yet, as was made very clear above, intuitions about the nature of explanation *are* playing a very substantial role in our theorising about the nature of grounding, *particularly* for unfiltered theorists who believe that there are no grounding relations in cases where we lack priority intuitions.

I will draw out the illicit nature of this move by looking once again to the literature on scientific explanation. My argument, in short, is as follows:

⁶⁶ It is surprising, here, to see Raven expressing the intuition that it is in the nature of explanation to form chains. My intuition runs contrary to this (consider the chain of causal/scientific explanations running from the big bang to the writing of this sentence). This simply seems to be a clash of intuitions.

1. Causation* is a primitive, mind-independent, *in re* relation, which obtains only in the cases in which we find causation to be explanatory. (stipulation)
2. It is illicit to posit causation*.
3. Grounding is a primitive, mind-independent, *in re* relation which obtains only in the cases in which we find the traditional modal relations to be explanatory.
4. Positing grounding is analogous to positing causation*.
5. If positing grounding is analogous to positing causation*, then if it is illicit to posit causation*, it is illicit to posit grounding.
6. Therefore, it is illicit to posit grounding.

As such, it is not the fact that the unfiltered grounding-based theory does so well with the epistemic desiderata that should arouse our suspicion. It is the fact that it does so while insisting that we are tracking a mind-independent relation, and thus securing the ontic desideratum of objectivity.

In order to defend premise (2), let's look once again to the space of views on scientific explanation. It is notable that, in that literature, there is no analogous view to the unfiltered grounding-based theory of metaphysical explanation. Indeed, there are no views where we establish what instances of the candidate relation obtain by appealing to epistemic intuitions. Furthermore, when we sketch out how such a view would look in that context, it becomes increasingly clear that a worldly relation crafted with explanatory constraints in mind is a posit whose existence we might reasonably give low prior credence. To see this, we now make a detour via the causal and DN theories of scientific explanation discussed in §2.5.

Recall that, at the most general level, such theories suppose that there are candidate relations in the world, and that our scientific explanations track either some or all instances of a candidate relation. For example, the unfiltered causal process theory has it that wherever there are causal relations, there are corresponding scientific explanations. This view retains the objectivity of what explains what, yet runs into trouble with the relevance and understanding desiderata. There are many events which stand in causal relations to explananda to which they intuitively lack relevance. For example, many explananda causally depend upon events despite a lack of counterfactual dependence, which is a standard way to evaluate causal relevance. Furthermore, there are causal relations from the big bang to any proposed explanandum, yet the big bang happening

the way it did does little to increase our understanding of why many current events happen the way they do. The proponent of this view accepts its epistemic flaws, for she thinks that explanation is a matter of how the world is, rather than an interest-relative affair that must impact agents' psychologies in the right way. Moreover, she has an independent characterisation of causation (see §2.5.2), which is used to establish what instances of that relation obtain. Such an independent characterisation is unfortunately lacking in the case of grounding, allowing the possibility of illicit pre-filtering.

Consider now the unfiltered DN theory. As well as struggling with providing only relevant explanantia, it struggles to achieve asymmetry. Recall that, in the light of the apparent time-reversal invariance of the physical laws, it seems that if A nomically entails B, B also nomically entails on A. So, the door is open to unwanted symmetrical DN derivations. This view also struggles with the same 'understanding' cases that beset the unfiltered causal theory. So, the unfiltered DN theory implies that there are irrelevant, symmetrical and unilluminating scientific explanations.

In order to retain objectivity of what explains what, the defender of either of these unfiltered theories must sacrifice other desiderata. But, says the defender of the epistemic conception of explanation, why call these *theories of explanation* at all? It is a serious flaw in a so-called theory of explanation that it allows irrelevancies, symmetries or unilluminating cases. So, we might choose to jettison objectivity, and seek the (perhaps more attractive) filtered versions of these theories. We note that, for creatures psychologically oriented as we are, accounting for events in terms of DN derivations running from the temporal direction we call 'past' to the one we call 'future' is much more useful than tracking those derivations running the other way. So we filter out the derivations heading 'the wrong direction'. Likewise, we might filter out the DN derivations and causal relations that include seemingly irrelevant explanantia, and those that fail to increase our understanding of their respective explananda. As they don't have the right epistemic characteristics, we refrain from tracking such relations with our scientific explanations.

Yet, all the while it is left open that, had our psychologies been different, different instances of these respective relations would have made the explanatory cut. For instance, had the entropy gradient been running the other temporal direction in this part of the universe, we might have preferred to appeal to the DN derivations running in the other temporal direction. Perhaps, were we *very* different psychologically, we would find

derivations running in *both* directions to be explanatory. In the case of causal relations, had we more advanced brains we might find that the way the big bang occurred *does* illuminate current small-scale events, and thus not filter out these causal relations between temporally distant relata. As such, the filtered theories are openly interest-relative. They give up on objectivity, and embrace the thought that *how we are* matters to what explains what.

This is all well and good. But imagine that someone proposes a theory which ostensibly requires no filtering—and thus retains objectivity—and yet captures only the explanations countenanced by the psychologically filtered theories. This new theory claims that scientific explanations reliably track ontologically robust, *in re* relations of causation*. Furthermore, as it turns out, there are instances of causation* running alongside all and only those causal processes that seem to us, to be explanatory.

Causation* is a primitive, explanatory relation, we are told. Thus its instances are limited by explanatory constraints. As nothing explains itself, causation* is irreflexive. As symmetrical explanations are problematic, causation* runs only from past to future. As an explanatory relation, causation* is constrained by relevance and only obtains in cases that increase our understanding. So, not every event which causes another can also be said to cause* it. Causation* is also non-monotonic: Just because A causes* B and (A & C) causes B, we cannot infer that (A & C) causes* B, as C might be irrelevant to B.

Yet, amazingly, proponents of the causal* theory of scientific explanations insist that what causes* what is not interest-relative. Causation* is simply a matter of what the world is like, for causal* relations exist *in re*. Thus the causal* theory of scientific explanation complies with the objectivity desideratum, while maintaining asymmetry, relevance and understanding. So, the causal* theory strictly does better according to the desiderata than any of the competing theories (both filtered and unfiltered). In fact, it meets *all* the desiderata.

Despite its success in meeting the desiderata, it is no surprise that nobody has built a theory around causation*, and nor does anybody believe in such a relation. This is because we are naturally hesitant to posit a primitive relation which just happens to coincide with the instances of causation which seem explanatory to us. Indeed, that they obtain alongside causal relations which seem explanatory to us seems to be the sole feature that unites the causal* relations. If someone were to ask what causation* is,

beyond noting its logical features one would naturally point to a bunch of examples, and these examples would all be ones in which causation obtains, and that causal relation is explanatory. Thus, to posit the existence of causation* either seems highly chauvinistic (if there only exists an explanatory relation tailor-made to *our* psychological orientation) or a heavy cost to parsimony (if there is a primitive *in re* relation corresponding to any possible psychological orientation). Furthermore, in the latter case we can hardly claim to have developed an objective theory of scientific explanation. For, the resulting view is that agents with a psychology like ours reliably track causal* relations, whereas agents with different psychologies track causal** relations (and causal*** relations, and so on). As there is nothing privileged about causal* relations, why not use the filtered causal process theory to do this work?

So, it is clear enough that the theory that scientific explanations reliably track causal* relations is a bad one, for despite its claims of being an ontic theory of explanation, the ontology on which it is based seems to have been suspiciously custom-made on the basis of epistemic considerations. Even though, once posited, causation* unarguably allows us a theory of scientific explanation that does very well with the desiderata, it is surely a mistake to posit causation* in the first place. Thus goes my defence of premise (2).

Regarding premise (4), it is obvious that in the above story, causation* is supposed to be the analogue of grounding. Perhaps less obvious is that causation is the analogue of the traditional modal relation of necessitation. All parties to the dispute about causal process theories of scientific explanation can agree that causation exists and where it obtains (those who defend other theories—DN theories, for instance—may disagree, but we will leave these views to the side for now). Likewise, those who defend and those who despise grounding are equally happy to admit that there is necessitation, and they agree on what necessitates what. (Necessitation is merely modal correlation, after all.) Those who defend the unfiltered grounding-based theory, however, think that as well as necessitation relations, there are also grounding relations which exist *in re* and coincide with precisely those instances of necessitation which we find explanatory (assuming a necessitarian view of grounding). Positing grounding, thus understood, looks very much like positing causation*. Thus, as positing causation* is illicit, and grounding is analogous to causation*, positing grounding is illicit.

The argument by analogy to causation* can be thought of as one horn of a dilemma. If the defender of grounding cannot identify some non-anthropocentric commonality—some unifying feature—of those instances of necessitation where we have intuitions of priority, then grounding is objectionably similar to causation*. On the other horn, imagine that (unlike causation*, say), there are some non-anthropocentric features shared by all of the instances of necessitation where we have corresponding priority intuitions. This is what defenders of grounding would have us believe. If this is so, then a filtered yet objective version of the modal relations theory considered in §2.2 emerges. Such a view will have it that ‘x *because* y’ is true just in case y bears an appropriate modal relation to x, *and* said modal relation has such and such non-anthropocentric features. Such a view will satisfy the very same desiderata as a grounding-based theory, without the ontological cost of positing grounding.

That is to say, *even if* there are some non-anthropocentric features which unify the necessitation relations that obtain where we have priority intuitions, the above view looks superior to the grounding-based view. As such, *whether or not* there is some non-anthropocentric commonality between the instances of necessitation where we have priority intuitions, there is no call to posit grounding relations. Thus, my arguments against grounding-based theories go through whether or not there is some such unifying feature. However, I am sceptical—not agnostic—regarding whether there is a non-anthropocentric commonality to be found. That is, in the cases where necessitation is symmetrical, there seems to be nothing shared by the relations going ‘the right way’ except our intuitions regarding the direction of explanation. Similarly, there is nothing in common between the relevant and illuminating instances beyond their effect upon us.⁶⁷

For, it is a contingent feature of our psychology that we find explanatory those instances of necessitation which we, in fact, do. Indeed, there may be Martians who find the existence of sets to be a good metaphysical explanation of the existence of their concrete members, and think that all necessary truths are mutually explanatory. On what basis can we say that the in-the-world grounding relations show that we are getting the explanatory order right, whereas the Martians are simply wrong? We have used our own explanatory intuitions to judge where these grounding relations are! Thus, a theory which

⁶⁷ Maybe something can be said to rehabilitate the idea that there is some objective notion of relevance, such that it is not anthropocentric to prefer the necessitation relation between [Michael exists] and [a man exists] over that between [Michael exists and is wearing a hat] and [a man exists]. However, I am not sanguine about the prospects of extending this idea to find something non-anthropocentric in common between all of the instances of grounding.

makes explicit the role of human psychology in determining our intuitions of what metaphysically explains what, will be advanced in Chapter 5.

That, then, is the explicit argument behind my accusations of illicit pre-filtering. A potential objection to my argument by analogy is as follows. It is odd that the kind of suspicion which would surely arise if someone proposed causation* has not been aroused in the context of grounding. Now, this might be thought to be evidence for a disanalogy between the cases: nobody would buy causation*, but lots of people buy grounding, so there must be some relevant difference between the two. I have three things to say about this objection. The first is that there is some published scepticism about grounding (see Wilson, 2014; Koslicki, forthcoming; Daly, 2012). The second is the anecdotal report that many, in fact, don't buy grounding. However, due to their view, such people are not sufficiently interested in grounding to publish their scepticism. Thus there is a selection effect whereby the vast majority of papers published on grounding are friendly to the notion, simply because people only want to write about it if they think it's worth writing about.

Finally, I suspect that there is something in the scientific respectability of causation that makes it so obviously laughable that philosophers might claim that there is this other relation that runs alongside some instances of that relation. In other words, our motivation for positing causal relations stems, at least in part, from our scientific understanding of what causes what, and the fact that there is no notion of causation* in science rightly raises questions about whether it is legitimate to posit such a relation in the service of a theory of scientific explanation. In contrast to the causation literature, the literature on grounding is relatively unconstrained by scientific practice, for science is mostly in the business of providing scientific, rather than metaphysical, explanations. As such, that grounding relations lack scientific respectability has not been viewed to their detriment, whereas causation*'s lack of scientific respectability would contribute to its downfall, were anyone ever to posit such a relation.

In sum, the analogy to causation* poses a serious challenge, not only for defenders of the unfiltered grounding-based theory of metaphysical explanation, but for defenders of grounding more generally, given the way that explanatory intuitions are used in order to establish the features of the grounding relation. Moreover, those who insist that grounding runs alongside only the instances of necessitation we find explanatory while

maintaining that what grounds what is an objective feature of the world owe us a substantial story about the commonality between these instances. Furthermore, this commonality cannot be anthropocentric. Only a non-anthropocentric story of this kind can show why grounding should be treated differently to causation*.

Relating this argument back to the broader dialectic of the thesis, while the grounding-based *vindication* of our priority intuitions has a certain attractiveness that stems from the neat matchup between our priority intuitions and the grounding relations, this matchup issues from the fact that grounding relations are tailor-made to do certain explanatory work. Indeed, any vindication of our intuitions about scientific explanation framed in terms of causation* would be similarly attractive, but this is insufficient reason to posit causation*. Ultimately, we should be sceptical that there is anything in the world that does *such* a good job of vindicating this kind of intuition. So, the impressive matchup between instances of grounding and our priority intuitions should ultimately undermine our confidence that such relations exist, rather than bolstering said confidence.

3.5 Evaluating the Filtered Grounding-Based Theory

The second theory of metaphysical explanation that becomes available when we posit grounding is the filtered grounding-based theory. To reiterate, the idea here is that our metaphysical explanations track grounding relations, but not every grounding relation has an associated metaphysical explanation. We filter the grounding relations in order to isolate only those that have the right epistemic character, and this subclass of the grounding relations backs the metaphysical explanations.

The first thing to note about this kind of view is that, if explanatory considerations are appealed to in establishing the characteristics of grounding, and where it obtains, the argument by analogy to causation* (§3.4.1) will apply here, too. In fact, it will apply with renewed force, for it is even more bizarre that we would use explanatory intuitions to *first* establish the features of a relation that we then go on to *further* filter according to our explanatory intuitions. Thus, the charitable way to understand the view is one whereby grounding is not beholden to our explanatory intuitions. For, the instances of grounding being filtered are not the same instances identified by the unfiltered theory. The filtered theory identifies the same class of explanations as the unfiltered theory by positing *more*

grounding relations, and then psychologically filtering them.

This view fulfils the understanding desideratum where the unfiltered view arguably could not. Recall the putative problem case where the arrangement of microphysical particles grounds a high-level economic fact. If cases like this do, in fact, serve to increase our understanding, then the filtered view will not filter them out, and thus the view will identify the same class of metaphysical explanations as the unfiltered view (but without the pretensions of objectivity). On the other hand, if such cases fail to increase our understanding, the filtered theorist has the power to rule them out. For, on the current view, there can be a transitive chain of grounding relations, and yet the objectionable explanation is ruled out by psychologistic filtering. In this case, the view will identify a smaller class of metaphysical explanations than the unfiltered view.

Thus, it is open to the filtered grounding-based theorist to allow that, while grounding is transitive, metaphysical explanation is not. This is why Trogdon claims that it is “natural to interpret Schaffer’s [counterexamples to the transitivity of partial grounding] as targeting metaphysical explanation. This is compatible, however, with grounding proper being transitive” (2013a:9-10). Thus, on this view, the world can be organised hierarchically (grounding can have a partial ordering structure) without commitment to epistemically dubious metaphysical explanations.

Transitivity is paradigmatic of any kind of ordering relation. Thus, the transitivity of grounding can be motivated by an allegiance to the idea that grounding limns the world with a hierarchical structure, giving rise to Schaffer’s great chain of being. Though, as we saw above, Raven (2012) appeals to explanatory considerations to motivate transitivity, one need not do so, for it can be motivated by appealing to the notion of structure. Likewise, one could motivate asymmetry and irreflexivity on structural grounds (as it were). I have suggested, previously, that such appeals to structure ultimately collapse into appeals to the priority intuitions, but for now let us allow that some other story might be told.

Actual defences of the filtered view tend to be motivated by reasons to reject the epistemic desiderata as features of grounding, while keeping them as features of metaphysical explanation. For instance, Carrie Jenkins’ (2011) proposal that grounding is ‘quasi-irreflexive’ can be re-imagined as the view that grounding is not irreflexive—some facts ground themselves—while metaphysical explanation is irreflexive. As metaphysical

explanation, once separated from grounding, is all about how we conceive of the world, it is a hyperintensional phenomenon. In other words, names for the very same fact sometimes cannot be substituted into sentences of the form ‘x *because* y’ *salva veritate*. So, the ‘mode of presentation’ of the names for the facts matters for the truth of metaphysically explanatory claims.

To use Jenkins’ example, assume that the mind-brain identity theory is true. We might say that while there is only one thing there, it grounds itself. Yet, while [Gina is in mental state X] is in some sense the same fact as [Gina is in brain state X], on the (quasi-irreflexive) filtered grounding-based theory it can be true that ‘Gina is in mental state X *because* Gina is in brain state X’ but false that ‘Gina is in brain state X *because* Gina is in brain state X’. So, metaphysical explanation can remain irreflexive while tracking non-irreflexive grounding relations.

Precisely such a view is defended by Rodriguez-Pereyra (2015). He presents counterexamples to the asymmetry and irreflexivity of grounding. For example, he thinks that the truth-teller proposition (<this proposition is true>) is *self-grounding* and that there is *mutual grounding* between <this proposition is true> and <<this proposition is true> is true>.⁶⁸ Rodriguez-Pereyra’s pre-emptive response to those who will appeal to the explanatory nature of grounding to insist that he is mistaken is to allow that:

...it is true that explanations, in the sense of propositions put forward by an enquirer to advance our understanding, are asymmetric [...] But there is no reason to extrapolate the formal features of epistemic explanations to grounding. [...M]any explanations track grounding relations [...] But it does not follow that grounding must be asymmetric because epistemic explanations are. (2015:531).

Similar considerations apply to the relevance desideratum. On the unfiltered view it is clear that grounding must be non-monotonic. Nevertheless, Trogdon notes that the filtered view “is compatible with metaphysical explanation being non-monotonic and grounding monotonic.” (2013a:11).

So, there is a view here that finds some endorsement in the literature. However, I want to press three objections against this view. The first is aimed at views such as Rodriguez-Pereyra’s, which abandon the irreflexivity and asymmetry of grounding

⁶⁸ This can be translated into the language of facts for those who are uncomfortable with talk of grounding relations between propositions.

(indeed, it is aimed at any view which retreats from the irreflexivity, asymmetry, and non-monotonicity of grounding). The problem with such views is that they substantially compromise the power of the grounding-based theory to explain the priority intuitions. For, one reason that positing grounding seems to provide such a good explanation of our intuitions is that the direction of our priority intuitions is aligned with the direction of grounding. This is what makes the grounding-based explanation superior to the explanation in terms of the modal relations (§2.2).

If, however, grounding is a non-symmetric relation with symmetric instances, and if it obtains symmetrically between some x and y , the appeal to grounding cannot explain why we have the priority intuition according to which ‘ x because y ’ is true, while ‘ y because x ’ is false. Furthermore, if grounding turns out to be monotonic, it cannot explain why our priority intuitions are so sensitive to considerations of relevance. Indeed, if grounding turns out to be a non-symmetric relation, only some instances of which are accompanied by metaphysical explanations, the grounding relations are starting to look a lot like necessitation relations—relations that the defender of grounding has explicitly rejected as being capable of backing metaphysical explanations.

The first objection to the filtered grounding-based theory, then, is that once we divorce the formal features of grounding from the features of metaphysical explanation, the posit of grounding does far less work to explain the priority intuitions. For, if one chooses not to delimit the instances and formal features of grounding by appealing to features of explanation, the notion of grounding with which one is left is such that the explanatory and predictive power of the grounding-based theory is diminished, and the view opens itself up to the same criticism we levelled at the modal relations theory considered in §2.2. Specifically, the criticism is that there are instances of grounding where there is no corresponding priority intuition. Thus, the difference between those instances where we have priority intuitions and those where we do not remains in need of explanation. Consider that Raven (2015), a defender of the unfiltered view, insists on the need for grounding by noting that the work of backing our metaphysical explanations cannot be done by a relation that sometimes obtains symmetrically. The modal relations theory was found wanting, in part because not all instances of the modal relations seem explanatory. If grounding obtains in cases where there is no associated explanation, the grounding-based theory suffers the same explanatory disadvantage.

The natural response to this, on behalf of the filtered grounding-based theorist, is to appeal to her preferred way of filtering the grounding relations. She will note that those grounding relations where we have corresponding priority intuitions are those with a certain epistemic character, and it is this epistemic character which explains our intuitions. But—though we did not explore this possibility in Chapter 2—this very same move is available to the proponent of a modal relations-based theory. That is, the modal relations theorist can say that the reason that we have priority intuitions regarding only some of the modal relations is that these have a certain epistemic character. This leads us to my second objection, namely that the filtered grounding-based theory is strictly worse than the ‘filtered modal relations theory’ explored in Chapter 5. For, the following argument presents itself:

1. The filtered grounding-based theory and the filtered modal relations theory satisfy the very same desiderata (namely, every desideratum but objectivity).
2. The filtered modal relations theory is more parsimonious than the filtered grounding-based theory.
3. When all else is equal, we should prefer a more parsimonious theory.
4. Therefore, as all else is equal, we should prefer the filtered modal relations theory to the filtered grounding-based theory.

As I hinted in §3.4.1, I defend a psychologically filtered version of the modal relations theory in Chapter 5. On that view—roughly—for Bob’s utterance of ‘x *because* y’ to be true, y must bear the appropriate modal relation to x, and y must stand in the appropriate epistemic relation to x for members of Bob’s community.⁶⁹ The second criticism of the filtered grounding-based theory, then, stems from an unfavourable comparison with the filtered modal relations theory. The point, here, is that once we are filtering the grounding relations to find those that are explanatory, the advantage of positing grounding as well as traditional modal relations evaporates. For, while the unfiltered grounding-based theory boasts the (putative) advantage of objectivity over the filtered modal relations theory proposed in Chapter 5, the filtered grounding-based theory gives up on objectivity and thus loses this advantage. Thus, the filtered modal relations theory and the filtered grounding-based theory do equally well in fulfilling the desiderata.

⁶⁹ A much more substantial story is given in Chapter 5. Here I hope only to have given the flavour of the theory.

The picture, then, is as follows: there is a class of modal relations. Some of them seem explanatory. So, we could use psychologistic filtration to zone in on the explanatory subclass (à la the filtered modal relations theory). *Or*, we could posit relations that run alongside a subclass of the modal relations (posit grounding relations, that is) not all of which are explanatory, and *then* psychologistically filter *these* grounding relations to zone in on the explanatory subclass (à la the filtered grounding-based theory). It is clear that grounding is doing nothing to make the second picture better than the first. It is an entirely unneeded middleman which serves only as a cost to the parsimony of our theory.

So, while Trogon motivates the need for grounding by telling us that “As was noted years ago, however, there isn’t a corresponding connection between explanation and either supervenience or modal entailment” (2013a:4), if it turns out that we need to filter the grounding relations just like we filter the modal relations, it seems like the connection between grounding and metaphysical explanation is no different in kind than that between the modal relations and metaphysical explanation.

Once again, a useful comparison can be made to causation*. In this case, rather than perfectly matching our explanatory intuitions, causation* is playing an awkward middle-ground role: it runs alongside all those instances of causation where we have explanatory intuitions, and some other instances too. The theory is that our explanations track a psychologistically filtered subset of the instances of causation*. It is hard to see what work causation* is doing in this theory, and why we should believe in it.

The third objection proceeds by pointing out that the already flimsy epistemology of grounding becomes substantially flimsier when the relation is separated from metaphysical explanation. That’s because the way we find out about what grounds what, if indeed we can do so, is via our priority intuitions, which are intuitions about metaphysical explanation. The filtered grounding-based theorist is therefore in an awkward position whereby she must tell us a story about how we come to know about the grounding relations which *aren’t* tracked by our metaphysical explanations. This cannot be done through empirical investigation, which can at best give us insight into the obtaining of modal relations (what supervenes upon what, for example). Perhaps some story can be told about how we extrapolate from those cases where we *do* have priority intuitions to help determine what other cases of grounding there may be. However, what features would make such cases relevantly similar is unclear.

We can find both a useful statement of the problem, (and a confusing denial of it) in a recent paper from Trogdon. The problem, as expressed by Trogdon's (ms-a) defence of separatism, is that

...if intuitions about explanation are off the table when it comes to theorizing about grounding then grounding is theoretically unconstrained in an objectionable way. The idea is that these intuitions provide a crucial fixed point with reference to which we can evaluate claims about grounding, such as the claim that grounding is asymmetric.

However, Trogdon thinks that intuitions about explanation are *not* off the table, even in the context of the filtered/separatist view. He says:

If grounding backs explanation then it makes perfect sense to appeal to intuitions about the nature of explanation—intuitions governed by epistemically constrained conceptions of explanation—in our theorizing about the nature of grounding (assuming, of course, that intuitions in general have a role to play in our theorizing). Returning to the feature of asymmetry, if separatism is true then we have reason to think that grounding is asymmetric given our intuition that explanation itself is asymmetric. Roughly speaking, the idea is that the formal features of structure-imposing explanation-backing relations mirror the formal features of the relations they back.

The idea, I take it, is that even if we grant that some grounding relations aren't explanatory, we can still have reason to think that grounding has the formal features of explanation because metaphysical explanations track grounding relations. Indeed, maybe those are the only reasons we could have to ascribe features to grounding. Yet, as argued by Rodriguez-Pereyra (2015) it is not obvious that Trogdon is correct in supposing that the formal features of structure-imposing explanation-backing relations mirror the formal features of the epistemic explanations they back. Consider, for example, that the nomic determination relations proposed by the defender of the DN theory are symmetrical, and yet (appropriately filtered) back asymmetric explanations. Likewise, perhaps Price (1996) is right that, when characterised in a mind-independent way, causal relations turn out to be symmetrical. Still, these relations can back asymmetric explanations. Thus, we should at least be cautious, and not assume, for example, that

“The fact that grounding is systematically connected with explanation in [the filtered] way suggests that grounding, like explanation generally speaking, is asymmetric.” (Trogdon, ms-a).

While I think that this latter objection is a substantial challenge for the filtered grounding-based theorist, I will be focussing on the former two. The following chapter will develop a psychological explanation for our priority intuitions, and Chapter 5 will use this explanation to build a psychologically filtered version of the modal relations theory. The filtered grounding-based theory has no advantages (and a substantial disadvantage when it comes to parsimony) in comparison to the filtered modal relations theory, and is thereby strictly a worse theory.

3.6 Conclusion

I concede that the posit of grounding can help provide us with an explanation for the priority intuitions, and that this explanation also serves to vindicate these intuitions. However, the Sophisticated Grounding Theory is less theoretically virtuous than the Grounding-free Theory I develop in the following chapter. Thus, the indispensability of grounding cannot be established by appealing to the theoretical virtues.

Neither does the appeal to the connection between grounding and metaphysical explanation serve to establish the superiority of the Sophisticated Grounding Theory. For, there are three ways grounding could relate to metaphysical explanation, none of which are promising.

Firstly, there could be no connection between the two. In that case, theorising about metaphysical explanation gives us no reason to believe in grounding.

Secondly, grounding might coincide almost perfectly with our priority intuitions, as per the unfiltered grounding-based theory. Then, as a putatively mind-independent feature of the world which mysteriously exhibits various epistemic virtues (relative to our psychological orientation), grounding is analogous to causation*. Grounding, in other words is an illicitly pre-filtered ontological posit. Thus, though it can do good explanatory work once posited, we should refrain from doing so.

Thirdly, grounding might fail to coincide perfectly with our priority intuitions, as per the filtered grounding-based theory. Then, appealing to grounding does little to explain why we have the priority intuitions we do, and indeed it is quite mysterious how we could come to know about grounding. Furthermore, the filtered grounding-based theory of metaphysical explanation is strictly worse than the filtered modal relations theory.

In sum, no matter the putative connection between grounding and metaphysical explanation, the fact that we care about the latter gives us no reason to posit the former. In the following chapters I will develop a grounding-free explanation of the priority intuitions, alongside some accompanying grounding-free theories of metaphysical explanation. I will show how the former is more theoretically virtuous than the grounding-based explanation, and how the latter are preferable to the grounding-based theories of metaphysical explanation.

Chapter 4: A Psychologistic Explanation of the Priority Intuitions

We are now in a position whereby we have carefully explicated the explananda, and narrowed our focus to the explanation of the priority intuitions. Moreover, we have seen that, while positing grounding allows us to explain and vindicate these intuitions, there are independent reasons to be hesitant to posit grounding relations, which are, as I have put it, *illicitly pre-filtered*. Recall that, in order to demonstrate the dispensability of grounding relations, I need to show that the Grounding-free Theory can explain the priority intuitions in a way that is empirically on a par with the Sophisticated Grounding Theory, yet more theoretically virtuous (Colyvan, 1999; §1.3). Here, I will develop an explanation of the priority intuitions that appeals to the modal relations and evolved psychological mechanisms. Not only can this Grounding-free Theory predict that we would have the priority intuitions that we do, it is also empirically supported, and it is more parsimonious (as it eschews positing primitive grounding relations) and elegant (as it eschews the ideology associated with grounding). On the assumption that the defender of the Sophisticated Grounding Theory helps herself to my psychological story, that theory and the Grounding-free Theory are equally unified, each appealing to mechanisms that also serve to explain other phenomena of interest. In particular, the mechanisms described here also serve explain our intuitions of *diachronic* (causal) priority.

Here is the plan. In §4.1 I quickly recap the intuitions in need of explanation, and in §4.2 note how my preferred explanation shall proceed. §4.3 introduces the correlation and causal detection mechanisms that will do the explanatory work, while §4.4 shows how the former can detect modal correlations. §4.5 describes how the causal detection mechanism, which serves primarily to filter diachronic correlations, has been co-opted to filter the modal correlations in search of non-symmetries. §4.6 provides an overview of the empirical literature regarding the tendency of the causal detection mechanism to *overgeneralise*. This sets the stage for §4.7, where I show how those cases where a modal relation obtains symmetrically, yet we have an asymmetric priority intuition, are precisely those which have the features which cue the causal detection mechanism to overgeneralise. Thus, the mechanism leads us to believe that some instances of the modal relations are non-symmetric when they are, in fact, symmetric.

This synopsis is, admittedly, all too brief. However, as noted, I will provide a more comprehensive pre-emption of the structure of my explanation in §4.2.

4.1 The Observations

Recall that, according to the master argument, in order to demonstrate the dispensability of grounding relations, I need to show how my Grounding-free Theory can do the same explanatory work as the Sophisticated Grounding Theory in a more theoretically virtuous way. We have already seen, in Chapter 3, that the grounding-based explanation (provided by the Sophisticated Grounding Theory) proceeds by piggybacking off the psychological story told here, and further supposing that these mechanisms are good grounding-detectors. Here, I develop an alternative, more theoretically virtuous explanation. Of particular note is that the view developed here refrains from positing any primitive entities, and is thereby more parsimonious than its grounding-based rival. However, what I offer in this chapter may not give us everything we want from such an explanation, for it does not, by itself, serve to vindicate our intuitions, nor provide a theory of metaphysical explanation. Extending the grounding-free explanation so as to achieve these further goals is the purpose of Chapters 5, 6 and 7. For now, we focus on providing a grounding-free explanation of the data.⁷⁰

However, before launching into the grounding-free explanation, we should quickly review the explanandum at issue. The explanandum concerns widely shared, convergent judgements of priority in certain cases of non-diachronic correlation. Let's begin with a reiteration of the representative list of priority intuitions from §2.1.2. There, we noted that many people judge the following to be true:

- A. The flower is red *because* the flower is maroon
- B. The bicycle exists *because* of the existence and arrangement of the wheels, spokes, handlebars, etc
- C. <a man exists> is true *because* Pythagoras exists
- D. {Pythagoras} exists *because* Pythagoras exists
- E. <Pythagoras exists> is true *because* Pythagoras exists
- F. God loves X *because* X is good
- G. [Pythagoras exists] obtains *because* Pythagoras exists
- H. $2+2=4$ *because* 2 exists and 4 exist

whilst judging the following to be false:

⁷⁰ Recall that 'the grounding-free explanation' refers to the explanation offered by the Grounding-free Theory.

- a) The flower is maroon *because* the flower is red
- b) The existence and arrangement of the wheels, spokes, handlebars, etc exist *because* the bicycle exists
- c) Pythagoras exists *because* <a man exists> is true
- d) Pythagoras exists *because* {Pythagoras} exists
- e) Pythagoras exists *because* <Pythagoras exists> is true
- f) X is good *because* God loves X
- g) Pythagoras exists *because* [Pythagoras exists] obtains
- h) 2 exists and 4 exists *because* $2+2=4$

These are the representative intuitions I shall work with in elucidating the grounding-free explanation. I take it that a good explanation of these intuitions will explain both (i) why there is an appearance of asymmetry in these cases (i.e. we judge that (A) is true and (a) is false) and (ii) why it seems to us that there is an explanatory connection between what lies to the left, and what lies to the right, of the '*because*'. I contend that my preferred explanation can do so. However, recall that the Sophisticated Grounding Theory makes use of the story I offer here to do this work, so the mere fact that I can explain these intuitions without grounding (divorced from considerations arising from the theoretical virtues) does little to distinguish between the theories.

The priority intuitions have, in some sense, two components:

- 1) We observe that certain objects, facts, or properties, are non-diachronically correlated.
- 2) We observe that, for many of these correlations, there are widely shared priority intuitions.

Here, I shall provide an explanation for 1) and 2) that does not appeal to the presence of grounding relations. Rather, the explanation appeals to the functioning of a set of two inter-connected cognitive mechanisms: a correlation detection mechanism, and a causal detection mechanism. All this is, of course, in service of defending premise 2 of the master argument: that grounding relations are *not* indispensable to the best explanation of our observations.

4.2 General Schema of an Explanation

I will begin by laying out the general form my explanation will take. I aim to explain our priority intuitions in terms of the functioning of two cognitive mechanisms—a correlation detection mechanism and a causal detection mechanism—where the latter acts as a filter on the outputs of the former. The idea is that the causal detection mechanism functions by filtering the outputs of what I call the *correlation detection mechanism* by using certain cues to search for asymmetries amongst the correlations.⁷¹ I defend this way of thinking about the relevant mechanisms in §4.3. I then expand on the following five steps that jointly constitute a grounding-free explanation of our priority intuitions.

First, I propose that we use our correlation detection mechanism to detect both diachronic and non-diachronic correlations: this mechanism explains (1) above. For, I suggest, given that there is utility in tracking non-diachronic correlations, it is plausible in the extreme that a general correlation detection mechanism will detect not only diachronic but also non-diachronic correlations.

The second step appeals to the thought that in detecting non-diachronic correlations we are sometimes thereby detecting modal correlations: traditional modal relations such as necessitation and supervenience. The thought is this: some⁷² instances of non-diachronic correlation are also instances of modal correlation. So, for instance, when we detect a non-diachronic correlation between certain brain states and certain mental states, we thereby detect a particular modal correlation between those states

⁷¹ A refresher on the relevant terminology might be useful, here. Symmetric relations are ones in which, for any x, y , if $x R y$, then $y R x$. Asymmetric relations are ones in which, for any x, y , if $x R y$, then it is not the case that $y R x$. And non-symmetric relations are ones in which, for some x, y , if $x R y$, then it is not the case that $y R x$. Hence, all asymmetric relations are non-symmetric, but not *vice versa*. In what follows there will be a need to extend this terminology so as to apply to instances of relations. Accordingly, I will call an instance of a relation in which $x R y$, and $y R x$, a *symmetric instance* of that relation, or, alternatively, I will say that the relation *obtains symmetrically*. Likewise, I will call an instance of a relation in which $x R y$ and it is not the case that $y R x$, a *non-symmetric instance* of the relation, or, alternatively, I will say that the relation *obtains non-symmetrically*. Thus, a symmetric relation will only have symmetric instances, an asymmetric relation will only have non-symmetric instances, and a non-symmetric relation will have at least one non-symmetric instance. Given this terminology, there is no such thing as an asymmetric instance of a relation: asymmetry is a property of relations, not instances. I hope that this terminology draws attention to the fact that non-symmetric instances of a relation share the same formal features, whether they are instances of a non-symmetric relation or an asymmetric relation. That will subsequently become important.

⁷² I say ‘some’ here because there are many actual non-diachronic correlations that are not instances of modal correlations: for instance, the correlation between the set containing me, and the set containing you. Yet, correlations such as this are not the kind that attracts the attention of our correlation detector. Those that do tend to be indicative of modal correlations.

because the non-diachronic correlation is also a modal correlation. I spell out the details of steps 1 and 2 of my explanation in §4.4.

The third step demonstrates how the causal detection mechanism can identify, amongst the modal correlations, *non-symmetric* instances of supervenience or necessitation. Recall from Chapter 2 that I am conceiving of the traditional modal relations purely in terms of modal correlation. The modal relations are non-symmetric, for they have instances that are symmetric, and instances that are non-symmetric (§2.1.1). This is in contrast to dependence relations like grounding and causation, both of which are typically thought to be asymmetric.

I suggest that the causal detection mechanism has been co-opted to filter the non-diachronic correlations detected by the correlation detection mechanism. It successfully does so by seeking the same sorts of cues it uses to separate causal relations from mere diachronic correlations. Since any instance of causation is non-symmetric, the causal detection mechanism searches for instances of non-symmetrical relations amongst the correlations; in doing so it detects instances of causation. When it applies the same procedure in the case of non-diachronic correlations, it filters non-symmetric instances of non-diachronic relations from symmetric instances. Where it does so successfully, our priority intuitions are explained by us successfully detecting non-symmetric instances of a non-diachronic relation: a modal relation (this is how I explain cases (A) and (B)).

Recall that in making those judgements, individuals express the belief that an asymmetric explanatory relation obtains between the relata. I can explain the appearance of *asymmetry* by noting that the relation in question holds *non-symmetrically* in such cases, and a non-symmetric instance might be an instance of either a non-symmetric⁷³ relation, or an asymmetric relation. Since the mechanism is tuned to detect causal relations (which are asymmetric), it is unsurprising that we infer that the instances thus filtered are instances of an asymmetric, rather than a non-symmetric, relation. Further, I explain the appearance of *explanation* in these cases as the result of a trigger produced by the causal detection mechanism. That mechanism evolved to signal the presence of a dependence relation (causation), which *does* back explanation, and, in the process, to trigger a phenomenology as of there been an explanation present. When the same mechanism detects the presence of a non-symmetric instance of a non-diachronic relation, that same phenomenology is often triggered. I develop this aspect of the explanation in §4.5.

⁷³ But not asymmetric.

Step four of the explanation appeals to empirical evidence which shows that our causal detection mechanism sometimes overgeneralises, signalling the presence of a non-symmetric instance of a relation—causation—where no such instance obtains. It does so because failing to detect an instance of causation that obtains, is more costly than signalling the presence of causal relations where there are none. If such mistakes occur when filtering the diachronic correlations, there is every reason to suppose they also occur when filtering the non-diachronic correlations, since the latter filtration utilises the same mechanisms as the former, co-opted for a slightly different purpose. So we should expect cases in which our causal detection mechanism mistakenly signals, amongst the non-diachronic relations, the presence of a non-symmetric instance where no such instance obtains.

Step five of the explanation points out that there are environmental conditions under which we should expect the causal detection mechanism to signal that a non-diachronic relation is non-symmetric. Moreover, these conditions are those we find obtaining with respect to the objects, properties or facts mentioned in (C) through (H). So we should *expect* that, having detected the correlations in those cases, the causal detection system will signal that those correlations imply the obtaining of non-symmetric instances of a modal relation. And that is exactly what we find. But we would expect this even on the supposition that the instances of those modal relations are symmetric. So we can explain the appearance of *asymmetry* expressed by our priority intuitions in the same way it is explained in cases (A) and (B), except in (D) through (H) the relevant mechanism *mistakenly* signals the presence of a non-symmetric instance of a relation (rather than correctly signalling its presence). With respect to (C) the mechanism gets the right answer—it detects a non-symmetric instance of a modal relation—but it does so for the wrong reasons, namely, because it is sensitive to cues that would have resulted in it signalling the presence of such a relation even if one had not been present. Finally, I explain the appearance of *explanation* in the same way as it was explained for cases (A) and (B), as the result of a trigger produced by the causal detection mechanism. I defend steps 4 and 5 in §4.6 and §4.7.

By the end of §4.7, then, the priority intuitions will be explained in a way that appeals, in an indispensable manner, to the existence of certain psychological mechanisms and modal relations, but does not appeal, indispensably or otherwise, to the existence of any relation of ground. In §4.8, I note that, as this explanation can account

for the empirical data just as well as the grounding-based explanation, yet is more parsimonious and elegant, and enjoys empirical support, it is to be preferred. Furthermore, while the grounding-based theory arguably provides a simpler explanation of the *priority intuitions*, the Grounding-free Theory is simpler overall.

4.3 The Correlation and Causal Detection Mechanisms

There is a good deal of evidence that we have a correlation detection mechanism, though exactly how that mechanism functions (and which brain processes subserve it) is more controversial. What is agreed is that this mechanism allows us to distinguish information bearing patterns from random patterns, and to quantify the information bearing patterns. It functions by taking in inputs, namely, the frequency of the presence, or absence, of certain features in the environment. In the literature, these frequencies are represented by what is known as a contingency table—a matrix that displays the frequency distribution of variables. A simple version of a contingency table is below (figure 2).

		Factor 2	
		Present	Absent
Factor 1	Present	A	B
	Absent	C	D

Figure 2. (From Arkes and Harness (1983)).

Attempts to understand the correlation detection mechanism focus on determining which heuristic we use to detect correlations. The assumption is that we represent (perhaps sub-personally) something like a contingency table of data, then use a heuristic to determine whether a correlation obtains between the relevant data. Investigators have identified four candidate heuristics. The first (known as the Cell A rule) focuses entirely on the frequencies in Cell A (see, e.g., Smedlund, 1963; Nisbett and Ross, 1980). The second, called the A minus B rule, holds that we are sensitive to frequencies in Cells A and B: the more the frequencies in cells A and B diverge the higher the correlation is

judged to be (Shaklee and Mims, 1982). The third rule is known as the sum of diagonals. Here, we compare $A + D$, with $B + C$ (Shaklee and Tucker, 1980). The strength of the correlation signalled depends upon the extent to which these sums differ. According to the fourth rule, we use conditional probabilities (a Bayesian calculation) to evaluate two competing hypotheses: H_1 is the hypothesis that the data was produced by a random process, and H_2 is the hypothesis that the data was produced by some systematic process. The mechanism then uses Bayes' rule to combine prior beliefs about these hypotheses, with the evidence the data provides. More recently, studies suggest that subjects use all of these rules flexibly (Arkes and Harness, 1983).

In what follows I argue that the correlation detection mechanism is the basis of our ability to detect causal relations, and that it is the mechanism responsible for the detection of non-diachronic correlations. For the former claim to be plausible, it needs to be that the correlation detection mechanism is *good enough* at detecting correlations that, when filtered, it will generate the observed causal judgements. I think that very plausible. Research on the correlation detection mechanism often focuses on why the mechanism is so *inaccurate*. This inaccuracy, however, takes two forms. The first is the presence of false positives: cases where the mechanism signals the presence of correlations where none exist (known as illusory correlations; see Redelmeier and Tversky, 1996). The second are not false positives *per se*; rather, they involve the correct detection of a correlation, but an over-estimation of the strength of the correlation (Chapman and Chapman, 1967). It is easy to see why misjudgements of strength will sometimes occur if we use either the Cell A heuristic or the A minus B heuristic, since in either case we are ignoring important data (that contained in cells C and D).

Notice, though, that what matters for our purposes is that the correlation detection mechanism typically detects correlations that are there, not that it is always accurate in detecting their strength. If the causal detection mechanism filters the outputs of the correlation detection mechanism, then it is a virtue if the latter system is highly sensitive—if it is more inclined to produce false positives than false negatives—since false positives can be filtered out by the causal detection mechanism. One way in which the mechanism is thought to be highly sensitive is that (at some sub-personal level) we deploy Bayes' theorem (see Williams and Griffiths, 2013). Since the likelihood of most data sets is higher on the hypothesis that the data is non-random, than that it is random, the mechanism tends to yield false positives. Another way to put this is that some kind

of pattern can be detected in almost any data set, and the probability of that data set conditional on its being the result of some structure in the world, is typically higher than the probability of that data set conditional on it not being the result of some structure in the world. What this means is that the mechanism is highly sensitive to possible patterns in data, and will readily signal the presence of those patterns. This feature (as I will now argue) makes it ideal for detecting both diachronic and non-diachronic correlations.

4.4 Detecting Non-Diachronic Correlations

There is overwhelming evidence that we have cognitive mechanisms adapted to identify and track *causal dependencies*, as well as mechanisms that produce sophisticated *causal reasoning*. It seems to me that the best way to understand such claims is to suppose that our causal detection mechanism operates by filtering the outputs of our correlation detection mechanism. The latter signals the presence of (*inter alia*, diachronic) correlations. The causal detection mechanism then searches for cues to filter out those that are *mere* correlations, leaving those which are indicative of a causal relation. It does so by searching for *non-symmetries* amongst the correlations. I take it that searching for non-symmetries crucially involves searching for correlations in which: (a) changing X is likely to change Y and b) changing Y does not change X and (c) X occurs before Y (Sloman, 2005). The idea, here, is that the mechanism searches for non-symmetries amongst the correlations because any instance of the causal relation is non-symmetrical. So the mechanism aims to detect non-symmetries amongst the correlations, as a way of detecting causation.

There is a range of empirical data that supports the idea that the causal detection mechanism seeks out non-symmetries via various environmental cues. One such cue is the way in which the environment responds to an intervention. Since more than one causal model is consistent with any observed correlation, the only way to discover which causal model is the right one is for agents to perform an intervention.⁷⁴ This is because interventions cut off the thing upon which one intervenes, from any prior causes, but not

⁷⁴ See Gopnik et al. (2004); Kushnir, Gopnik, Lucas and Schultz (2010), Lagnado and Sloman (2004); Steyvers, Tenenbaum, Wagenmakers and Blum (2003).

from any later effects and thus have the capacity to reveal asymmetric dependencies.⁷⁵ As Haggmeyer et al. put it:

Interventions often enable us to differentiate amongst the different causal structures that are compatible with an observation. If we manipulate an event A and nothing happens, then A cannot be the cause of event B, but if a manipulation of event B leads to a change in A, then we know that B is a cause of A, although there might be other causes of A as well. (2007:87)

The process of intervention is, in effect, a process that aims to determine whether there is a non-symmetry present. If intervening on A intervenes on B, and not the converse, there is a non-symmetry present. Interventions, then, are one cue that the causal detection mechanism uses, to filter correlations. Where there is a non-symmetry detected, via intervention, this is a cue that the correlation is associated with a non-symmetric instance of some relation: in this case causation.

A second cue is temporal order (Sloman, 2005:6). Here again, the aim is to detect a non-symmetric instance of a relation, among the correlations. If x occurs before y, then y does not occur before x. This non-symmetry is a cue that there is a non-symmetric instance of a relation present: causation.

Finally, there are other cues that, among the correlations detected, some are backed by non-symmetric instances of a relation; namely prior knowledge, and an existing hypothesis about causal structure (Waldmann and Haggmeyer, 2013:745). Where some (or all) of these cues are present, the causal detection system signals the presence of a non-symmetric instance of a relation amongst the diachronic correlations. The non-symmetric instance in question is, in each case, an instance of causation.

4.5 Co-opting the Causal Detection Mechanism

So, we have a correlation detection mechanism that detects both diachronic and non-diachronic correlations, and a causal detection mechanism which filters the detected diachronic correlations. The next step of my explanation is to argue that the causal detection mechanism has been co-opted to also filter the *non*-diachronic correlations detected by the correlation detection mechanism. First, however, I suggest that in

⁷⁵ Sloman (2005).

detecting non-diachronic *correlations* we are sometimes detecting non-diachronic *relations*, of which the modal relations are a subset. Just as sometimes our detecting of a diachronic correlation is, *ipso facto*, detecting a causal relation, and just as detecting a particular determinable is, *ipso facto*, detecting a particular determinate, so too sometimes detecting instances of a non-diachronic correlation is, *ipso facto*, detecting a non-diachronic relation, and hence, in some cases, detecting a modal relation. If that is right, then our correlation detection mechanism is a mechanism that allows us to detect modal correlations. So we have an explanation for how it is that we track those relations (or at least, their actual world instances).

In other words, we have an explanation for (1), above. However, as noted, this does not yet confer an advantage over grounding-based theories, for the defender of grounding may well endorse the above as an explanation of how we come to be aware of the traditional modal relations.

In what follows I argue that our causal detection mechanism filters the non-diachronic correlations. If what I have just said is right, then in filtering these correlations it thereby filters instances of modal relations into those that are symmetrical and those that are non-symmetrical. Now, one might object, causation is an *asymmetric* relation. How could a mechanism evolved to track causal relations do the work I am suggesting? Well, notice that our detection mechanisms detect, and filter, *instances* of relations. So the causal detection mechanism is really a mechanism evolved to detect non-symmetric *instances* of a diachronic relation. By reliably detecting these non-symmetric instances, the mechanism is thereby detecting an asymmetric relation: causation.

Since that mechanism is sensitive to the formal features of *non-symmetry*, features shared by both diachronic and non-diachronic relations, it is easy to see how it could be co-opted to track non-symmetrical non-diachronic relations, of which the relevant relations, for our purposes, are the modal relations. Now, it might be that detecting non-symmetrical instances of modal relations is an adaptation of the causal detection mechanism. At worst, I think, our capacity to filter non-diachronic correlations is an *exaptation* of our causal detection mechanism. Exaptations are traits that are a by-product

of adaptive selection (they are not selected *for*) but which nonetheless come to be useful to the organism (Gould 1991:43).⁷⁶

How does the causal detection mechanism filter the non-diachronic correlations? Well, the sorts of environmental cues that allow it to discern whether there is an underlying non-symmetric instance of a relation are, by and large, the same sorts of cues for both diachronic and non-diachronic correlations. As I briefly noted previously, there are four important environmental cues to which the causal detection mechanism is sensitive.

- (i) Temporal order
- (ii) The result of intervention
- (iii) Prior knowledge
- (iv) An existing hypothesis about causal structure

Clearly (i) will only apply in the case of diachronic correlations. The remaining three cues, however, are relevant. Consider, first, interventions. Return to our example of the chair. Upon noticing that there is a correlation between the chair and its parts we can engage in an intervention. We can notice that there is no way of intervening on the chair without intervening on its parts. We can see this by trying to wiggle one thing (the chair) *without* wiggling the other (its parts) and by trying to wiggle one thing (the chair) *by* wiggling the other (its parts). Similarly, imagine an evolutionary ancestor of ours trying to establish the relationship between the deadliness of her spear and its other properties: she can try to wiggle the deadliness of the spear *by* wiggling those properties upon which the deadliness supervenes, or *without* wiggling these properties. In this way, the response of the environment, to interventions, can serve as a cue to the co-opted causal detection mechanism.

So too, presumably, can prior knowledge. Just as one might have causal knowledge that one can bring to bear in determining whether a diachronic correlation is due to causation, so too one might have knowledge one could bring to bear in determining whether a particular non-diachronic correlation is due to a non-symmetrical instance of a modal relation. Perhaps once I see the non-symmetric relation between chairs and their

⁷⁶ There has been an attempt to explain our tendency towards religious belief as a spandrel (an exaptation that is not useful): it arises not because it is adaptive but as a by-product of a host of other cognitive processes (Atran, 2002; Barrett, 2004; Boyer, 2001; Pyysiäinen, 2001; Pyysiäinen and Anttonen, 2002; Gould 1991:58). Notice that if this explanatory strategy is right, we do not need to suppose that we are good deity-trackers and that there are deities.

parts, I find it easier to see the non-symmetric relations between other sorts of objects and their parts. Finally, an analogue of (iv) might be relevant. It might be that an existing hypothesis about *modal* structure serves as a cue to help filter the non-diachronic correlations. For instance, I might have a prior hypothesis that token mental events are identical to token physical events; or I might have a prior hypothesis that any physical event like this one, will be correlated with a mental event like this one, but not *vice versa*. Either of these existing hypotheses might serve as a cue to the causal detection mechanism.

So far, then, I have argued that in detecting non-diachronic correlations, we thereby (sometimes) detect modal relations, and, in filtering those non-diachronic correlations we thereby identify non-symmetric instances of modal relations. Thus we are on our way to explaining (2), the observation that there are widely shared priority intuitions. In particular, I can now explain some of those judgments: namely (A) and (B). These are cases in which there is a non-diachronic correlation between the events, properties, or facts, in question. Features of these correlations correctly cue the co-opted causal detection mechanism to signal the presence of a non-symmetric instance of a modal relation.

Consider (B). At every world where those bicycle parts exist and are arranged appropriately, there is a bicycle, but there are worlds where the bicycle is composed of different parts, so the existence of the bicycle does not guarantee that *those* parts exist in *that* arrangement.⁷⁷ Plausibly, our causal detection mechanism is cued to this non-symmetry via the result of interventions (some counterfactual) since the only way of intervening on the bicycle is by intervening on its parts.

Now consider (A). Every possible maroon flower is a red flower, yet there are red flowers that are not maroon (crimson flowers, for example). Thus necessitation obtains non-symmetrically. In effect, redness is multiply realisable, supervening on the set of determinate properties. As such, the only way to intervene on redness is to intervene on a determinate property. Of course, not every way of intervening on a determinate property will change whether redness obtains (it will just change which shade obtains). But since the only way to intervene on redness is to intervene on a determinate colour-

⁷⁷ I am assuming, here and elsewhere, that the fact that the bicycle exists does not rigidly designate that particular bicycle.

property, the interventionist information will cue the presence of a non-symmetric instance of a modal relation.

In each case we can explain the apparent *asymmetry* by noting that each instance of the modal relation in question is a non-symmetric instance. Since it is non-symmetric instances of causation that the mechanism evolved to track, and since causation is an asymmetric relation, it can hardly be surprising that we mistakenly conclude that the relation we are tracking in these cases is itself asymmetric (when in fact it is non-symmetric, and not asymmetric).

What explains our belief that there is an explanatory relation obtaining between the relata is that the signal from the causal detection mechanism triggers, or is otherwise associated with, a phenomenology as of such an explanatory relation obtaining. It is associated with that phenomenology either because it directly causes it, or because what cues the signal is a common cause of the signal and the phenomenology. It makes good sense for the output of the mechanism to trigger that phenomenology because the mechanism evolved to track causal dependencies, and those dependencies are genuinely explanatory. But the signal triggers the phenomenology even when it is filtering non-diachronic correlations, and so we experience a phenomenology as of the obtaining of an explanation when the causal detection mechanism cues us to the presence of a non-symmetric instance of a relation.

What remains to be explained, then, are judgements (C) through (H). In what follows I argue that what explains the apparent asymmetry, and the phenomenology of explanation, is the same for these cases as in (A) and (B): the functioning of the causal detection mechanism. The difference lies in the fact that in cases (D) through (H) the mechanism *mistakenly* signals the presence of a non-symmetric instance of a relation, where no such instance obtains. In case (C), the mechanism correctly identifies a non-symmetric instance of a modal relation, but, as I will show, it does so for the wrong reasons. We arrive at these mistaken judgements because, as I will now argue, the causal detection mechanism overgeneralises.

4.6 A Filter that Overgeneralises

Step four of the grounding-free explanation appeals to empirical evidence to show that our causal detection mechanism sometimes *overgeneralises*. This can hardly be surprising. Evidence suggests that causal reasoning—that is, reasoning in terms of causal models—is typically very successful: it affords agents a good deal of predictive and explanatory power, and is often fast, automatic and unreflective (Sloman, 2005:77-78,80). Not only are subjects much more ready to accept an explanation if it fits with a causal model, but the best way to convince someone of something is actually to provide them with the resources to create their own causal explanation (Ross and Anderson, 1982). Moreover, it will come as no surprise that subjects have difficulty understanding a narrative in which events are inconsistent with their view of the causal structure of the world (Sloman, 2005:88). Plausibly, then, the use of causal models in reasoning is important, and according to some, fundamental to how we understand the world (Schafer, 1996).

Here is something we know. In general, we ought to expect that where the costs of a false negative (failing to detect some stimulus) significantly outweigh the costs of a false positive (signalling the existence of a stimulus where there is none), we typically develop cognitive systems that set the threshold for detecting that stimulus quite low.⁷⁸ Consider life as an animal that is predated upon. The cost of failing to detect a predator could well be death. The cost of misidentifying something as a predator is not nil (since you might run away, thus using up energy) but it is lower than failing to detect a predator. For this reason, prey animals have predator detection systems that have a very low threshold for detecting predators. That is, the features that something in the environment needs to have in order to set off the predator detection system are relatively minimal. In cases such as these I will say that the cognitive systems in question tend to *overgeneralise*.

There are plenty of examples of overgeneralisation in the human cognitive system. Detecting faces is important. So we have very sensitive facial detection systems that can be triggered by something as simple as an arrangement of three dots in a particular

⁷⁸ Notice that I am making a very general claim here: I am not defending anything like the thesis that we should expect cognitive systems to be optimal in the manner in which they make these trade-offs. Clearly there are developmental constraints on the ways in which cognitive systems can solve problems which mean that systems often are not optimal.

configuration (roughly a triangular configuration).⁷⁹ This is why Facebook is replete with images of ‘faces’ in the clouds, and on pieces of toast. Our facial detection mechanism rarely fails to detect a face, but will often give false positives where there are no faces. The same can be said, *mutatis mutandis* for our agency detection module (Guthrie, 1993).⁸⁰

Since our causal detection mechanism is such a powerful tool, I hypothesise that the benefits of sometimes signalling that a diachronic correlation is due to the presence of a non-symmetric instance of diachronic relation, when in fact it is not, outweigh the costs of sometimes failing to signal that the relation obtains, when in fact it does. Indeed, there is strong evidence that our causal detection mechanism does, indeed, overgeneralise in this way.

We sometimes impose causal structure where none exists. For instance, we seem to perceive causation where there is none. Subjects reliably describe the interaction of two moving dots on a screen in terms of one dot causing the other dot to act in certain ways (Michotte, 1956). People impose beliefs about the causal structure of the world onto the correlational data they are trying to understand (Waldman, 1996) even when imposing a causal framework distorts their representation of the world. For instance, mathematical equations are symmetric in the sense that any variable can appear on either side of an equality.⁸¹ Despite this, subjects find certain ways of expressing an equation more natural than others: namely, those ways that fit best with their causal model (Sloman, 2005:72). Further, if subjects are allowed to write equations in any form they like they will put the cause on the left hand side of the equality sign and the effect on the right hand side (Sloman, 2005:72). While this is merely an aesthetic preference, the fact that subjects are so disposed indicates that their causal model might be generating unwarranted judgements about asymmetries. At the very least, it shows that the way subjects conceptualise and understand their environment is significantly affected by the causal model they overlay onto the world. Similarly, it is argued that the overgeneralisation of our causal detection mechanism leads us to misjudge probabilities. For instance, suppose subjects are asked which of the following is more probable:

⁷⁹ Rigdon et al. (2009), for example, showed how such an arrangement of dots significantly increased participants’ giving behaviour. They hypothesised that this is because the dots are sufficiently face-like to cue people to act as though someone is watching to see how generous they are.

⁸⁰ Indeed, there are those who think that our tendency towards religious belief is to be explained by an overactive agency detection model which ‘detects’ agency where there is none, leading us to posit supernatural agents (Barrett, 2004).

⁸¹ Algebra is all about how to permute the order of the variables without changing the relations the equation expresses.

Case 1

- 1) A man has a history of domestic violence if his father has a history of domestic violence.
- 2) A man has a history of domestic violence if his son has a history of domestic violence.

Subjects report that (1) is more probable than (2) even though they are equally probable. It is thought that (1) appears more probable because the direction of causation goes from father to son, not son to father. The causal detection mechanism responds to certain cues present in the presentation of this data, and signals that these probabilities are asymmetric, when they are not.

It is also thought that conjunction fallacies might sometimes be due to the misapplication of causal models. Subjects are asked which is more probable:

Case 2

- 3) An adult chosen at random from the US population will die of a heart attack within the next 10 years.
- 4) An adult chosen at random from the US population will be obese and will die of a heart attack within the next 10 years.

Subjects report that (4) is more probable than (3) (Sloman, 2005:104). One explanation for this is that since obesity is a known cause of heart attacks, subjects mistakenly turn the statement into a causal claim and judge that someone randomly chosen will die of a heart attack because they are obese. To put it another way, the way the information is presented cues the causal detection mechanism, and it signals that what is in fact less likely, is more likely. It mistakenly interprets the question in terms of a causal model, yielding the incorrect answer.

What sorts of cues might lead our causal detection mechanism astray in these cases? Consider Case 1. Recall the four cues to which the mechanism is sensitive.

- (i) Temporal order
- (ii) The result of intervention
- (iii) Prior knowledge
- (iv) An existing hypothesis about causal (modal) structure.

Consider (i). The event of the son engaging in domestic violence occurs after the event of the father engaging in domestic violence (at least in most cases). So temporal order will cue the mechanism. Consider (ii). Evidence shows we are adept at doing counterfactual interventions (Sloman 2005:80). Were we to imagine performing an intervention in this case, we would likely conclude that if we want to change whether a son is violent, we intervene upon whether his father is, not *vice versa*. So interventionist information will cue the mechanism. Consider (iii): prior knowledge. Some people will know that one way in which people become abusive is by witnessing it in the home, and that prior knowledge might feed into (iv) and create an existing hypothesis about the causal structure—namely that the father’s being abusive causes the son to be abusive. All of these cues result in the causal detection mechanism signalling the presence of a non-symmetric instance of a relation—causation—and the salience of that instance overrides the symmetrical correlation that is relevant in making probability judgements, leading to mistaken judgements. Analogous considerations apply to Case 2.

Prima facie, then, if the causal detection mechanism overgeneralises in signalling the presence of non-symmetric instances of a diachronic relation, we should expect it to similarly overgeneralise in signalling the presence of non-symmetric instances of non-diachronic relations. The next step of my explanation builds on this idea by showing that there are certain environmental conditions under which we should expect the causal detection mechanism to signal the presence of a non-symmetric instance of a modal relation amongst the non-diachronic correlations whether or not such an instance obtains—and these conditions are precisely those associated with cases (C) through (H).

4.7 Fooling the Co-opted Causal Detection Mechanism

Cases (D) through (H) are ones in which a *symmetric* instance of a modal relation obtains between the relevant objects, properties, or facts. (D) claims that the existence of the singleton set containing Pythagoras depends upon the existence of Pythagoras.⁸² However, every world in which Pythagoras exists is a world in which {Pythagoras} exists, and *vice versa*. So the existence of Pythagoras necessitates the existence of {Pythagoras} and *vice versa*: they modally co-vary. (E) has a similar structure to (D). <Pythagoras exists> is true at precisely those worlds at which Pythagoras exists, and thus

⁸² This kind of claim tends to be defended by noting that on the iterative conception of sets, they are *generated* from their members (see, e.g., Schaffer, 2009).

necessitation holds symmetrically. (F) stems from Plato's *Euthyphro*.⁸³ It is assumed that, for any possible action, if that action is good then God loves it, and if God loves it then it is good. Thus, the obtaining of the relationship between God's attitude towards the action and the value of the action is symmetrical, insofar as supervenience and necessitation are concerned.

(G) describes the relationship between the fact that Pythagoras exists, and the fact that the fact that Pythagoras exists, obtains.⁸⁴ Again, these facts obtain at all and only the same worlds, so the relevant modal relation obtains symmetrically. Finally, if numbers exist, then presumably they exist of necessity, and likewise for mathematical facts.⁸⁵ As these facts obtain at every world, the necessitation relations between them will obtain symmetrically. The existence of 2 necessitates the existence of $[2+2=4]$, and $[2+2=4]$ necessitates the existence of 2.

In all these cases, though we are inclined to say that an asymmetric explanatory relation obtains, it cannot be that we are correctly tracking non-symmetric instances of modal relations, since no such instances obtain. By contrast, in case (C) the relevant modal relation obtains non-symmetrically. Every world in which Pythagoras exists is a world in which \langle a man exists \rangle is true. Yet there are worlds in which Pythagoras does not exist, but \langle a man exists \rangle is true nonetheless, due to the presence of some other man. However, it doesn't seem right to say that we are 'successfully tracking' this non-symmetrical instance, since (C) through (H) have features that we should expect to cue the causal detection mechanism and result in it signalling the presence of a non-symmetric instance of a modal relation *whether or not one obtains*. To see why this is, let us go through each case.

Consider again cases (C), (D) and (E).

- C. The proposition \langle a man exists \rangle is true *because* Pythagoras exists
- D. $\{\text{Pythagoras}\}$ exists *because* Pythagoras exists
- E. The proposition \langle Pythagoras exists \rangle is true *because* Pythagoras exists

⁸³ Plato (translation published 2002).

⁸⁴ This case is adapted from Trogdon (2013a), who in turn adapted it from an example employed by Rodriguez-Pereyra (2005). Trogdon claims that $[\text{this fact obtains}]$ *grounds* $[[\text{this fact obtains}] \text{ obtains}]$ as a way of avoiding the possibility of mutual grounding between the two facts.

⁸⁵ Though see Colyvan (2000, 2001) for a view to the contrary. See Miller (2012) for a response.

Now consider the kinds of cues that can trigger the filtering system. (i), temporal order, is irrelevant. So consider (ii): The result of intervention. We have already seen that the result of actual (or counterfactual) interventions can help trigger the causal detection mechanism (recall the case of the chair and its parts). Interventions on non-diachronic correlations are, however, not always straightforward. With respect to some of those correlations, one of the correlated properties, facts or objects, cannot be intervened upon. This is what we find in cases (C) through (E), where one relatum is an abstract object, and is, therefore, an object we cannot intervene on.

Let's reflect on how interventions work to provide cues to the causal detection mechanism. In the case of diachronic correlations between, say, events of kind x and kind y, we are able, on one occasion, to intervene on a token x and see whether this wiggles a token y, and on another occasion to wiggle a token y, and see if this wiggles a token x. It is the result of this pair of interventions that acts as a cue for the causal detection mechanism. In particular, if in wiggling x we can wiggle y, but not *vice versa*, this tends to cue the mechanism to signal that there is an underlying non-symmetric instance of a relation.

In cases (C) through (E) it is clear enough that we can intervene upon whether Pythagoras exists, and thereby wiggle whether {Pythagoras} exists or <Pythagoras exists> is true or <a man exists> is true.⁸⁶ However, it is false that we can intervene upon whether {Pythagoras} exists or <Pythagoras exists> is true or <a man exists> is true and thereby wiggle whether Pythagoras exists. This is because propositions and sets are abstract objects upon which we cannot intervene. We can't wiggle whether {Pythagoras} exists in order to wiggle whether Pythagoras exists because we can't wiggle whether {Pythagoras} exists at all! Likewise for the true proposition in cases (C) and (E). We cannot directly intervene on the alethic value of a proposition. How could we? Rather, we intervene on what is true and what is false *indirectly*, by changing the nature of the world. Thus, even though, *were we* to directly intervene on these abstract objects, we would see the resulting wiggles in whether Pythagoras exists, we never get this feedback, for we simply cannot perform the requisite interventions. In this way, the causal detection mechanism gets the same non-symmetric feedback about interventions that it does when there are causal relations underpinning a diachronic correlation, but for a very different reason.

⁸⁶ Not every intervention on Pythagoras will wiggle <a man exists> but one way to wiggle <a man exists> is to make it the case that Pythagoras exists.

Causal relations do not obtain between abstract relata, and thus the causal detection mechanism has not evolved to distinguish between a situation where wiggling y does not wiggle x and a situation where y simply cannot be wiggled. Thus, triggered by an apparent non-symmetry in the results of counterfactual interventions, the mechanism signals that the correlations in these cases are indicative of an underlying non-symmetric instance of a modal relation. In cases (D) and (E) the signal is mistaken: the modal relation obtains symmetrically. Fortuitously, in (C) it correctly signals the presence of a non-symmetric instance of a modal relation. It gets things right, in that case, but for the wrong reasons: it would signal the presence of non-symmetry even if no such non-symmetry were present. For that reason I will not describe this as a case of the mechanism successfully tracking a non-symmetric instance of a modal relation.

Let us now consider (F), the Euthyphro case:

F. God loves X *because* X is good

We are asked to imagine that God and goodness co-exist in all worlds, and to wonder whether things are good because God loves them, or he loves them because they are good. This is a case in which there are no cues of kind (i) or (ii) (there are no cues of kind (ii) because we cannot intervene on *either* relatum. We cannot wiggle what in the world is good, and we surely cannot wiggle God's attitudes). But consider (iii) and (iv):

(iii) Prior knowledge

(iv) An existing hypothesis about modal structure

Both of these kinds of cues could be expected to play a role in determining the output of the causal detection mechanism. Plausibly, we should expect prior causal knowledge to influence a subject's existing hypothesis about modal structure. Each of us is familiar with agents' intentional states depending on the way the world is. (We typically hope our beliefs are like this). If this is the most salient piece of a subject's prior knowledge, it might lead her to have a hypothesis about modal structure, according to which just as the causal direction of fit goes from the world, to the mind so too does the modal direction of fit. Subjects will have a prior hypothesis that God's attitudes depend on the distribution of goodness, and in the absence of any other cue to the contrary, this will cue the causal detection mechanism to signal the presence of a non-symmetric instance of a modal relation.

On the other hand, subjects are also familiar with cases in which agent's intentional states cause the world to be a certain way (we typically hope our desires are like this, when appropriately conjoined with our beliefs). Where *this* causal knowledge is most salient, we can expect it to lead to a hypothesis about modal structure according to which, just as the direction of causal fit goes from God's attitudes to goodness, so too does the direction of modal fit. In the absence of any other cues to the contrary, this hypothesis about modal structure also cues the causal detection mechanism to signal the presence of a non-symmetric instance of a modal relation.

As I noted in Chapter 2 (see footnote 23), judgements about this case vary. This is an awkward datum for the defender of the grounding-based explanation to accommodate, for she must contend that at most one participant in the Euthyphro dialogue is correct: the participant whose intuitions align with the grounding relations. The current explanation, however, serves to account for why there is more divergence in people's intuitions about this case than the prior cases. As neither (i) nor (ii) offer us any useful cues, it could have been predicted that people would form very different modal hypotheses given different salient background causal knowledge, when confronted with the description of the case.

Now consider (H):

H. $2+2=4$ *because* 2 exists and 4 exists

This is another case in which we cannot intervene on either relatum. So, once more, (i) and (ii) provide no helpful cues. But, again, we might expect prior knowledge to lead to the triggering of the causal detection mechanism. We have prior knowledge of relations and relata, itself stemming from experience with interventions. We typically know (even if only implicitly) that in order to intervene on some relation, we need to intervene on the relata. If we want to intervene on the relation of 'being next to' obtaining between Bill and Ben, we can only do this by moving either Bill, or Ben. We can't directly intervene on the 'being next to' relation. So prior knowledge tells us that, in general, if you want to intervene on the obtaining of a relation, then you need to intervene on the relata. Addition is a relation. So prior knowledge suggests that in order to intervene on whether it obtains between 2 and 2, one would need to intervene on the relata of the addition function. Thus in the absence of being able to discern any other cues, this prior knowledge leads to an existing hypothesis about modal structure which cues the causal

detection mechanism, which indicates that there is more than mere correlation here, and that (H) is true.

This cue, however, is relatively weak. After all, cues of kind (i) and (ii) are absent. So it is relatively easily overturned. We can witness that, by the fact that mathematical structuralists find (h), rather than (H), intuitive.⁸⁷ Their prior knowledge has altered the cues that the causal detection mechanism receives, and changed the output so that although it signals the presence of a non-symmetric instance of a modal relation, the ‘direction’ of the non-symmetry is taken to be the reverse of (H).

Moving on, let’s revisit the relationship between a fact and the fact that that fact obtains:

G. [Pythagoras exists] obtains *because* Pythagoras exists⁸⁸

Insofar as one is prepared to believe that facts about facts obtaining exist, there is an intuitive pull to suppose that one would intervene on the fact on the right, in order to intervene on the fact on the left, but not *vice versa*. Here is why we might expect to have that intuition. The fact on the right is a *constituent* of the fact on the left. Talk about facts is sophisticated, and almost certainly something we come to after we have learned about more mundane things such as, saliently, parts and wholes. Prior knowledge of wholes and parts tells us that we intervene on the parts to intervene on the whole. The logical form, if you will, of (G) strongly suggests that we ought to intervene on the fact on the right, in order to intervene on the fact on the left. Since we have little other experience at intervening on facts, the logical form of (G) is a salient cue, which leads us to make a particular hypothesis about modal structure, and ultimately cues the causal detection mechanism to categorise (G) as more than mere correlation.⁸⁹

4.8 Summing Up the Grounding-free Explanation

So far, I have shown why we should expect our cognitive apparatus to signal the presence of a non-symmetric instance of a modal relation when presented with non-

⁸⁷ Mark Colyvan (personal communication).

⁸⁸ I take the expressions flanking the ‘because’ here to pick out facts. This can be made clear by expressing what I take to be the equivalent claim: [[Pythagoras exists] obtains] *because* [Pythagoras exists].

⁸⁹ Similar considerations apply to the putative explanation of a ‘conjunctive fact’ in terms of its conjuncts (see Raven, 2012). That is, prior knowledge of wholes and parts tells us that we should intervene upon a conjunct—a ‘part’—in order to intervene on the conjunction—the ‘whole’.

diachronic correlations in certain conditions; those conditions we find obtaining between the objects, properties, and facts, mentioned in priority intuitions (A) through (H). In cases (A) and (B) this is (in part) because interventionist feedback cues the mechanism to correctly signal that there is a non-symmetric instance of a traditional modal relation. In cases (C) through (H) by contrast, we should expect our causal detection mechanism to signal the presence of non-symmetry whether or not it obtains (and it does obtain, but only in (C)). On the assumption that these non-symmetric instances do not obtain, however, this overgeneralisation is largely cost free in the following sense: feedback that can be gained via interventions cannot reveal that the system is generating false judgements because the relevant interventions that would reveal this cannot be performed. Thus the environment can never provide feedback that would allow us to correct that judgement.

So there is, effectively, no cost associated with the causal detection mechanism erroneously being triggered by these correlations. Thus the appeal to the sorts of cues the environment sends the causal detection mechanism has a dual role. On the one hand, it explains why we should expect to have the intuitions we do, regardless of whether there are non-symmetric instances of modal relations underpinning these correlations. But examination of those cues also reveals why there is no cost to being misled in this way, on the assumption that we are, indeed, being misled (that is, on the assumption that there is no grounding relation that we are tracking): namely, that we should also expect this overgeneralisation in such cases, because there is simply no cost to being wrong.⁹⁰

In sum, I hope to have accounted for the priority intuitions via a theory that appeals, in an indispensable manner, to the existence of the traditional modal relations and certain psychological mechanisms, but does *not* appeal, in an indispensable manner, to the existence of any relation of ground. Furthermore, this explanation exemplifies the theoretical virtues.

⁹⁰ One might worry that, unlike other cases where the causal detection mechanism overgeneralises, our grounding observations in cases (D) – (H) are highly resistant to change on the basis of reflection. So, for example, making salient the probabilistic symmetry between abusive fathers and sons effectively constrains our tendency to overgeneralise. One difference is that in my explanation of (D) – (H), there are not two relations, one symmetric and one non-symmetric, such that the salience of one swamps the salience of the other. Instead, a single, symmetric relation, is misclassified by a cognitive system as non-symmetric. We know that the output of some sub-personal cognitive systems are (largely) immune to change on the basis of personal-level reflection. No amount of reasoning makes the lines look the same length in the Muller-Lyre illusion. I think the outputs of the causal detection mechanism are like this (or at least, lie towards this end of the spectrum).

Firstly, the Grounding-free Theory is straightforwardly more parsimonious than the Sophisticated Grounding Theory, simply because it refrains from positing any primitive ontology. It might be thought that my explanation of the priority intuitions is more complex (less elegant) than the explanation offered by the Sophisticated Grounding Theory. In response to this, I remind the reader that the Sophisticated Grounding Theorist will (as I hinted in Chapter 3) likely wish to help herself to a story very much like mine in order to provide an epistemology of grounding (that is, *if* she thinks for some reason that the cognitive mechanisms described here are well suited to track grounding relations, if there are any). Regardless, she will need to appeal to some such story to explain the epistemology of grounding. Thus, she is saddled with the same complexity on this front also.

The defender of either of the grounding-based theories can, plausibly, insist that she has an advantage when it comes to providing a simple *vindication* of our intuitions, but as I demonstrated in Chapter 3, there are strong reasons to dislike the grounding-based theories of metaphysical explanation. Ultimately, the complexity of the machinery used here is complexity that the grounding-based theorist is, presumably, committed to in order to explain our judgements of causal dependence. In other words, the correlation and causal detection mechanisms are posited by the Background Theory, and thus all the theories considered here are committed to their existence. Thus, the complexity of the overall theory is not increased by using this machinery to also account for the priority intuitions. Hence, I contend that the Grounding-free Theory is more elegant than the Sophisticated Grounding Theory, which is saddled with the ideology that governs grounding.

The explanation I offer here is, in several respects, virtuously unified. Firstly, it unifies the explanations of our intuitions of both diachronic and non-diachronic priority, as they are explained by the very same cognitive mechanisms. Secondly, it is unified with the general strategy of appealing to evolved mechanisms to explain animals' beliefs and behaviours, which does a good deal of explanatory work in a lot of domains. Finally, it is unified with a large class of evolutionary explanations. However, insofar as the Sophisticated Grounding Theorist uses a story like mine, her explanation also exemplifies this virtue.

I foresee an objection to the claim that my explanation is as unified as my

opponent's. According to that objection, my explanation is less unified because it appeals to the obtaining of a non-symmetric modal relation in the explanation of our intuitions in cases (A) and (B), whereas it appeals to the existence of an overgeneralising causal detection mechanism to do the explanatory work in cases (C) through (H). To this, I reply that it is the *same* pair of mechanisms doing the work to explain our intuitions in *all* of the exemplar cases. The difference between the cases is that, in cases (A) and (B), both relata are such that we can intervene upon them, whereas in cases (C) through (H), there is at least one relatum upon which we cannot intervene. To be clear, then, I do not claim that the non-symmetry of the modal relation directly explains our intuitions about cases (A) and (B). Rather, it is our cognitive mechanisms that explain our intuitions in all cases, and in cases like (A) and (B) there are further reasons to think that our intuitions are successfully tracking non-symmetrical relations. Similar thoughts apply to case (C), where we end up making what I take to be a true judgement (that there is a non-symmetry at play), but our mechanisms deliver this result for the wrong reasons. Thus, I take the virtue of unification to tell in favour of my psychological explanation, deployed in the service of the Sophisticated Grounding Theory *or* the Grounding-free Theory.

Returning to the master argument, I hope to have achieved two goals in this chapter. Firstly, I hope to have achieved the crucial initial step of establishing that the Grounding-free Theory is empirically equivalent to the Sophisticated Grounding Theory. For, the explanation offered here can account for our making the judgements we observe ourselves to make. Thus, using Colyvan's (1999) framework for determining whether some ontological posit is indispensable to the best explanation of our observations, the comparison between the theories must be framed in terms of the theoretical virtues. Secondly, as the theoretical virtues tell in favour of my explanation over my opponent's, grounding is not indispensable to the best explanation of our observations. Thus, for those who care little about metaphysical explanation, my argument ends here.

However, my explanation lacks some features that many will deem desirable. For, as I flagged earlier, my explanation—framed as it is in terms of *overgeneralisation*, of being *misled*, of *erroneous* judgements—is an *explaining away* of the priority intuitions, rather than a vindication thereof. Now, the reader who finds the explanation offered here to be compelling might find it acceptable that many claims of the form 'x *because* y' that she intuits to be true turn out, in fact, to be false (or at least, not true). For, if this is the correct explanation of our intuitions, it really does seem that we are projecting non-

symmetries onto a world where many modal relations are symmetrical. If one has the intuition that we must be tracking something ontologically robust (like grounding) for our priority intuitions to be vindicated, then the present story cannot provide the (possibly desired) vindication.

Nevertheless, if we think of the priority intuitions as expressing propositions about what *explains* what, and we think of explanation as an epistemic phenomenon, then the explanatory story told here can, in fact, help to vindicate these intuitions. This can be done by developing a psychologicistic theory of metaphysical explanation, and an attendant relativist semantics. If the truth-conditions for claims of the form ‘x *because* y’ don’t require that there is an asymmetric dependence relation between x and y, but rather, that creatures like us find y to be a good explanation of x, many of these claims will be true. The building of such a theory, and the exploration of several ways of tightening this loose idea, is the focus of Chapter 5.

After I develop the psychologicistic theory, the thesis concludes with an exploration of some metaphysical variants of the DN theory of scientific explanation. I contend that those seeking a theory of metaphysical explanation that satisfies the ontic desideratum of objectivity should explore these avenues, rather than the suspicious grounding-based theories of metaphysical explanation.

Chapter 5: The Filtered Modal Relations Theory of Metaphysical Explanation

It is orthodoxy to suppose that grounding relations are indispensably required to account for our priority intuitions. However, in Chapter 4 I developed an explanation of this observation in terms of the functioning of a pair of evolved psychological mechanisms. My preferred explanation—the grounding-free explanation offered by the Grounding-free Theory—is empirically on a par with the grounding-based explanation offered by the defender of the Sophisticated Grounding Theory. Moreover, this grounding-free explanation is more theoretically virtuous than its rival, for it is more parsimonious and elegant. Thus, according to Colyvan’s (1999) criteria for establishing the indispensability of some entity, grounding relations are dispensable and should be dispensed with.

However, as I noted in §4.8, the explanation in terms of psychological mechanisms is more of an *explaining away* of our priority intuitions than a vindication thereof. Furthermore, I take it that failing to vindicate said intuitions will, for some readers, amount to a failure of the explanation. Likewise, the lack of a theory of metaphysical explanation will be seen by many as a substantial drawback of my explanation. Thus, in the remaining chapters I develop a series of grounding-free theories of metaphysical explanation which might accompany the Grounding-free Theory. Each of these serves to vindicate the priority intuitions in a different way. Which theory one should adopt, of those I develop, will depend on how one weights the desiderata for a theory of metaphysical explanation, as well as one’s prior metaphysical commitments.

In this chapter I develop a (psychologically) filtered modal relations theory of metaphysical explanation. According to this theory, it is the modal relations of necessitation and supervenience which we track with our metaphysical explanations, but only those instances about which our causal detection mechanism generates intuitions of priority count as metaphysical explanations. The theory is explicitly relativistic, such that relative to different communities, with different psychological apparatuses, different instances of the modal relations will count as explanatory (different claims of the form ‘*x because y*’ will be true). It is this theory to which I unfavourably compared the filtered grounding-based theory (§3.5). The present view will be attractive to those who think of explanation as an epistemic notion, particularly those who already endorse a filtered causal process theory in the context of scientific explanation (of which my view is a

natural extension).

Here is the plan. §5.1 will introduce the filtered modal relations theory in opposition to the grounding-based theories. §5.2 makes a general case for the plausibility of a psychologistic and contextualist account of metaphysical explanation. §5.3 then lays out the specifics of the account and defends my preferred semantics. To show that my account is a *plausible* alternative to grounding-based theories I need to show both that it assigns propositions the correct truth-values and that the truth conditions it proffers for propositions of the form ‘*x because y*’ are plausible. §5.4 considers the indeterminacy that results from my preferred theory. §5.5 will conclude by evaluating the filtered modal relations theory against the desiderata, and showing how this account can be unified with a filtered causal process theory to develop a theoretically virtuous overall theory.

5.1 Grounding-based versus Grounding-free Theories

As I explored in Chapter 3, the orthodox view in the contemporary literature is that ‘*x because y*’ is, if true, true in virtue of the obtaining of a *grounding* relation between the referent of ‘*y*’ and the referent of ‘*x*’. These are the grounding-based theories of metaphysical explanation (see §3.2.3; §3.4; §3.5). Grounding-based theories of metaphysical explanation thus combine two claims: (a) that the only plausible candidate that could make true our claims about metaphysical explanation is the existence of grounding relations and (b) grounding relations exist and do make true at least some of those claims. On the unfiltered view, which metaphysical explanations are true and which are false is purely a matter of putatively mind-independent facts: in this case, facts about what grounds what. On the filtered view, only those grounding relations with the right epistemic character back metaphysical explanations.

I have also argued against positing grounding, on the basis that grounding is not indispensable to the best explanation of our observations. As the Sophisticated Grounding Theory is empirically equivalent to, and less theoretically virtuous than, the Grounding-free Theory, we should, according to Colyvan’s (1999) criteria, be grounding sceptics. But if we are sceptics about grounding, and if grounding-based theories of metaphysical explanation are the only game in town then it looks like we should embrace an error theory about metaphysical explanation: sentences of the form ‘*x because y*’ are

never true because what is required to make them true—grounding relations—fails to obtain (see §5.1.1).

However, grounding-based theories are not the only game in town. I reject the contention that the only plausible candidate to make true our claims about metaphysical explanation is the existence of grounding relations. Chris Daly, a grounding sceptic, notes that “if we chose to talk of metaphysical explanation, it is open to the sceptic to understand such talk in other ways, such as in terms of supervenience or metaphysical necessitation, thereby creating a fissure between talk of grounding and talk of metaphysical explanation.” (2012:94). Here and in the following chapters, I explore various ways of implementing Daly’s suggestion of understanding such talk in other ways.

I suspect that, while some readers who accept arguments against positing grounding will respond by becoming error theorists and abandoning talk of metaphysical explanation, others will desire the continuation of such talk. That’s because such talk is, I suspect, sufficiently widespread and useful as to be highly resistant to elimination (See §5.1.1). This, in turn, is some evidence that the truth conditions for ‘x *because* y’ are not as tied up with grounding as grounding-based theories would have us believe, and lends plausibility to there being *some* alternative account. As we will see, the alternative account presented here is one that appeals to, *inter alia*, psychological facts.

In what follows, I will show that sceptics about grounding can also tell a plausible story which judges intuitively true claims of the form ‘x *because* y’ to be true, and intuitively false such claims of the same form to be false. To use the terminology introduced in §2.5, the view I propose in this chapter is that the candidate relations we track with our metaphysical explanations are the traditional modal relations of necessitation and supervenience. We briefly considered the unfiltered version of such a view in §2.2, and it was found to fall short of several desiderata. However, the *filtered* version of this view remains a strong contender. It will give up on the ontic desideratum—the objectivity of what metaphysically explains what—in favour of securing all of the other desiderata.

Of course, it is incumbent on those who defend a filtered theory of explanation to say something about the filtration to which they appeal. The view I offer here is that the instances of the modal relations are filtered by psychological facts about communities,

such that only if certain psychological facts obtain does a particular instance of a modal relation count as explanatory. Of course, what I have said here leaves open which particular psychological facts are those that matter when determining the truth or falsity of an utterance of ‘x *because* y’. In §5.3 I will consider several proposals regarding *whose* psychological states should enter into the truth-conditions for such utterances. This will be done within the framework of a relativist semantics.

So, the story is, roughly, that metaphysical explanations are true, in part, in virtue of psychological facts about the agents who utter them. In other words, I offer a *psychologistic* theory of metaphysical explanation, whereby what makes true claims of the form ‘x *because* y’ are the obtaining of traditional modal relations (necessitation and supervenience) alongside dispositions to have certain mental states. Since these mental states can be expected to vary across individuals, on this view claims about what metaphysically explains what are not true or false *simpliciter*, but rather, relative to a context of assessment. Thus, while I am engaged in the same project as those who defend grounding-based theories (insofar as I am seeking to identify which states of the world serve to make true our claims about what metaphysically explains what), I abandon the idea that said truthmakers are mind-independent.

Recall that Chapter 4 provided an empirically supported, grounding-free explanation of how we come to have the psychological state of intuiting that ‘x *because* y’. The story told here builds upon that explanation, for I suppose that it is the outputs of our cognitive mechanisms that constitute the relevant psychological states that filter the modal relations when contemporary metaphysicians utter metaphysical explanations. However, I want to flag that for our current purposes it doesn’t matter which cognitive mechanisms actually do the work of producing the relevant mental states, as long as those states are produced.

What *does* matter is that the current account is truly an *alternative* to grounding-based theories, and that will be true so long as it is plausible that we would have the relevant mental states in the absence of grounding relations. For, if the best explanation for our having those mental states is that we are good trackers of what grounds what, then my account is parasitic on the existence of grounding relations and thus not a genuinely grounding-free theory of metaphysical explanation. In sum, I shall continue on the assumption that the story told in Chapter 4 is true, but the view here is compatible

with an alternative (though grounding-free) psychological story better capturing why we have the intuitions that we do.

5.1.1 Why isn't this an Error Theory?

Chapter 4 defends the claim that it is plausible that we would have priority intuitions even in the absence of the existence of relations of ground. Yet, as I mentioned in §4.8, the account of the cognitive mechanisms involved in producing the relevant mental states makes mention of *overgeneralisation* of a mechanism that signals the presence of asymmetric relations where there are only symmetric instances of non-symmetric relations. One might object that such an account motivates an error theory more than it motivates a psychologically filtered theory. After all, the suggestion is that our mental states are systematically in error.

In what follows, I argue that a psychologically filtered theory of metaphysical explanation will be appealing insofar as one finds it plausible that there is something subjective, or psychologically dependent, about what makes true, claims of the form 'x *because* y'. More generally, any plausible account of the conditions under which claims of the form 'x *because* y' are true needs not only to correctly assign truth-values to instances of the schema, but in addition, those truth conditions need to be plausible. But, proceeds the current objection, if what I say in Chapter 4 is right, then the filtered account does not offer plausible truth conditions. Rather, our cognitive machinery massively misleads us: and if that is the way things are, then we should conclude that any instance of 'x *because* y' is false. So, why should we seek vindication of our intuitions rather than embracing an error theory?

I am not unmoved by this thought. Perhaps accepting the explanation provided in Chapter 4 should constitute the end of this enquiry, and seeking some way of vindicating our intuitions is a mistake. If that's right, I have already achieved the overarching goal of the thesis, for there will be no remaining reasons to posit grounding. However, I intend to show, more ambitiously, that *even those who still seek vindication* can have what they desire within the grounding-free framework. For, I think it is an empirical matter whether, if what I say in Chapter 4 is true, we will be error theorists about metaphysical explanation. That is because I think that whether or not we will be error theorists about some discourse (in this case the metaphysical explanation discourse) is settled by facts about

the content of our concepts and the meanings of our terms, and those facts are, themselves, empirical facts.

So, for instance, I think it is an empirical matter whether we ought to be error theorists about witches, because I think it is an empirical matter both (a) what the content of our concept of witch is, and (b) whether given that content, there is anything that satisfies that concept. Since I also think that the content of our concepts is discoverable (at least in principle) by what we are disposed to say and do when we make certain discoveries about the world, I think we can empirically discover whether or not we will be witch error theorists by considering what we are disposed to say if, for instance, we discover that there are no supernatural powers, no devil with which one can have carnal relations, and so on.⁹¹

By parity, then, in the case under consideration I think that we have evidence about whether or not error theory is true about metaphysical explanation in the form of what we (each of us singly, and collectively as a community) are disposed to say about the presence or absence of metaphysical explanation, under the supposition that the mechanistic story I told in Chapter 4 is correct. If we, as a community, are disposed to say that some instances of ‘*x because y*’ are true, on the supposition that the explanation I proffered in Chapter 4 is correct, then we have evidence that the explanation I offered is not one that will lead to error theory. Whether or not we are so disposed, as a community, is, of course, an empirical matter.

Why think that we will be disposed to endorse an error theory of metaphysical explanation, given the truth of the explanation I provided in Chapter 4? I suppose that, in part, what might so dispose us is the thought that although explanations are rightly about increasing agents’ understanding, in this case it turns out that there *is* no increase in understanding: given the way we form our beliefs about what metaphysically explains what, there is merely an appearance *as of* an increase in understanding. And that’s just not good enough to make true any instance of ‘*x because y*’.

Why think that we will be disposed to continue to judge that some instances of ‘*x because y*’ are true? Perhaps what might so dispose us is the thought that increasing understanding is, at least in part, a matter of increasing our capacity to manipulate the world. In worlds where agents can directly manipulate abstract, but not concrete, objects,

⁹¹ Of course, some of this is controversial. Not everyone thinks that we have even defeasible access to the content of (some of) our concepts. This is not a claim I can defend here.

their (very different) causal detection mechanisms can be expected to signal the presence of an asymmetry that runs the opposite direction to the asymmetry our mechanisms signal. We, and these agents, will disagree (sometimes faultlessly, in my view) about what metaphysically explains what. But the *understanding* that each agent gets via the mechanisms in question is real: each is able to manipulate the world in a way she would not otherwise have been able, in the absence of detecting the correlation, and her understanding of the direction of manipulability is real and important. The fact that it is not grounded (pardon the pun) in any ontological asymmetry does not matter. So we might be inclined to say that there *is* genuine understanding that comes from the functioning of the relevant mechanisms.

To be sure, God, for instance, would understand things quite differently: God would see that in many of these cases there is a symmetric instance of a non-symmetric relation obtaining and at some level God's understanding is clearly superior. Nevertheless, understanding is a perspectival matter: something can genuinely increase understanding despite not being part of what is understood from the omniscient and non-perspectival God-like perspective. And if that is the right way to understand understanding, then it is plausible that, as a community, we may be disposed to think that even if what I say in Chapter 4 is true, nevertheless the truth conditions the psychologistic account offers are plausible: there are metaphysical explanations after all.

In any case, my argument is conditional. I only contend that *if*, despite the nature of the grounding-free explanation of our priority intuitions, we still seek vindication rather than an error theory, such vindication can be achieved.

5.2 Motivating a Psychologistic Theory

As noted, in §5.3 I will develop several semantic proposals for claims of the form 'x *because* y', in order to make precise several ways in which the modal relations might be psychologically filtered. But what is the connection between the metaphysical story (whereby there are candidate modal relations that back metaphysical explanations, and we filter out the non-explanatory instances of such relations by appealing to psychological facts) and the semantic proposals?

The metaphysical story is intended to parallel a filtered causal process theory of scientific explanations. Indeed, the parallel is so close that I conclude this chapter by proposing a unified filtered theory of scientific and metaphysical explanation. But merely knowing that psychological facts serve as part of what makes true our claims about metaphysical explanation very much underdetermines the nature of the theory. *Which*—and indeed *whose*—psychological states matter, on this picture? I will use the development of the semantic proposals as a vehicle through which to propose and evaluate several answers to this question. While the discussion will be framed in terms of semantics, each semantic proposal corresponds with a way of thinking about the metaphysics—with a way of thinking about which psychological states matter when we are trying to determine what explains what.

In particular, if one thinks that a particular semantic proposal provides intuitive verdicts regarding the truth and falsity of various putative metaphysical explanations, that will provide a reason to endorse the corresponding metaphysical picture. By contrast, if one finds a semantic proposal to provide counterintuitive truth-conditions, this constitutes a reason to refrain from endorsing the corresponding metaphysical picture.

This is all to say that, while I speak as though there is a monolithic psychologicistic filtered modal relations theory, there is actually a class of such theories, differentiated by which psychological facts they appeal to when filtering the modal relations. The semantic proposals, then, serve as ways of comparing the plausibility of the various incarnations of the filtered modal relations theory. Before getting into the specific proposals, however, I want to say a little about the relativistic nature of the theories that emerge when we start to appeal to psychological considerations in our theory.

Appealing to psychologicistic filtering in a theory of metaphysical explanation places such a theory in the company of psychologicistically filtered theories of scientific explanation, and similarly subjectivist theories such as those of taste (gastronomic), aesthetics, epistemic modals and ethics. Of course, these latter views are not naturally thought of as a kind of filtering. I don't mean to suggest that there are a bunch of aesthetic relations, moral relations, etc., that are then filtered according to some individual's (or community's) psychological orientation to isolate those instances of these relations that are fit to make true moral or aesthetic claims.

Rather, the filtered modal relations theory is in the company of such views insofar as in each of these domains some form of realism⁹² is defended, but the truthmakers for the relevant claims within the discourse include psychological facts (broadly construed). As I discuss below, that there is a relativistic element is not equally obvious in each of these discourses. For example, imagine someone who unreflectively endorses a filtered causal process theory. She supposes that it is causal relations that back scientific explanations, when appropriately filtered according to some psychological properties we (humans) all share. Such an individual might not think that the truth of claims of scientific explanation is *relative*, for she might not have thought about what to say about agents with dissimilar psychologies and intuitions. In other words, if the relevant psychological states of all the agents one cares about (all the agents around here, say), are the same, one might not be motivated to develop a relativist semantics, because the same claims are true for all of us. Yet, to be upfront, we should notice there is an implicit relativity to this view, for if human psychology determines whether explanatory claims are true when uttered by any possible agent, with any psychological orientation, the resulting view is, by my lights at least, unattractively chauvinistic.

By contrast, in the other domains I mentioned, it is more obvious that there is a relativistic element, for there is interpersonal variation (as opposed to the possible *inter-species* variation alluded to above, which is easy to forget about). I think that the case of metaphysical explanation falls somewhere between these two extremes, for there are quite clearly modal relations that we are tracking with such explanations (like scientific explanations, but unlike aesthetic judgements), but there is also some disagreement in the actual human community (the Euthyphro dialogue, for example, and the debate between mathematical structuralists and certain mathematical Platonists, for another). There is enough variation in our intuitions about metaphysical explanation to suspect that agents that are very different to us could have very different intuitions indeed, and insofar as we want our theory to cover claims made by such agents, our theory should be a relativistic one.

For, there seem to be three general reasons to endorse subjectivist or relativistic views, and each applies to metaphysical explanation. First, one might think it is obvious, or at least plausible, that there is something subjective, or psychologically dependent, about what makes true certain utterances in the relevant discourses. Second, one might

⁹² At least, realism in the sense that the relevant claims are taken to be truth-apt and sometimes true.

think that the relevant claims are truth-apt, and sometimes true (one rejects error theory) but be sceptical of the existence of some entirely mind-independent fact that could do the truthmaking work. Third, one might find it plausible that there are sometimes faultless disagreements between parties, both of whom are speaking truly. All three considerations militate in favour of the truthmakers in question being at least partially psychological in nature, and the third also militates in favour of some sort of contextualist or relativist semantics according to which parties that are (or appear to be) disagreeing, can both be speaking truly. Ultimately, I think that these three general reasons also militate in favour of a psychologically filtered theory of metaphysical explanation. Let us consider each in turn.

As noted above, I do not think that claims of the form ‘*x because y*’ seem inherently subjective in the way that, to some at least, claims of the form ‘*x is tasty*’ seem inherently subjective. But I do think that the truth of such claims strike many as psychologically dependent: that is, it seems plausible, to many, that part of what it is for one thing to explain another, is for some (relevant) agent to find the explanans to increase her understanding of the explanandum (perhaps in the sense clarified in §5.1.1: she gains power that enables her to manipulate the world in certain ways). If that is right, then what explains what (in both the scientific and metaphysical domains) is, in part, dependent on psychological facts.

The second reason to think that ‘*x because y*’ is like ‘*x is tasty*’ or ‘*x is beautiful*’ or ‘*x is wrong*’ is that in all these cases the claims seem to be truth-apt, and sometimes true, but one is sceptical of the existence of some entirely mind-independent fact that could do the truthmaking work. Now of course, not everyone is sceptical. Those who defend an unfiltered grounding-based theory of metaphysical explanation think there are such mind-independent facts, just as some moral realists think there are mind-independent moral facts. But I have already argued against grounding; I am offering this account to those already sceptical of grounding, yet still interested in vindicating the priority intuitions and developing a theory of metaphysical explanation. And such folk are precisely those who are likely to be sceptical that there exist any mind-independent facts that could make true our claims about what metaphysically explains what. Likewise, those who think that causal processes outstrip scientific explanations (and thus dislike the unfiltered causal process theory) and are sceptical of causation* (as they should be; see §3.4.1) are likely to think that the truthmakers for scientific explanations are, in part,

mind-dependent. (For those who are sceptical of grounding yet still seek a non-psychologistic/objective theory of metaphysical explanation, I develop non-psychologistic DN theories in Chapters 6 and 7).

Any plausible account of the truth conditions for claims of the form ‘*x because y*’ needs to correctly assign truth-values to instances of the schema. Correctly assigning truth-values involves assigning truth-values that typically match what, after reflection, we are inclined to say about the truth-value of some particular instance of ‘*x because y*’. But recall the only candidate thus far considered for a mind-independent yet grounding-free theory of what makes true claims of the form ‘*x because y*’: the modal relations theory. According to the modal relations theory, ‘*x because y*’ is true just in case one of the traditional modal relations obtains between the referent of ‘*y*’ and the referent of ‘*x*’.

As we saw in Chapter 2, the modal relations theory of metaphysical explanation fails because it frequently assigns truth-values to ‘*x because y*’ which differ from the truth-values we would, after reflection, be disposed to assign.⁹³ Thus, this view would deliver the undesirable result that many intuitively false metaphysical explanations are true. As we will see in Chapter 6, the unfiltered DN theory has similar flaws. Thus, if we want to develop the most plausible semantics for claims of the form ‘*x because y*’, we should suppose that mind-dependent facts are doing some of the truthmaking work.

The third reason to think that ‘*x because y*’ is like ‘*x is tasty*’ or ‘*x is beautiful*’ or ‘*x is wrong*’ is that in each case there appear to be faultless disagreements between parties, both of whom are speaking truly. I find it plausible that agents with significantly different psychologies and interests may faultlessly disagree about what metaphysically explains what, since what counts as a metaphysical explanation for one agent may not count as such for another. But I recognise that this will not seem intuitive to those whose intuitions are more aligned with the unfiltered grounding-based theorist’s. Here, I am inclined to say that claims of the form ‘*x because y*’ are more like ‘*x is wrong*’ or ‘we all ought to ϕ ’ than they are to ‘*x is tasty*’ or ‘*x is sublime*’ (this is even clearer in the context of scientific explanation, where intuitions are even more convergent than they are regarding metaphysical explanation). To many, while the latter claims obviously seem to

⁹³ Another reason to be hesitant about this approach is that the traditional modal relations are typically thought not be explanatory unless they are, themselves, evidence for some further dependence relation. The thought is that all the traditional modal relations tell us is that certain modal correlations obtain. They don’t tell you anything about *why* they obtain. Correlation itself is not explanatory. See, for instance, Kim (1993).

admit of faultless disagreement, the former do not. In this regard I take the psychologistic view of metaphysical explanation to be more closely aligned with subjectivist theories in ethics.

According to the sophisticated subjectivist about moral facts, claims about what one morally ought to do are (roughly speaking) made true by what one's more rational and better informed self would advise one's current self to do, given one's current non-instrumental desires. That makes sophisticated subjectivism a good candidate for a relativist semantics. According to such a view, although you and I genuinely disagree when I say "one morally ought to φ " and you say "it is not the case that one morally ought to φ " we may disagree faultlessly since if we have different non-instrumental desires then our more rational selves will differently advise our current selves. Nevertheless, on the assumption that there is often significant overlap in the non-instrumental desires of psychologically similar creatures like you and me, we can expect our more rational selves to often give us the same advice. This explains why there is a certain amount of agreement regarding the moral truths, why frequently disagreement is not faultless, and why we might be inclined mistakenly to believe that moral truths are mind-independent.

Likewise, I think there is a certain amount of agreement regarding what metaphysically explains what because the relevant psychological features at play, in creatures like us, are quite similar. This creates a good deal of actual convergence in judgements, and can lead us mistakenly to believe that truths about metaphysical explanation are made true by mind-independent facts. In both cases, however, I think that further reflection reveals that faultless disagreement (even if only between psychologically very dissimilar beings (imagine the Borg)) is possible.

For all these reasons, I think some kind of psychologically filtered theory of metaphysical explanation is *prima facie* plausible. Furthermore, it is plausible that the relations we track with our explanations are the modal relations, for in every case where we intuit that 'x *because* y', y necessitates x. The resulting filtered modal relations theory is one according to which what makes it the case that 'x *because* y' is true when uttered by contemporary metaphysicians is that there is a modal relation between y and x, and that beings like us are disposed to come to have a mental state according to which x *because* y. That is, there is some relation between the referent of 'y' and the referent of 'x'—a modal

relation—and agents like us find that knowing that *y* obtains increases our understanding of why *x* obtains.⁹⁴ That is all metaphysical explanation consists in. The devil is, however, in the details, and in the following section it is to these details that we turn.

5.3 The Devil is in the Details

What counts as a mental state according to which *x because y*? I want to be appropriately liberal about this. I will say the following. An individual has a mental state according to which *x because y* iff that agent either (a) has a belief that *y* explains *x*⁹⁵ or (b) finds *y* to increase her understanding of *x*. An individual is *disposed* to have a mental state according to which *x because y*, iff, were that individual to detect a modal relation between *y* and *x*, she would (in general, absent any defeaters) come to have a mental state according to which *x because y*. That raises the question of *whose* mental state or states are relevant in determining the truth of some claim of the form ‘*x because y*’?

It could be that the only mental states that are relevant are ours—roughly, contemporary Homo sapiens. Then if *we* typically have mental states according to which *x because y*, while aliens, or the Mayans, or merely possible Klingons, do not have mental states according to which *x because y*, this latter set of mental states is just irrelevant with regard to whether or not *y* metaphysically explains *x*. But that seems hopelessly parochial. Moreover, since the key idea upon which I intend to build is that claims about what is metaphysically explanatory are, in part, claims about what is found to increase understanding, and since finding something to increase understanding is relevantly like finding something tasty, or finding something beautiful, it seems plausible that truths about what metaphysically explains what are relative, or context dependent, in the way the truths about what is tasty, or beautiful, are often taken to be relative or context dependent.

As such, two models naturally present themselves as plausible ways of spelling out the truth conditions for claims of the form ‘*x because y*’: relativism and contextualism. I will articulate the views below in terms of relativism, though nothing hangs on this

⁹⁴ See Trout (2002) for a discussion of the relationship between explanation and understanding.

⁹⁵ Notice that my aim, here, is to provide truth conditions for claims of the form ‘*x because y*’. The aim is not to provide a reductive account of metaphysical explanation that does not appeal to notions of explanation more broadly. So I assume that there are beliefs of the form ‘*y* explains *x*’.

choice and the reader with contextualist leanings should feel free to translate what I say into contextualist semantics.

I take relativism to be the thesis that some utterances express a proposition that determines a truth-value relative to a <world, time, individual> triple. On such a view the proposition *expressed* is the same in all contexts—that is, the content of what is said is the same in all contexts—but relative to different individuals these contents determine different truth-values.⁹⁶ So, it is useful to think of the individual parameter as being an ‘assessor’ parameter. Consider the triple <w, t, Jeff> and the triple <w, t, Mary>. Now consider the proposition *vegemite is tasty*. On the assumption that Jeff and Mary have different gastronomic standards, the content determines different truth-values in this case. Relative to Jeff’s standards the content is (let us suppose) true, and relative to Mary’s the content is false.

Suppose that utterances of the form ‘x *because* y’ express a proposition that determines a truth-value relative to a <w, t, i> triple. Then the earlier question of whose mental states are the ones that matter can be reframed as the question of how to give truth conditions for ‘x *because* y’ relative to a centred world <w, t, i>. In what follows I consider a number of options. Ultimately I think that the view I call *Dispositional Community Relativism* does the best job of correctly assigning truth conditions to ‘x *because* y’ and therefore should be preferred. But it will be instructive to see why this is so. Moreover, I leave it open that others might prefer the truth conditions provided by one of my other proposals. Let us consider, first, Radical Individualism.

Radical Individualism:

‘x *because*⁹⁷ y’ is true relative to a centred world <w, t, i> iff:

- (1) i has the mental state according to which x *because* y, and
- (2) Either (i) y necessitates x or (ii) x supervenes on y.

As noted, condition (1) requires that i (a) has a belief that y explains x or (b) finds y to increase her understanding of x. Notice that condition (2) rules out explanation in the case where the relevant individual has mental states according to which, say, Socrates exists *because* {Bert} exists. There is no modal relation between Socrates and {Bert}, so

⁹⁶ Here, I take propositions to be functions from worlds to truth-values.

⁹⁷ The relevant ‘because’, for now, is restricted only to metaphysical explanation. I will later broaden the theory to also cover scientific explanation.

‘Socrates exists *because* {Bert} exists’ is not true according to this analysis. Similarly, consider an assertion of ‘the flower is scarlet *because* it is red’ at a world centred on an individual who has the mental state according to which the flower is scarlet *because* it is red. There is a modal relation between the flower’s being scarlet and its being red, but that relation obtains non-symmetrically, and in the wrong direction. The necessitation goes from the flower’s being scarlet to its being red, so (2) is not met and the claim is not true, according to this analysis.

Despite these virtues, this proposal strikes me as too radically individualistic. In essence, any utterance of ‘x *because* y’, at a centre is true, as long as some relevant modal relation obtains and the individual at the centre has the mental state according to which x *because* y. I think it unlikely that such a semantics assigns truth-values to utterances of ‘x *because* y’ that typically accord with our post-reflection judgements about these matters. For instance, if Bill, my neighbour, is considered as the centre of our world, Bill’s assertion of ‘Socrates exists because {Socrates} exists’ comes out as true given that he has the relevant mental state. While some might find this plausible, I suspect that most of us will not, and therefore Radical Individualism will fail adequately to assign truth-values to utterances. The obvious amendment is to bring in the standards of some relevant community of individuals. One option is the following:

Community Relativism:

‘x *because* y’ is true relative to a centred world $\langle w, t, i \rangle$ iff:

- (1) Typically, members of the community in which i is embedded, have the mental state according to which x *because* y, *and*
- (2) Either (i) y necessitates x or (ii) x supervenes on y.

The problem with Community Relativism is that perhaps nobody in the community has entertained whether x *because* y. Yet this should not debar it from being the case that ‘x *because* y’ is true at $\langle w, t, i \rangle$. There is an easy fix for this, leading us to my preferred view, Dispositional Community Relativism:

Dispositional Community Relativism:

‘x *because* y’ is true relative to a centred world $\langle w, t, i \rangle$ iff:

- (1) Most members of the community in which i is embedded are disposed to have the mental state according to which x *because* y, *and*

(2) Either (i) y necessitates x or (ii) x supervenes on y .

Condition (1), by appealing to dispositions, deals with the situation described above where the community has simply failed to ponder some modal relation. Thus, Dispositional Community Relativism captures the conditions under which we actually judge ' x because y ' to be true. However, the question remains as to what to say when condition (1) is false.

Consider the Euthyphro case. Arguably, this a case in which no utterance of 'God loves X because X is good' or of ' X is good because God loves X ' expresses a true proposition. At least, on the assumption that God necessarily fails to exist, this is so, since on that assumption clause (2) is false. That seems to be the right result. But suppose, for the sake of an example, God does exist, and exists of necessity. Consider, then, some community in which 50% of people have the mental state according to which God loves X because X is good and the other 50% have the mental state according to which X is good because God loves X . Then at a world centred on a member of this community, any utterance of either sentence fails to be true since (1) is not true. Consideration of this case, I think, reveals that we ought to distinguish between occasions in which an utterance of ' x because y ' is false, and occasions in which an utterance of ' x because y ' is neither true nor false. If at $\langle w, t, i \rangle$ either (1) or (2) does not hold, then ' x because y ' is not true relative to that $\langle w, t, i \rangle$. But it does not follow that ' x because y ' is false. I suggest the following as falseness conditions.

Dispositional Community Relativism:

' x because y ' is false relative to a centred world $\langle w, t, i \rangle$ iff:

- (1) Most members of the community in which i is embedded are not disposed to have the mental state according to which x because y , or
- (2) It is not the case that y necessitates x , or the case that x supervenes on y .

Given these falseness conditions, there is room for cases in which relative to $\langle w, t, i \rangle$ ' x because y ' is neither true nor false. That will be the case, for instance, if an appropriate modal relation obtains between y and x , but it is not the case that most members of the community in which i is embedded are disposed to have the mental state according to which x because y , nor the case that most members of that community are not disposed to have the mental state according to which x because y . This is just what we see in the Euthyphro case on the assumption that God exists of necessity. Since the community in

question is split with respect to which mental state they have, utterances of ‘God loves X *because* X is good’ at $\langle w, t, i \rangle$ turn out to be neither true nor false, as do utterances of ‘X is good *because* God loves X’.

That, then, is the (psychologically) filtered modal relations theory of metaphysical explanation. Next, I consider what to say about the vagueness inherent in the theory.

5.4 Vagueness

We can expect there to be indeterminacy regarding whether clause (1) in the truth and falseness conditions is met, since ‘most’ is itself a vague notion. What percentage of the community needs to have a mental state according to which x *because* y , in order for most members of that community to have that mental state? There is surely no one answer to that question. So it might sometimes be indeterminate whether (1) of the falseness conditions is met, and indeterminate whether (1) of the truth-conditions is met. This will result in meta-indeterminacy. These will be cases in which it is indeterminate whether ‘ x *because* y ’ at $\langle w, t, i \rangle$ has a determinate truth-value, (true or false) or is gappy (is neither true nor false).

There is also an indeterminacy regarding how to understand the locution ‘the community in which i is embedded’. Suppose the individual at i is Wendy. Suppose that at the age of 13 Wendy is abducted by aliens and that she spends the next 40 years living amongst these aliens. Is her community the community of aliens? Is it the community she was born into and lived in for the first 13 years? Is it the conjunction of both, or the disjunction of either? Suppose that Wendy’s community up until age 13 is disposed to have the mental state according to which x *because* y . Suppose the alien community is not disposed to have the mental state according to which x *because* y .

When Wendy asserts ‘ x *because* y ’ at the age of 30, does she assert a truth or not? It is my view (though there are others that one could coherently take) that it will always be at least somewhat vague which individuals are members of the community in which i is embedded. It seems natural to model this by saying that the set of individuals that compose the relevant community, on any occasion, is a fuzzy set—its members are members of that set to some degree greater than 0 and up to and including 1. If all

members of the relevant fuzzy set share the same dispositions with respect to having a mental state according to which *x because y*, then that the set is fuzzy has no implications for whether '*x because y*' takes a determinate truth-value at $\langle w, t, i \rangle$. If however, some of the members of the relevant fuzzy set have different dispositions to have mental states of these kinds, then it will sometimes make a difference to whether '*x because y*' has a determinate truth-value at $\langle w, t, i \rangle$.

For it seems very plausible that we ought to differentially weight the opinions of members of the community, depending on the degree to which they are members of the relevant fuzzy set. Someone, for instance, who is a member of a set to degree .1, ought to have less 'weight' with respect to determining which mental state most of the community have, than an individual who is a member of the set to degree .9. I will not, here, offer a schema for weighting the input of different members of the relevant fuzzy set, though there are clearly a number on offer (one might just move directly from degree of membership to a weight). The point is that it might be a vague matter whether most of the community in which *i* is embedded has the mental state according to which *x because y*, even in cases where a clear majority of members of the set of individuals that compose that community have that mental state: for it could be that all of those individuals are members of the set to a very low degree. This, again, may lead to cases of meta-indeterminacy, in which it is indeterminate whether '*x because y*' at $\langle w, t, i \rangle$ takes a determinate truth-value, or lacks a determinate truth-value.

While I wish to remain neutral about which is the best way of dealing with these issues of indeterminacy, I think that Dispositional Community Relativism is the right way to spell out the truth conditions for utterances of '*x because y*' according to the filtered modal relations theory. Using these semantics we can vindicate, as true, actual plausible utterances of '*x because y*' made by contemporary metaphysicians, such as (A) through (H). Assuming that most ordinary people share the same dispositions as metaphysicians about such cases,⁹⁸ it will often be the case that an utterance of '*x because y*' at a centre of our world, is true, and an utterance of '*y because x*' is false, and that will be so even where the modal relation that obtains between *x* and *y*, obtains symmetrically. While my account vindicates, as true, most of the claims that contemporary metaphysicians wish to

⁹⁸ It is an interesting empirical question how the folk are disposed to respond in any given case. My account leaves it open that claims (A) through (H) turn out to be false if, as it turns out, most of the community are not disposed in the manner that contemporary metaphysicians are disposed.

make, it is also sufficiently liberal that it allows that relative to other centred worlds, quite different utterances of ‘x *because* y’ come out as true (or false).

This means that, while the semantics developed here delivers the same verdicts as grounding-based theories about actual human utterances of ‘x *because* y’, the theories will diverge when considering utterances made by beings with different psychological mechanisms. I think that this is an attractive feature of the view, though others will surely disagree. For instance, I doubt that I have convinced any defenders of the ontic view of explanation to adopt my proposal. However, I suspect that many who are convinced by my anti-grounding arguments will prefer a psychologistic, relativistic, theory of metaphysical explanation over an error theory.

5.5 Revisiting the Desiderata and the Theoretical Virtues

With our first grounding-free theory of metaphysical explanation now in hand, we can weigh the view against the desiderata for a theory of explanation (§2.4), and consider the theoretical virtues of a version of the Grounding-free Theory that endorses this theory of metaphysical explanation. The filtered modal relations theory satisfies precisely the same desiderata as the filtered grounding-based theory (namely, all but objectivity), but without positing grounding relations. Indeed, as the filtered modal relations theory can do so as part of a more parsimonious overall theory, this was presented as an objection to the filtered grounding-based theory in §3.5.

The view does well with covering cases, for all the metaphysical explanations have the modal scope of metaphysical necessity. This is why the *unfiltered* modal relations theory ascribed counterintuitive truth-values only by judging too many instances of ‘x *because* y’ to be true, rather than judging some intuitively true explanations to be false. However, the filtered view fails to achieve the objectivity of what explains what. Such is the nature of a psychologistic account. This is why those who think of explanation as an ontic phenomenon will not find the view attractive. However, those who think of explanation as an epistemic phenomenon that is tied up with increasing agents’ understanding, sacrificing objectivity will not be a substantial cost.

Furthermore, the view inherits from our priority intuitions the characteristics of irreflexivity, asymmetry and relevance. As such, none of the true metaphysical

explanations uttered around here will be reflexive, symmetrical or include what we deem to be irrelevant idle wheels. However, this is not built into the account itself. It is left open that, at some centred worlds, there might be self-explanatory states of affairs, and mutually explanatory states of affairs, if the community is disposed to have the relevant kinds of mental states. Thus, the view only satisfies these epistemic desiderata when applied to communities in which these *really are* appropriate epistemic desiderata. I think this is a feature of the view, not a flaw. This is most clearly so in the case of the relevance desideratum, for it is natural to think that what is relevant to what is something that should be evaluated relative to a particular psychological orientation. Finally, the view secures the understanding desideratum, as those modal relations about which we have priority intuitions are those where knowing about one state of affairs helps to illuminate, for us, another state of affairs.

In sum, the view satisfies every desideratum except for objectivity, placing it on a par with the filtered grounding-based theory in this respect. It also satisfies the very same desiderata as the filtered causal process theory of scientific explanation (§2.5.2). Indeed, those friendly to the filtered causal process theory of scientific explanation will likely be friendly to the filtered modal relations theory of metaphysical explanation. To such people I suggest a *unified* theory of scientific and metaphysical explanation.

Consider that, if the story I told in Chapter 4 is right, we use the very same causal detection mechanism to generate our intuitions of both diachronic/causal priority and non-diachronic/metaphysical priority. Thus, the following unified theory naturally presents itself:

Scientific/Metaphysical Dispositional Community Relativism:

‘x because y’⁹⁹ is true relative to a centred world $\langle w, t, i \rangle$ iff:

- (1) Most members of the community in which i is embedded are disposed to have the mental state according to which x because y, *and*
- (2) Either (i) y necessitates x or (ii) x supervenes on y or (iii) y causes x .

Moreover, ‘x because y’ is false relative to a centred world $\langle w, t, i \rangle$ iff:

- (1) Most members of the community in which i is embedded are not disposed to have the mental state according to which x because y, *or*

⁹⁹ Here, the relevant ‘because’ is underlined, indicating that it has been appropriately broadened so as to include both scientific and metaphysical explanations.

(2) It is not the case that y necessitates x , or the case that x supervenes on y , or the case that y causes x .

As the very same psychologic machinery can be used to elucidate the truth conditions for the filtered causal process theory, it makes a lot of sense to unify these theories in this way. While the resulting theory will not serve as a generic theory of explanation, there is potential to broaden its scope so as to include mathematical explanations, logical explanations, etc. (though I shall not explore this here).

That these theories can be so neatly amalgamated in this way serves to demonstrate the theoretical virtue of the Grounding-free Theory. Not only is this a clear case of a virtuously unified theory, it is both parsimonious and elegant (by appealing to the very same mechanisms and conceptual machinery to which we were already committed for our filtered causal process theory).

Thus, for those who desire vindication of the priority intuitions, and are willing to accept that what metaphysically explains what is not an objective matter, the Grounding-free Theory, paired with the filtered modal relations theory of metaphysical explanation, is a very strong theory overall which nicely explains and vindicates our priority intuitions.

Chapter 6: Deductive-Nomological Theories of Metaphysical Explanation

In earlier chapters I identified the priority intuitions as one of two explananda whose best explanation is thought to indispensably require grounding. I explored the explanations of these intuitions offered by the Sophisticated Grounding Theory and the Grounding-free Theory (the grounding-based and grounding-free explanation, respectively), and argued that while they are empirically equivalent, the latter is more theoretically virtuous. Thus, I concluded, we have no reason to believe that grounding relations exist. However, I also noted that some will want to explain the priority intuitions in such a way as to vindicate these intuitions—to show that, in many cases at least, those claims of the form ‘*x because y*’ which we intuit to be true are, in fact, true. I have already shown that a grounding-based theory of metaphysical explanation is not the only way to achieve this vindication. The psychologically filtered modal relations theory is more than up to the task, as shown in Chapter 5. Moreover, I argued in Chapter 3 that both the filtered and unfiltered grounding-based theories of metaphysical explanation are seriously flawed.

However, there are no doubt those who, when evaluating a prospective theory of explanation, heavily weight the ontic desideratum of objectivity. From this perspective, nothing I have offered thus far will be adequate. That’s because, if I am right, the ‘unfiltered’ grounding-based theory is, in fact, illicitly (psychologically) pre-filtered, and thus cannot truthfully claim to be objective. Moreover, the filtered grounding-based theory and the filtered modal relations theory both explicitly jettison objectivity in order to secure the epistemic desiderata. As such, the remaining chapters explore the prospects of developing objective, yet grounding-free, theories of metaphysical explanation. To reiterate, these chapters are *exploratory*, and I don’t claim to have developed full-fledged theories that are free of outstanding kinks and problems. Nevertheless, I contend that the success I have had with what follows indicates that these are promising avenues for future research.

In this chapter, we move away from theories of metaphysical explanation based around modal relations and grounding relations, and explore the prospects of generalising Hempel and Oppenheim’s (1948) deductive-nomological (DN) theory into the metaphysical realm. I show how a filtered variant of this theory can fulfil the very same desiderata as the filtered modal relations theory. For those who think of

explanation as an epistemic notion, but also think that *laws* should play a role in explanations, the theory proffered here is preferable to that proffered in the previous chapter. It is worth noting, however, that this filtered theory secures the epistemic desiderata at the cost of the ontic desiderata. Given that, my main focus is on the more interesting question of how far an *unfiltered* DN theory—a theory which will meet the ontic desideratum—can go towards satisfying the epistemic desiderata. The chapter will be framed around this investigation, for, given any unfiltered DN theory, additional filtration of the kind described in Chapter 5 will suffice as a way of sacrificing objectivity in order to jump any epistemic hurdles.

The theories of metaphysical explanation built here can be unified with the scientific DN theories. This unification can be thought of in one of two ways. If the metaphysical explanations fall under the umbrella of the scientific explanations, I am simply showing how the DN theory extends to cover these cases. If they do not, I am showing how the DN theory can be appropriated to cover non-scientific cases. Nothing of substance hangs on this difference. In addition, along the way I'll show how the theories I develop here are superior to Wilsch's Nomological Account of Ground—the one extant metaphysical DN theory in the literature.

Without further ado, the plan is as follows. §6.1 sets to the task of building a classical DN theory of metaphysical explanation. §6.2 presents my preferred account of metaphysical laws as metaphysically necessary generalisations. §6.3 distinguishes material and formal entailment, and argues that it is the latter notion that is relevant to the theory we are building. With our theory in hand, §6.4 weighs it against the desiderata. I conclude by comparing my theory to others on offer. §6.5 quickly shows that deRossett's (2013) early foray into this terrain is not really a DN theory, while §6.6 explores Wilsch's (2015, 2016) reduction of grounding relations to DN arguments. Although Wilsch provides an interesting explication of metaphysical laws, his theory inherits the illicit pre-filtering/anthropocentrism of the unfiltered grounding-based theory, and the non-classical notion of entailment he employs constitutes additional ideology without advantage. In §6.7 I'll sum up and point the way forward.

6.1 A Classical DN Theory of Metaphysical Explanation

According to Hempel and Oppenheim (1948:137–138), there are four conditions of adequacy that must be satisfied in order for an explanation to be sound. Together, these four conditions constitute a basic DN theory. The conditions can be stated as follows:

1. The explanans must entail the explanandum.
2. The explanans must contain physical laws of nature, and these must be required for the derivation of the explanandum.
3. The explanans must have empirical content; i.e. it must be capable, at least in principle, of test by experiment or observation.
4. The propositions constituting the explanans must be true.

Conditions (1) and (4) identify explanations with sound arguments. Conditions (2) and (3) constrain the space of arguments that count as explanations: only those arguments that feature laws and have empirical content count. Note that a law of nature is *required* for the derivation of an explanandum just when it is essential to the argument in which it appears, in this sense: removing the law would break the entailment of the conclusion by the premises.

Conditions (1) and (4) can be carried over to the metaphysical case without alteration, though there will be substantially more to say in §6.3 and §6.6.2 about the relevant notion of entailment. The remaining conditions of adequacy will need to be modified to handle cases of metaphysical explanation. First, cases of metaphysical explanation outstrip the physical laws. There is no physical law of nature underwriting, for instance, the relationship between a singleton and its urelement, or between natural and moral properties. Indeed, it is difficult to see how there could be: at best, physical laws possess the modal strength of physical necessity. As noted, however, metaphysical explanations encode modal relationships that are stronger than physical necessity. So any physical law will be too weak. An extended discussion of the requisite laws takes place in §6.2 and §6.6.1.

Second, while some metaphysical theses are open to empirical test and observation, many are not, and yet explanations arise nonetheless. There is, for instance, no empirical

observation to be had of sets and urelements. Similarly, it is not obvious that one can empirically test via experiment or observation the mereological relations between entities and so on. Rather, our ‘observations’ in this domain proceed via imagining *counterfactual* interventions, as discussed in Chapter 4. So the requirement of empirical content in condition (3) is misguided in this context. Rather, what we’re interested in is whether the premises have *ontological* content. In other words, do they (tacitly, perhaps) involve quantification over some entities, such as objects, properties, events, facts (as in: parts of the world) and so on, where those entities may or may not be open to empirical observation?

Let us, then, replace the third condition with a plea for content of this kind. In this way, the model is restricted to cases where the existence (and, perhaps, properties) of one entity putatively explains the existence (and, perhaps, properties) of another entity. To modify the second condition we must replace the reference to physical laws of nature in (2) with reference to something stronger: a metaphysical law. Indeed, the idea that anything like a deductive nomological theory of metaphysical explanation can be developed pretty clearly presupposes the idea that we can make sense of something sufficiently law-like that it can play a role in underwriting metaphysical explanations via deduction.

Thus, the conditions of adequacy for a metaphysical DN theory are as follows:

1. The explanans must entail the explanandum.
2. The explanans must contain metaphysical laws, and these must be required for the derivation of the explanandum.
3. The explanans must contain ontological content.
4. The propositions constituting the explanans must be true.

As noted, the laws do a lot of heavy lifting in a DN theory. Thus I now offer my preferred account of the metaphysical laws.

6.2 Metaphysical Laws as Metaphysically Necessary Generalisations

I propose the following account: metaphysical laws are metaphysically necessary¹⁰⁰ generalisations that govern the relationship between some entity/entities and another entity/entities, where these entities are of the kind required by the third condition of adequacy. In other words, the metaphysical laws simply describe the law-like relationships captured by the modal relations. For instance, they relate parts to wholes, numbers to sets, truths to truthmakers, natural properties to moral properties and so on. The metaphysical laws have a particular logical structure. As generalisations, they are in universal form. Moreover, I will suppose that the metaphysical laws are also invested with some kind of conditionality, and so are best represented in either conditional or biconditional form.

There are three reasons for the conditionality requirement. Importantly, however, these reasons do not settle whether the laws are conditional or biconditional, merely that they involve conditionality in some sense. First, it is common to treat physical laws as possessing conditionality. They tell us that *if* an event of type A occurs, *then* an event of type B will follow. If the metaphysical laws are in any way analogous to the physical laws then we can expect a similar logical form.

Second, it seems plausible that the metaphysical laws would hold, at a world, *w*, whether or not the relevant entities they refer to exist in *w*. For example, even if the proposition that snow is white were false, because snow is not white, *if* snow were white (and hence the fact that snow is white were to exist), <Snow is white> would be true. Similarly, even if sets and numbers do not exist (i.e. mathematical nominalism is true), nonetheless: if numbers did exist, then so would the corresponding sets containing those numbers. And so on for each case where there is a modal relation. The fact that A necessitates B, for example, can be described in conditional form: if A exists, B exists. This is true even in worlds where A does not exist. Conditionality captures this feature of the laws since a conditional remains true when the antecedent conditions it describes, such as the existence of an entity of a certain kind, are not met.

Third, it is hard to see how there could be any non-conditional metaphysical laws, at least assuming the restriction to universal form. For instance, consider the non-conditional claim that $\forall xAx$. Such a claim would be extremely strong if it were a law. It

¹⁰⁰ Metaphysical necessity is, here, considered to be the broadest modality, on a par with logical possibility. There are no possible worlds that are metaphysically impossible.

would be saying that everything has a property of a certain kind. It is quite hard to see how there could be any laws along these lines. Furthermore, such laws would not be particularly useful for explanation, since explanations are typically interested in the relationships between restricted classes of entities, and such restrictions standardly require conditionality.

Of course, none of this is decisive. There may be metaphysical laws that are not conditional in structure.¹⁰¹ However, if we restrict our focus to metaphysical explanations that involve ontological dependence then it would seem that all of the relevant laws involve conditional generalisations that hold of metaphysical necessity.

Still, you might ask (perhaps a little more pointedly), *what are metaphysical laws?* It's hard to say. Moreover, the prospects for giving an entirely satisfactory account of metaphysical laws are dim, since the associated project in the philosophy of science—giving an adequate account of a physical law—has not fared all that well (Cartwright, 1983). On the other hand, this should give us hope. For there may not be any special reason to worry about the notion of a metaphysical law. If it is equally hard to give an account of the physical laws, despite valiant attempts to do so, then perhaps the real trouble, if there is one, lies with the concept of a law, rather than with the broadening of that concept to cover metaphysical cases as well.

At any rate, this is obviously a thorny issue. However, in order to get a basic DN theory of metaphysical explanation up and running, all we really need are some candidate generalisations with conditional structure and ontological content that can be used to derive the various explananda in which we are interested, and that possess the modal strength of metaphysical necessity. Fortunately, it is a relatively straightforward matter to reverse-engineer candidate generalisations of the right modal strength from putative cases of metaphysical explanation such as those identified in Chapter 2. So, let's consider those cases again, and identify some putative laws that might be at play. To be clear, I am not defending these laws as the *right* laws. Indeed, in some cases, there are reasons to be sceptical about the particular laws I identify below. However, the below is sufficient to indicate how the *right* laws might look.

¹⁰¹ For instance, if presentism is necessarily true it is plausibly a metaphysical law that for all x , x is present. However, such a law could be rendered conditionally: if x exists, then x is present. Moreover, derivations made using this law (i.e. the derivation of an entity's presentness from its existence) don't seem too implausible as metaphysical explanations. Regardless, I shall continue on the assumption that the metaphysical laws are conditional in form.

1. Action X is wrong because it will cause a great amount of suffering and no happiness.

The law in this case could be:

[L1] Necessarily, if action ∂ possesses natural property of kind K , then ∂ is wrong.

Where the ‘necessarily’ here—and below—is read as metaphysical necessity. Similarly, consider:

2. The composite object James exists because of the existence and arrangement of the simples that compose him.

The law in this case could be:

[L2] Necessarily, for any entities P_1, \dots, P_n , if the P_n exist and relation R obtains between the P_n , then W , the object composed by the P_n , exists.

In this case, relation R tells us the way things need to be arranged, in order for there to exist a composite whole. We need not take a stand, here, on whether composition is restricted, universal, or never obtains at all.

3. James is in a state of pain because his C-fibres are firing.

The law in this case could be:

[L3] Necessarily, for any entity E , if E exists and possesses a physical property of kind K , then E possesses the mental property of being in pain.

Obviously, a lot more needs to be said about kind K , but [L3] will suffice for our current purposes.¹⁰²

4. The existence of the urelement Socrates explains why the singleton set $\{\text{Socrates}\}$ exists.

The law in this case could be:

[L4] Necessarily, for any entity E , if E exists then the singleton set $\{E\}$ exists.

¹⁰² You might, reasonably, be sceptical that [L3] is the right law, since you might think that there can be C-fibres firing without pain (i.e. in animals in which C-fibres don't have anything to do with pain). However, the example still gives an idea of how the correct law might work.

Indeed, for some of the putative cases of metaphysical explanation identified, generalisations of this kind have already been provided. For instance, with respect to:

5. The proposition <A friendly Airedale terrier exists> is true because of the existence of Jasper, the friendly Airedale.

We already have:

[Truthmaker] Necessarily, for each proposition P there is some entity E such that if E exists then P is true. (Armstrong, 1997, 2004)

Yet [Truthmaker] is not quite apt to figure in a DN derivation of example 5, as we shall see in §6.4.1.

Finally,

6. The champagne glass is fragile because of the crystalline bonds between its component molecules.

The law in this case could be:

[L6] Necessarily, for any entity E , if E exists and possesses a categorical property of kind K , then E possesses the dispositional property of being fragile.

Once again, you might disagree that this is the right law, but it will serve as an example. That many of these laws help themselves to *kinds* implies that there are other principles in the background delineating what properties are of what kind, rather like Wilsch's (2015) *ontological principles* (see §6.6.1).

When we reflect on the six putative laws above, natural divisions in kind emerge. There seem to be (at least) three types of laws: predicative, ontic, and mixed laws. An ontic law expresses the fact that if an entity of one kind exists, then another entity (of the same or another kind) also exists. The entity to singleton law is an example of an ontic law: it connects urelements to sets. A predicative law expresses the fact that, for a particular entity, or particular kind of entity, E , if E has a particular kind of property, then it also has a property of another kind. The natural property to moral property case is an example of a predicative law: it connects some properties of an act to other properties of the same act. Finally, a mixed law takes us from the existence of an entity to the possession of a particular kind of property by another entity (or *vice versa*). The

truthmaker law is a mixed law: it takes us from the existence of a truthmaker to the possession of an alethic property by a proposition.

You get the point: candidate metaphysical laws are pretty easy to come by. We can build them, and we can sort them. While I have not provided an overarching theory of metaphysical laws (I have, in effect, simply identified some putative metaphysically necessary generalisations that might be at play in various metaphysical explanations), it is notable that with respect to physical laws, Hempel ultimately takes a ‘we know ’em when we see ’em’ approach. The thought being that we are able to identify candidate physical laws fairly easily, despite lacking a theory of *what they are*, and that’s enough to get a DN theory of scientific explanation off the ground. A similar approach is available in the metaphysical case. With respect to metaphysical laws, perhaps we do know ’em when we see ’em. The putative laws above make good examples: if true, they certainly seem to be metaphysical laws.

Still, one might demur, and seek to press for some account of the metaphysical laws. One tempting move is to attempt to generalise the standard account of a physical law of nature. According to that account, a physical law is a generalisation describing a *non-accidental* physical regularity. Likewise, a metaphysical law might be thought of as a generalisation describing a non-accidental metaphysical regularity. In other words, the metaphysical laws could be those that govern only the *non-accidental* modal correlations identified in §2.1.1.

Of course, in the context of developing a grounding-free DN theory of metaphysical explanation, it is not viable to call upon grounding to distinguish the accidental from the non-accidental modal correlations. Moreover, my preferred explanation of our judgements regarding these correlations—which I have not explored in this thesis—is not well placed to underwrite this distinction in the context of an objective DN theory.

Indeed, I suspect that any attempt to restrict the metaphysical laws such that accidental modal correlations are not lawful will result in a theory of metaphysical explanation that is illicitly pre-filtered in the same way that the unfiltered grounding-based theory is illicitly pre-filtered. For, while we have strong intuitions according to which we want to distinguish these correlations, finding a non-anthropocentric difference between them remains elusive. Thus, labelling some of these correlations as accidental

and then building a supposedly objective theory of metaphysical explanation where only those derivations that make use of non-accidental generalisations count as explanations is potentially illicit.

Nevertheless, perhaps we are tracking something metaphysically substantial when we judge some modal correlations to be accidental and others non-accidental. It would be useful if we were, for by constraining the class of metaphysical laws along these lines, we can develop a DN theory that, while retaining objectivity, does better with the epistemic desiderata. However, I suspect that an upstanding DN theory of metaphysical explanation will not help itself to this distinction. Moreover, this is not a problem that is particular to the notion of a metaphysical law: no definition of a non-accidental regularity has been found adequate for the case of physical laws, either.

For now, I will proceed on the assumption that the metaphysical laws are metaphysically necessary generalisations with ontological content that are conditional in form. However, when we measure the theory against the desiderata, I'll note that if some respectably non-anthropocentric difference between the different kinds of modal correlations could be established, the resulting DN theory would do better according to some of the epistemic desiderata.

6.3 Material and Formal Entailment

Before we continue, it is important to distinguish two different kinds of classical entailment. This is because the notion of entailment in the first condition of adequacy might be read in one of two ways:

Formal Entailment: Premise set Σ formally entails Π if there is no possible interpretation of the non-logical expressions in Σ or in Π under which the premises are true and the conclusion is false.

Material Entailment: Premise set Σ materially entails if, holding fixed the interpretation of the logical and non-logical expressions in that argument, there is no possible world in which the premises in Σ are true and Π is false.

If an argument satisfies formal entailment, then it is formally valid; if it satisfies material entailment then it is materially valid. The relevant notion for the theory we are building

here is formal entailment. This is because material entailment forces us to strip the metaphysical laws from the arguments, leaving no arguments which satisfy all the conditions of adequacy, and therefore no metaphysical explanations! The problem is that the lawful premise must be *required* for the derivation of the conclusion, but metaphysical laws—understood as metaphysically necessary generalisations—are idle premises with regards to material consequence. To see this, consider the following argument:

(1) 2 exists.

[L4] For any entity E , if E exists then the singleton set $\{E\}$ exists.

Therefore,

(2) $\{2\}$ exists.

This argument is materially valid, and looks like a good explanation. But now consider the argument:

(1) 2 exists.

Therefore,

(2) $\{2\}$ exists.

Given the modal correlation between 2 and $\{2\}$, it is metaphysically impossible for (1) to be true and (2) false. So material entailment holds for this second argument also. So [L4] in the first argument is not essential to the entailment of the explanandum. As this directly violates the second condition of adequacy, we are forced to strip the lawful premise from the argument. Call this the *suppression* problem.

Assuming that the metaphysical laws are metaphysically necessary, the suppression problem is widespread. Indeed, it seems to rule out the viability of *any* putative metaphysical explanation. This is because for any such argument, the non-lawful premise(s) will materially entail the conclusion without the help of the lawful premise(s). Thus the second condition of adequacy will always require that the lawful premise(s) be suppressed. But the resulting arguments lack laws, which also violates the second condition, leaving no metaphysical explanations which satisfy all four conditions.

It might, then, be suggested that demanding that the laws must be required for the derivation of the conclusion is too strong, and that the second condition be weakened

accordingly. This would alleviate the suppression concerns. However, in general we *want* to suppress premises not required for the derivation of the conclusion, on pain of allowing explanations with all kinds of irrelevant premises. So we need to keep the second condition unaltered.

The good news is that the suppression problem evaporates if we move from material to formal entailment. That's because the arguments which lack laws are materially valid but formally invalid. Let's look further into the urelement/singleton case. Under material entailment, the second condition of adequacy forced us to remove the lawful premise. However, if we remove the lawful premise in the context of formal validity, the resulting argument would not be a valid argument form (it would have the form *A therefore B*). Regardless of the fact that A and B happen to exist at all of the same worlds, this argument is formally invalid, and thus the conditional lawful premise is not formally suppressible. The solution here is quite general. For there is good reason to think that the laws will never be formally suppressible. As noted, it is natural to treat the laws as conditional or biconditional statements. We can therefore expect DN arguments to have a common structure, namely that of *modus ponens*, or of a series of instances of *modus ponens* forms chained together. Because conditionals are essential to the *modus ponens* form, however, it won't be possible to remove the laws without doing violence to the formal entailment structure of the relevant arguments. So, the DN theory of metaphysical explanation must, like the DN theory of scientific explanation, invoke formal validity.

6.4 Evaluating the DN Theories Against the Desiderata

6.4.1 Covering Cases

DN theories do well in covering the kinds of metaphysical explanations we actually believe to be true. We have already seen in §6.3 how we can generate DN derivations that plausibly constitute metaphysical explanations of the existence of singleton sets. Furthermore, we have noted some putative laws that seem to cover the other exemplar cases. Let's see these other derivations in action, with the order adjusted for convenience of exposition:

- (1) Action X will cause a great amount of suffering and no happiness.

(2) The property of causing a great deal of suffering and no happiness is of kind K .

[L1] Necessarily, for any action ∂ possessing natural property of kind K , if ∂ exists then ∂ is wrong.

Therefore,

(3) Action X is wrong.

(1) Relation R obtains between the collection of entities $P_1...P_n$.

[L2] Necessarily, for any entities $P_1...P_n$, if the P_n exist and relation R obtains between the P_n , then W , the object composed by the P_n , exists.

Therefore,

(2) W , the object composed by the P_n , exists.

(1) James' C-fibres are firing.

(2) The property of having C-fibres firing is of kind K .

[L3] Necessarily, for any entity E , if E exists and possesses a physical property of kind K , then E possesses the mental property of being in pain.

Therefore,

(3) James is in pain.

(1) The champagne glass has crystalline bonds between its component molecules.

(2) Crystalline bonds are of kind K .

[L6] Necessarily, for any entity E , if E exists and possesses a categorical property of kind K , then E possesses the dispositional property of being fragile.

Therefore,

(3) The champagne glass is fragile.¹⁰³

Now, consider the truthmaking case. We start with:

(1) Jasper, the friendly Airedale exists.

However, the conditional premise provided by:

[Truthmaker] Necessarily, for each proposition P there is some entity E such that if E exists then P is true.

is not apt to formally entail the conclusion that the proposition <A friendly Airedale terrier exists> is true. While [Truthmaker] may indeed be a metaphysically necessary generalisation, it won't allow us to derive our desired conclusion, and thus it isn't the right law for this derivation. The following does better:

(1) Jasper, the friendly Airedale exists.

[L5] Necessarily, for any entity E of kind K , if E exists then the proposition <an object of kind K exists> is true.

Therefore,

(2) <A friendly Airedale terrier exists> is true.

In this way, we can see how DN derivations can be constructed to cover each of our putative cases, which bodes well for the overall viability of the DN theory. However, like the unfiltered modal relations theory, our DN theory also covers cases that are not intuitive instances of metaphysical explanation. Let's now look at the other desiderata.

¹⁰³ Now, one might question whether [L3] and [L6] are, in fact, metaphysically necessary generalisations, with cases in mind such as Leuenberger's (2014) *chromoplasm* and *blocker*. I think the concerns raised by contingentists about grounding are genuine, and thus I reiterate that the derivations and laws I propose here are merely illustrations of how this kind of theory might work. The task of getting them exactly right is one to be attempted once the metaphysical DN theory is on the table. Here, I am merely offering the account and showing roughly how it would work, not defending any particular use of it—though I think that a plausible suggestion of how to deal with Leuenberger-style cases is to appeal to Armstrongian (2004) totality facts.

6.4.2 Irreflexivity

With the characterisation provided so far, it is a relatively simple matter to show that the DN theory will not generate reflexive metaphysical explanations. Consider the following argument:

(1) 2 exists.

Therefore,

(2) 2 exists.

Despite the fact that ‘2 exists’ formally entails itself, this does not make the cut as a metaphysical explanation. That’s because there are no laws at play. Yet, one might think that it is a metaphysical law that:

[L0] Necessarily, for any entity E , if E exists then E exists.

[L0] is a metaphysically necessary generalisation that, when added to the above argument, generates a formally valid derivation that includes a law:

(1) 2 exists.

[L0] Necessarily, for any entity E , if E exists then E exists.

Therefore,

(2) 2 exists.

[L0] is surely no *accidental* regularity. Perhaps we could exclude [L0] as a metaphysical law by insisting, for instance, that the laws must govern regularities between distinct existents, or that no claim in the premise set can feature in the conclusion. Yet there is no need for such restrictions. Fortunately, we can rule out the above derivation by noting that [L0] is not required for the entailment of the conclusion because, as noted, the lawless argument is also formally valid. Thus [L0] must be suppressed from the derivation. The resulting derivation lacks a law, and hence is not a metaphysical explanation. This solution is quite general: our DN theory will never generate reflexive explanations, and thus does well according to the irreflexivity desideratum.

6.4.3 Asymmetry

As we saw in §2.5.1, the symmetry problem, as it arises in the DN context, is that sometimes the conclusion of a DN derivation, together with the lawful premise(s), can be used to derive the non-lawful premise(s). This problem has always plagued scientific DN theories. Traditionally, the problem of symmetry results from deep symmetries in the laws of physics, allowing the interderivability of the explanans and the explanandum (van Fraassen, 1989). Yet the prevailing intuition is that the explanation only runs in one temporal direction in these cases. Symmetry problems remain in the context of the metaphysical DN theory, as our minimal characterisation of the metaphysical laws allows *biconditional* laws which permit derivation in either direction. Once again, consider the following argument:

(1) 2 exists.

[L4] Necessarily, for any entity E , if E exists then the singleton set $\{E\}$ exists.

Therefore,

(2) $\{2\}$ exists.

This, as noted, looks like a good metaphysical explanation. However, consider now the following, symmetrical argument:

(1) $\{2\}$ exists.

[L4*] Necessarily, for any entity E , E exists just in case the singleton set $\{E\}$ exists.

Therefore,

(2) 2 exists.

Both arguments are formally valid, have ontological content and feature a candidate metaphysical law. So both satisfy the four conditions laid down for a basic DN theory of metaphysical explanation. However, only the first argument strikes us as explanatory.

This is the symmetry problem, as it arises in the metaphysical context.

Our minimal characterisation of the metaphysical laws rules out neither of the following two candidate laws:

[L4] Necessarily, for any entity E , if E exists then the singleton set $\{E\}$ exists.

[L4*] Necessarily, for any entity E , E exists just in case the singleton set $\{E\}$ exists.

The conditional structure of [L4] is apt to capture intuitions of asymmetric explanatory priority: the existence of a set is thought to be explained by the existence of its member(s). [L4] will not allow the derivation of the existence of an entity E from the existence of its singleton set $\{E\}$. It can only be used to derive the existence of the singleton from the existence of its urelement. In contrast, the biconditional structure of [L4*] allows for symmetrical derivations, as we derive the existence of the singleton from the existence of its urelement and *vice versa*.

However, nothing in the account of metaphysical laws offered thus far rules out [L4*] as a metaphysical law. It is, after all, a conditional generalisation with the modal scope of metaphysical necessity that could easily govern the modal correlation between sets and their urelements. Indeed, even appealing to the accidental/non-accidental distinction does no work here, as it is surely no *accident* that an urelement exists in precisely those worlds where its singleton exists. If [L4*] is a metaphysical law, then there is nothing short of pragmatic filtering that can rule out the singleton-to-urelement derivation as an explanation.

So, the unfiltered DN theory of metaphysical explanation cannot satisfy the asymmetry desideratum. However, we are in a slightly better position in the metaphysical case than in the scientific case. That's because the biconditional versions of many of the other putative laws are false, showing that the symmetry problem is far less pervasive in the metaphysical context. While, as noted above, in the scientific case we have good reason to think that the physical laws are symmetrical across the board, the analogous assumption fails in the metaphysical case.

An interesting case is that of the law(s) governing the relationship between what exists and what is true. Consider the biconditional version of [L5]:

[L5*] Necessarily, for any entity E of kind K , E exists just in case the proposition <an object of kind K exists> is true.

Given [L5*] we would be able to derive both the truth of <an object of kind K exists> from the existence of a particular entity E , as well as the existence of a particular entity E , from the truth of that proposition. [L5*] is a bad candidate law, however, as it does

not hold of metaphysical necessity. Consider, for example, the proposition, P , that some arctic penguins are feisty. There is no entity E such that if this proposition is true then E exists. That's because the truth of the proposition is multiply realisable. For instance, suppose that P is actually made true by the existence of Sally the feisty penguin. Even if Sally actually makes the claim true, there is nevertheless a world, w , in which Sally does not exist (or is not feisty) and in which P is true, because there exists a distinct feisty penguin: Steve. That, however, flatly contradicts [L5*]. Notice, however, that if Steve or Sally exist, then P is true and necessarily so. So [L5] still holds.

However, there is another biconditional generalisation in the area which *does* appear to hold of metaphysical necessity. Consider:

[L5**] Necessarily, for any entity E , E exists just in case the proposition $\langle E \text{ exists} \rangle$ is true.¹⁰⁴

This principle has no counterexamples and allows for the derivation of the existence of E from the truth of $\langle E \text{ exists} \rangle$:

(1) $\langle \text{Socrates exists} \rangle$ is true.

[L5**] Necessarily, for any entity E , E exists just in case the proposition $\langle E \text{ exists} \rangle$ is true.

Therefore,

(2) Socrates exists.

Is [L5**] a law? Nothing in the account of metaphysical laws rules it out, and it is surely no accident. Arguably, neither does the presence of [L5] as a nearby, possibly more general, law that also tells us about the relationship between truth and existence, rule it out. So, there is a symmetry problem for a restricted class of truthmaking cases that cannot be resolved using the resources I have developed thus far (though it can, of course, be resolved via psychologistic filtering).

However, matters are more straightforward in the mereological case, where once again considerations of multiple realisability apply. Suppose that an object, W , is

¹⁰⁴ Similar considerations apply, *mutatis mutandis*, for [L5***]: Necessarily, an entity of kind K exists just in case $\langle \text{an entity of kind } K \text{ exists} \rangle$ is true.

composed of parts $P_1...P_n$. Now, consider the following two putative metaphysical laws governing composition:

[L2] Necessarily, for any entities $P_1...P_n$, if the P_n exist and relation R obtains between the P_n , then W , the object composed by the P_n , exists.

[L2*] Necessarily, for any entities $P_1...P_n$, the P_n exist and relation R obtains between the P_n , just in case W , the object composed by the P_n , exists.

Both [L2] and [L2*] can be used to formulate a DN argument to cover cases where the existence of the parts explains the existence of the whole. [L2*], however, would, if used, also give rise to the symmetrical explanation. But there are reasons to prefer [L2]: if [L2*] is true, then so is mereological essentialism, the view that composite objects necessarily have the parts that they in fact have. For what [L2*] tells us is that, for any object W composed of parts $P_1...P_n$, necessarily if W exists then so do the P_n . So, for instance, my bicycle, which is composed of certain parts: wheels, handlebars, gears, brakes, spokes and so on, would not exist if one of those parts didn't. Mereological essentialism is an implausible view, however; the intuition that a bicycle would still be the same bicycle, absent a spoke or two, is hard to shake.¹⁰⁵ But while mereological essentialism seems false, it nevertheless seems correct to say that a given set of parts under a composition relation necessitates the existence of a particular whole. So [L2] is acceptable.

The remaining three cases, involving as they do certain *kinds* of properties, do not suffer from the symmetry problem as characterised above. However, these are also derivations which run contrary to our explanatory intuitions by running 'the wrong direction', and thus are naturally grouped with the symmetric cases. We can explicate this nearby symmetry-like problem via a discussion of the case of categorical and dispositional properties, but similar considerations apply, *mutatis mutandis*, for the cases of physical and mental properties, and natural and moral properties. Consider the biconditional variant of [L6]:

[L6*] Necessarily, for any entity E , E exists and possesses a categorical property of kind K , just in case E possesses the dispositional property of being fragile.

¹⁰⁵ While it is a minority view, those who believe that composition is identity will also be friendly to the biconditional law. For defences of composition as identity see Wallace (2011a, 2011b) and Bohn (2009, 2014).

[L6*] appears to make the cut according to all of our criteria for a metaphysical law. However, it does not license a symmetrical derivation, as the following argument is invalid due to considerations of multiple realisability:

(1) The champagne glass is fragile.

[L6*] Necessarily, for any entity E , E exists and possesses a categorical property of kind K , just in case E possesses the dispositional property of being fragile.

Therefore,

(2) The champagne glass has crystalline bonds between its component molecules, and,

(3) Crystalline bonds are of kind K .

However, we can use [L6*] to form a valid derivation which runs contrary to our explanatory intuitions:

(1) The champagne glass is fragile.

[L6*] Necessarily, for any entity E , E exists and possesses a categorical property of kind K , just in case E possesses the dispositional property of being fragile.

Therefore,

(2) The champagne glass has a categorical property of kind K .

Kind K presumably is the class of categorical properties which give rise to the disposition of fragility. Thus, while we *have not* derived the fact that the champagne glass has crystalline bonds from the fact that it is fragile, we *have* derived that it must possess one of the properties of kind K . This is, intuitively, the wrong way around—the glass is fragile *because* it has one of these K -properties (namely, crystalline bonds). However, the argument satisfies all of the conditions we have placed on our DN theory, and cannot be ruled out using the resources I have developed thus far (though, once again, it could be ruled out via psychologicistic filtering).

Another problem in the vicinity of asymmetry is that of *contraposition*.

Contraposition does not require biconditional laws. It merely uses the *modus tollens* (rather

than *modus ponens*) argument form in combination with the conditional laws identified. For example, the following is formally valid:

(1) Action X is not wrong.

[L1] Necessarily, for any action ∂ possessing natural property of kind K , if ∂ exists then ∂ is wrong.

Therefore,

(2) Action X does not possess a natural property of kind K .

This contraposed use of [L1] generates an intuitively unexplanatory derivation: that action X lacks a particular moral property does not seem to explain why it has no natural properties of kind K .

The discussion of this subsection has shown that there are biconditional generalisations which hold of metaphysical necessity and which allow DN derivations in both directions, when our intuitions indicate that only one of these derivations is explanatory. What to do? Well, one option is to take our cue from the grounding literature and simply stipulate, on the basis of our explanatory intuitions, that the metaphysical laws must be asymmetric, and that the direction that suits our intuitions is the correct law. Such a stipulation is not open in the scientific case, because, as already noted, there is substantial evidence that the physical laws possess deep symmetries. As also noted, however, to make such stipulations on the basis of our intuitions of explanatory priority is to objectionably project our anthropocentric judgements into the world—it is to engage in illicit pre-filtering.

Thus, I do not recommend this stipulation. [L4*], for example, looks precisely like the metaphysically necessary generalisation which governs the relationship between sets and their urelements, and we should accept that this is the relevant metaphysical law. To be sure, this leaves us with no way, short of psychologistic filtering, of ruling out the symmetrical explanations. In other words, the DN theorist of metaphysical explanation faces a dilemma: she cannot, via licit manoeuvres, achieve both the objectivity desideratum and the asymmetry desideratum.

6.4.4 Relevance & Understanding

Like asymmetry, relevance and understanding have been perennial problems for DN theories (see §2.5.1 and Salmon, 1971). This is no different for the metaphysical variant. It is useful to consider these desiderata in tandem, for it is often the same putative derivations which fail to achieve both. The problem is that arguments with premises intuitively irrelevant to their conclusions can still abide by all the constraints thus far placed on metaphysical explanations. Consider the following argument:

(1) 2 exists.

[L4*] For any entity E , E just in case the singleton set $\{E\}$ exists.

(2) My left foot exists.

Therefore,

(3) $\{2\}$ exists.

Clearly, premise (2) is irrelevant to the conclusion. Yet, the derivation satisfies all of the conditions of adequacy. Fortunately, the fix is quite simple in this case. We can call upon the standardly appended condition of *minimality*. Minimality insists that it is not just lawful premises that must be essential to the derivation of the conclusion: the non-lawful premises must be essential also. Thus premise (2) is forcibly suppressed from the above argument. As it does not help with the entailment structure of the argument, it is identified as an idle wheel.¹⁰⁶

However, not all relevance problems are so easily dealt with. Trickier problems emerge as soon as we allow that some entities exist of necessity. Interestingly, this implies an interaction between one's background views and the extent of the relevance problem for one's DN theory of metaphysical explanation. To see this, consider that for mathematical Platonists such as Hale and Wright (1992), who believe that numbers exist (and exist at every possible world), the following is a candidate to be a metaphysical law:

[L7] Necessarily, for any entity E , if E exists then N exists, for any number N .

¹⁰⁶ The minimality constraint is not always intuitive. It denies, for example, that the fact that James exists and the fact that Reginald exists can both feature in an explanation of why at least one human exists. Nevertheless, James and Reginald can each enter into separate explanations of that fact.

Similarly, for those who follow Armstrong (1997) in thinking that universals exist necessarily, the following appears to be a metaphysical law:

[L8] Necessarily, for any entity E , if E exists then U exists, for any universal U .

[L7] and [L8] can be used to generate irrelevant and unilluminating derivations. Observe the apparent explanatory power of my left foot!

(1) My left foot exists.

[L7] Necessarily, for any entity E , if E exists then N exists, for any number N .

Therefore,

(2) 2 exists.

(1) My left foot exists.

[L8] Necessarily, for any entity E , if E exists then U exists, for any universal U .

Therefore,

(2) Redness exists.

Neither of these derivations strikes us as explanatory.¹⁰⁷ Of course, one way to avoid these cases is to deny that there are numbers or universals, or deny the necessary existence of such things. As noted, for those who do not believe in any necessarily existing entities, the problem of irrelevance shrinks substantially. However, for those already committed to numbers or universals, perhaps it is better to insist that the following generalisations are the correct metaphysical laws:

[L7*] Necessarily, for any entity E , if E is a number then E exists.

[L8*] Necessarily, for any entity E , if E is a universal then E exists.

Then we get the more palatable derivations:

(1) 2 is a number.

[L7*] Necessarily, for any entity E , if E is a number then E exists.

¹⁰⁷ I should note, perhaps, that my left foot is not red.

Therefore,

(2) 2 exists.

(1) Redness is a universal.

[L8*] Necessarily, for any entity E , if E is a universal then E exists.

Therefore,

(2) Redness exists.

How can we mediate in favour of [L7*] and [L8*] over [L7] and [L8]? They are all conditional generalisations with ontological content that hold of metaphysical necessity. One way to rule out [L7] and [L8] via an appeal to what the metaphysical laws *are* is to fall on the distinction between accidental and non-accidental generalisations considered in §6.2. The claims made by [L7] and [L8] describe intuitively accidental generalisations, due to the modal status of numbers and universals, rather than some deep relationship between these entities and any arbitrary existent. However, as was noted in §6.2, appealing to the accidental/non-accidental distinction to constrain the metaphysical laws would require giving an account of that distinction, and that is no easy task. In particular, if we want to avoid anthropocentric considerations creeping in the back door of our putatively objective theory, we have few resources with which to rule out metaphysical laws like [L7] and [L8].

Furthermore, one might think that there are no *truly* accidental generalisations that hold of metaphysical necessity. That is, while there is good reason to think that some *actual* regularities are merely accidental, when a generalisation holds across every possible world we ought to take it more seriously. Consider that if, actually, every human had blue hair, we would likely deem this an accidental regularity. If, however, every *possible* human had blue hair, *that* would be far less easy to dismiss as an accident: it would seem to imply that something in the essence of being human ensures that all possible humans have blue hair. If it is not even possible for a human not to have blue hair, the link between humanness and blue hair starts to look very law-like indeed. To be sure, some metaphysically necessary generalisations *strike us* as accidental. For instance, the metaphysically necessary generalisation that if there exists a number there exists a

universal. But such intuitions should enter into our theory only once we admit filtration. For only then can we be sure we are not simply projecting into the world the kinds of relationships we happen to find illuminating.

One possible avenue for the unfiltered metaphysical DN theorist to explore appeals to counterfactual and counterpossible truths. For, while [L7] and [L8] are true generalisations, it is notable that the existence of numbers and universals does not counterfactually depend on the existence of my left foot, nor upon any contingently existing entity. This lack of counterfactual dependence lends some support to the thought that our intuitions about accidental generalisations might not be so anthropocentric.

This counterfactual manoeuvre is not so simple, however, when the entity doing the putative explaining also exists of necessity—counterfactuals with necessarily false antecedents are trivially true according to Lewis' (1986b) semantics. While we can develop semantic theories of counterfactuals that don't trivialise in this way when the antecedent is a necessary truth, the motivation for these alternative semantics stems from intuitions about *impossibility*. As such, it may be that we use certain intuitions about impossibility as a guide to judgements about metaphysical laws. Consider that, as a matter of fact, both 2 and {2} exist at every possible world, if they exist at all. Nonetheless, intuitively, in the impossible world where 2 does not exist, neither does the singleton set {2}. Compare this with the universal case. Universals and sets of numbers exist at every possible world. Intuitively, though, in the impossible world where redness does not exist, the singleton {2} exists nonetheless. While it is a tricky matter to render such counterpossible intuitions semantically precise, it is clear that we have intuitions of this kind and that they might be what is clueing us in to which generalisations count as laws (but it would seem that this only helps in a very narrow range of cases).

We should be wary, however, in the context of developing an unfiltered theory, of placing too much weight on our explanatory and counterpossible intuitions. Indeed, the counterpossible intuitions, while apparently more respectable, are plausibly merely a reiteration of the kneejerk 'we know 'em when we see 'em' intuitions, or indeed of the priority intuitions themselves. Thus, I contend that, in lieu of a convincing non-anthropocentric account of the metaphysical laws which excludes the intuitively accidental generalisations, the unfiltered DN theorist must accept that there are

metaphysical explanations that are intuitively irrelevant and fail to enhance our understanding.

Thus, once again the choice is between a theory that retains objectivity, yet delivers counterintuitive verdicts regarding apparently irrelevant and unilluminating explanations, or a theory that jettisons objectivity in favour of filtering out these unwanted explanations.

6.4.5 Summary

Thus concludes my characterisation of a classical DN theory of metaphysical explanation, and how the filtered and unfiltered variants fare according to the various desiderata. I hope to have shown the plausibility of such a theory, by showing how one might characterise the relevant notion of a metaphysical law, and how formal entailment can ensure that these laws are not suppressed.

As I mentioned in the introduction to this chapter, the primary goal here is to build a theory that satisfies the ontic desideratum of objectivity, and then see how it fares according to the epistemic desiderata. Having now completed this evaluation, we can see that while the unfiltered theory has the power to rule out reflexive explanations, we are still left with symmetrical explanations, irrelevant explanations and explanations that fail to increase our understanding. Thus, the following are the options for the DN theorist.

First, she might decide that, given how she weights the desiderata, the unfiltered DN theory is apt for her purposes, despite failing to fulfil some desiderata that others care about. Second, she might appeal to psychologistic filtering of the kind I elucidated in Chapter 5. In that case, the theory gives up on objectivity, and gains all of the epistemic desiderata in return. Thus it has the same profile, desiderata-wise, as the filtered grounding-based theory and the filtered modal relations theory. Those who think that laws ought to play a role in explanations will prefer the filtered DN theory to the filtered modal relations theory, but there is no standpoint from which the filtered grounding-based theory is preferable. Third, she might seek a non-anthropocentric account of the difference between accidental and non-accidental modal correlations that rules out some of the metaphysical laws allowed by the account I develop above. If this could be done, then relevance might be achieved, but symmetry problems (and the associated problems

with understanding) will remain, for the biconditional laws underpinning the symmetrical derivations are surely no accident.

I now move on to critically consider two DN theories of metaphysical explanation that have been proposed in the extant literature.

6.5 deRossett's Theory is Not a DN Theory

Now that we have a DN theory in hand, we can quickly show that deRossett's (2013) early foray into this terrain fails to make use of the resources offered by the DN theory: he denies the first and second condition of adequacy. While deRossett takes himself to be offering a DN theory, his only substantial agreement with Hempel and Oppenheim is that "It is plausible to think that a true explanation must be backed by an argument from *explanans* to *explanandum*" (2013:12). Yet, entailment and laws play no role in deRossett's arguments: "I am not, however, signing on to the deductive-nomological account in detail; in particular, I assume neither that the arguments in question are *deductive*, nor that they are *nomological* – laws need play no special role." (2013:footnote 36). This, presumably, is because deRossett's intention is to supplement a grounding-based theory with DN-ish *explanatory stories*, as opposed to crafting a viable alternative to grounding-based theories, which is the primary goal of this thesis.

Let's take a look at an example. Assume that B grounds A. deRossett's schema of an explanatory story goes as follows:

B

X1

X2

...

S₀, A

The Xs, deRossett tells us, are 'ancillary material'. They are explicitly not metaphysical laws, and neither are they part of the grounds for A. Yet they help to form an explanatory story linking B to A. We are invited to imagine that the disjunctive fact that [It is either chilly or windy] is grounded in the fact that [It is chilly]. deRossett claims that

the explanatory story will include not just the fact that it is chilly (the grounds), but also some ancillary material which makes the explanandum *intelligible*, given the explanans. In this case, deRossett tells us that this material “might include the observation that the *explanans* is a disjunction, a review of the truth table for disjunction, or a reference to a rule of disjunction introduction.” (2013:13). While the ancillary material looks like it plays the law-role in deRossett’s schemas, he rejects this thought.

Consider that deRossett steadfastly defends necessitarianism about metaphysical explanation, asserting that “if a fact P is completely explained by Q, then it is impossible that Q obtain and P not obtain” (2013:14). He further claims that “In general, an explanation has modal force: it’s supposed to indicate what *makes* the *explanans* obtain. If it is possible for the *explanandum* to obtain while the *explanans* does not, then the *explanandum* does not make the *explanans* obtain, and so the explanation fails.” (2013:15). Given this position, we might be surprised that he declines to sign up for laws and deduction. These are, after all, powerful tools with which to secure the desired modal force.

Yet, within deRossett’s theory, the modal force is secured by grounding relations, and the ancillary material is merely intended to assist with the understanding desideratum by adding something beyond the grounds to the explanantia. As such, this view is perhaps better characterised as an extension of the grounding-based theories considered in Chapter 3, rather than as a strange DN theory that jettisons both the D and the N!

6.6 Wilsch’s DN Theory

Tobias Wilsch (2015, 2016) has developed a DN theory that competes with the DN theory I developed above. However, Wilsch’s project has a significantly different starting point, insofar as his account is intended as a reductive account of grounding, which competes with the standard primitivist accounts presented in Chapter 3. That is, while he posits grounding relations, he rejects the majority view that those relations are primitive and unanalysable. His project is to reduce grounding to a relation of lawful determination, involving derivation in accordance with the metaphysical laws. To return to the analogy with scientific explanation: the goal of this chapter is analogous to providing a DN theory of scientific explanation. In contrast, Wilsch’s project is

analogous to reducing causation to a relation of nomic dependence.

One might think that this difference would evaporate once the grounding relation, so reduced, is then used to build a theory of metaphysical explanation, but Wilsch's theory inherits some of the criticisms levelled at grounding in Chapter 3, and ends up having to call upon primitive dependence relations, which Wilsch calls 'construction relations'. Thus there is a substantial difference between our projects. Nonetheless, as the only published DN theory of metaphysical explanation, the account is worth exploring in some detail. The following subsections will spell out and subsequently criticise Wilsch's proposal regarding metaphysical laws and his non-classical notion of 'grounding-entailment'.

6.6.1 The Constructional Conception of Metaphysical Laws

Wilsch tells us that the metaphysical laws guide the development of the world from the fundamental to the derivative in much the same way as the natural laws guide the world through the temporal dimension. In other words, they work 'up' his axis of fundamentality. The "important but initially elusive" (2015:3298) metaphysical laws play a pivotal role in Wilsch's theory, and thus his conception of them must be spelled out in detail. To begin with, the laws are supposed to satisfy three constraints: strength, generality and parsimony.

The *strength* constraint is simply that the laws, in combination with the fundamental truths, completely determine the derivative state of the world. There is no indeterminism as we move our way up the axis of fundamentality. The *generality* constraint is that the laws should be as simple as possible and apply to the broadest range of instances. This, presumably, is motivated by considerations of parsimony. For instance, a general law connecting determinate properties to a determinable property is more desirable than a collection of laws connecting, say, having colour X to being coloured, having colour Y to being coloured, having mass X to having mass, etc. Wilsch says that:

To ensure simplicity, the generality-constraint requires that the laws are generalizations that mention only a few select properties and objects. The generality-constraint poses a challenge for the proponent of metaphysical laws: the laws 'capture' every object and property and yet they only feature a

special minority of them ‘by name’. To state the laws we thus need some properties that can be used to pick out every object and property; and these select properties need to have an appropriately special status that makes them eligible for entering into the laws directly. (2015:3298).

Finally, the *modality* constraint is that the metaphysical laws must hold with metaphysical necessity. This sounds akin to a characterisation of laws as exceptionless generalisations about modal space, yet Wilsch leaves open that “metaphysical necessity is a restriction on logical necessity: the space of metaphysically possible worlds is defined (in part) as the space in which all the actual metaphysical laws hold” (2015:3299). So, on Wilsch’s view, it might be that metaphysical explanations do not (unrestrictedly) necessitate the truth of their explananda.

As such, Wilsch cannot characterise the laws as metaphysically necessary generalisations (for, recall, metaphysical necessity, in this thesis, is understood as the broadest modality). Instead, he tells us that “the laws of metaphysics are general principles that characterize the individual construction–operations” (3301). That is, the metaphysical laws stem from the various *construction relations* (these are the dependence relations with which grounding, if it exists, is intimately related: composition, constitution, set formation, realisation, etc. See Bennett’s (2011b) similar notion of *building relations*). We can understand each of these relations to imply a ‘construction operation’—notated as O—corresponding to the mechanisms by which the various relations take on more fundamental existents and crank out derivative ones. The composition relation makes for a relatively straightforward example. If object A is composed of $B_1 \dots B_n$, this can be expressed as the following:

$$A = \text{COM} (B_1 \dots B_n)$$

A is the result of the construction operation on $B_1 \dots B_n$. Similarly, the operation corresponding to the constitution relation can be expressed as:

$$A = \text{CON} (B)$$

An instance of this second operation, where the clay constitutes the statue, is ‘Statue = CON (Clay)’.

However, not just any old object or property can enter into these operations. Thus the

metaphysical laws work to constrain the construction operations. Wilsch divides the laws into two categories: *Ontological Principles* and *Linking Principles*.

The Ontological Principles determine which derivative objects and properties exist on the basis of more fundamental ones. Thus they constrain, for example, when some entity constitutes another or when some properties combine to form a conjunctive property. Each Ontological Principle will be of the form:

$$\text{(Ontological Principle)} \quad \forall_{xx}(C_o(xx) \supset \exists y(y = O(xx)))$$

C_o is the ‘application-condition’ for a certain construction relation; O is the construction operation associated with that relation. Thus, the above reads that for all xx s, if they satisfy C_o , then there exists some y that is the output of O on the xx s. Each determinate variant of this application conditional, with C_o specified, is an ontological principle. Ontological principles represent how construction relations do such work as constructing composites from their parts, constructing conjunctive properties like bachelorhood from their conjuncts (male and unmarried, in this case), and constructing statues from clay. Wilsch provides no characterisation of any particular C_o ; perhaps it depends on how one characterises the relevant construction relations. For instance, compositional universalists, restrictivists and nihilists will have very different views about the C_o that governs the composition relation.

In this way, one can imagine the application conditional associated with composition to be rather like [L2] above, insofar as [L2] claims that if a group of entities satisfied condition R then there exists a further entity that they compose. There, as here, the precise nature of R was left unspecified, for such a specification looks like a job beyond a theory of metaphysical explanation.

Moving on, Wilsch’s Linking Principles determine when these constructed, derivative, objects and properties combine as facts. Wilsch claims that these two kinds of laws/principles “work as a team: ontological principles determine the derivative ontology, and linking-principles determine the derivative facts.” (2015:3302). The Linking Principles are intended to build more complex facts/truths by linking together constructed objects and properties. The Linking Principles are supposed to show how constructed objects inherit their properties from the objects which construct them. Such things are best demonstrated by example. Here is Wilsch’s Linking Principle for

conjunctive property construction (2015:3304):

(Conjunction) $X = C_{\&}(Y_1 \dots Y_n) \& Y_1(x) \& \dots \& Y_n(x) \supset X(x)$

Notice that the Linking Principle is built on an Ontological Principle: properties $Y_1 \dots Y_n$ have satisfied $C_{\&}$, and thus generated the conjunctive property X . Furthermore, object x instantiates each of the $Y_1 \dots Y_n$. Thus x instantiates X . For example, perhaps the properties of being male and being unmarried have satisfied $C_{\&}$ and thus built the property of being a bachelor, so given that John is male and unmarried, John is a Bachelor (sidestepping, here, concerns about eligibility, etc.).

Another example provided is the Linking Principle for the construction of determinable properties. This principle states that to instantiate a determinable property an object must instantiate one of the determinate properties from which the determinable property is built. Where 'DC' indicates determinable-construction:

(Determinables) $\exists ZZ(X = DC(Y, ZZ)) \& Y(x) \supset X(x)$

In this case, the Ontological Principle of Determinable Construction (DC) has identified that there is a determinable property, X , constructed from Y and the ZZ s. Furthermore, object x instantiates Y . Thus x instantiates X . Perhaps being Y is the property of being red and the ZZ s are all the other determinate colour properties. So it follows from the fact that the ball is red that the ball is coloured. Again, Wilsch does not engage with the first order question of when some collection of determinate properties constructs a determinable property. One can simply substitute one's views on this question into the schema.

Wilsch's final example of a linking principle is for the construction of internal relations from pairs of monadic properties. This principle states that if a pair of objects each instantiates a property used for the construction of the relation (say, the relation of *being heavier than*), then the internal relation between them obtains. Where 'IRC' indicates internal-relation construction:

(Internal Relations) $\exists ZZ(R = IRC(\langle X, Y \rangle; ZZ)) \& X(x) \& Y(y) \supset R(x, y)$

Here, the Ontological Principle of IRC has identified that there is an internal relation R constructed from pairs of properties, including the ZZ s and $\langle X, Y \rangle$. Furthermore,

object *x* instantiates *X* and object *y* instantiates *Y*. Thus, the internal relation *R* obtains between *x* and *y* because of properties *X* and *Y*. So, if *X* is the property of weighing 60kg and *Y* is the property of weighing 50kg, and this ordered pair is one of the (many!) pairs that construct the ‘heavier than’ relation, then it follows from the fact that *x* is 60kg and *y* is 50kg that *x* is heavier than *y*.

The combination of the Ontological Principles and the Linking Principles (not exhaustively listed by Wilsch) are the metaphysical laws. With both in hand we are supposed to be able to derive the derivative ontology from the fundamental.

Let’s pause our Wilsch exposition here, and evaluate what has been said so far. Firstly, there’s the idea (that Wilsch neither endorses nor denies) that metaphysical necessity could be a restricted modality. That is, metaphysical possibility could simply be possibility delineated by accordance with the metaphysical laws, and the laws could have been otherwise (there are logically possible worlds with different metaphysical laws). Clearly, this runs against the idea that a metaphysical explanans ought to (unrestrictedly) necessitate its explanandum. The metaphysical laws are supposed to explain why certain relationships hold of necessity, but if the relevant sense of necessity is defined in terms of those laws, this looks unhelpfully circular. Regardless, Wilsch is not explicitly committed to a restricted version of metaphysical possibility, so we can leave this point here.

Secondly, and more importantly, Wilsch is trying to develop a reductive account of grounding, and yet his laws draw indispensably upon a class of so-called construction relations. Plausibly, what grounding *is*, if it exists, is a catch-all dependence relation that obtains whenever an instance of one of these more determinate relations obtains: grounding *unifies* these relations.¹⁰⁸ So it’s not clear what the primitive notions are, here. Grounding is supposed to be reduced to lawful entailment, but the requisite laws draw upon further primitive relations, which are themselves intimately tied to grounding.

As Wilsch seems content to draw the metaphysical laws from the construction relations, it seems that these form the primitive components of his view. For, it is from these relations that the laws derive, and from the laws (in combination with the notion of entailment discussed below) that grounding derives. The problem is that this opens up Wilsch’s view to the very same criticisms we levelled at the unfiltered grounding-based theory of metaphysical explanation in Chapter 3: there is illicit pre-filtering at play, here.

¹⁰⁸ Raven (forthcoming) argues against understanding grounding as this kind of unifier.

To see why this is so, consider that if you call upon the various dependence/construction relations to constrain the metaphysical laws, you had better have a story about why the modal correlations captured by your preferred class of relations are the ones worth generating laws from. Let's use the asymmetry desideratum, as applied to singleton sets and their urelements, as an example. The view we built, above, ran into some problems ruling out biconditional laws that license symmetrical derivations. For instance, it seems that the law that governs the relationship between singleton sets and their urelements is [L4*]:

[L4*] For any entity E , E just in case the singleton set $\{E\}$ exists.

Wilsch's view will not run into this problem. For, he can say that [L4*] does not stem from a construction relation, and is thus not a metaphysical law. Instead, something like [L4]—the law that only licenses the derivation of the existence of the set from the existence of the member—is a metaphysical law, for that generalisation stems from a construction relation. Let's call that relation *singleton set formation*. The singleton set formation relation is asymmetric: it runs *from* the urelement *to* the set. As the law is derived from this relation, it inherits this asymmetry and can thus help to secure the asymmetry of the theory of metaphysical explanation.

But this is just to move the bump in the carpet. For now the relevant question concerns why there is no construction relation that runs from sets to their urelements—no relation of *urelement formation*, with its associated law. Wilsch, presumably, believes in no such relation. The basis for this disbelief, however, is surely the priority intuitions. We don't intuit that there is a construction relation running from sets to their members because we intuit that the member is prior to the set. But, if we're seeking a non-anthropocentric vindication of these priority intuitions, said vindication had better not come from a theory that has already drawn metaphysical conclusions from these intuitions! How can a theory that takes as primitive that there are relations of singleton set formation but no relations of urelement formation provide the vindication we seek? For, if this assumption is made from the outset the move becomes directly analogous to the illicit pre-filtering involved in the unfiltered grounding-based theory.

In sum, it seems that anthropocentric priority intuitions are doing a lot of tacit work in this theory. For these intuitions underlie the assumptions about what construction relations there are, which in turn underlie what metaphysical laws there are,

which in turn underlie what metaphysical explanations there are. What is needed from Wilsch, in order to avoid this charge, is a story about why the urelement formation relation fails to make the cut as a construction relation, with its associated Ontological Principle which builds urelements from their singleton sets. The answer that ‘that’s not a construction relation’ will not do: we want to know *why* some modal correlations imply construction relations whilst others do not.

6.6.2 Grounding-Entailment

On the path to the DN theory he ultimately defends, Wilsch (2016) considers and discards another. It will be instructive for us to follow this path. His first pass at a DN theory is as follows (where \vdash_L is interpreted as ‘entails together with the metaphysical laws’):

$$(DN_1) p_1 \dots p_n \text{ lawfully determine } q \leftrightarrow \text{def } p_1 \dots p_n \vdash_L q$$

If this is right then all it takes for $p_1 \dots p_n$ to explain (and, for Wilsch, ground) q is that they entail q , on the assumption that all the metaphysical laws hold. However DN_1 is too weak for several reasons, which will be familiar from the development of our DN theory above. Firstly, given that the classical entailment relation is reflexive, the above implies that p explains p , as p entails itself. In other words, DN_1 fails to respect the motivation behind the second condition of adequacy: that the laws must be *required* for the derivation. Clearly enough, p entails itself without the help of any laws. So, DN_1 as it stands fails the irreflexivity desideratum. Secondly, the non-symmetry of classical entailment allows that if p entails q and q entails p then there is symmetrical explanation between the two, and Wilsch is keen to rule this out. Thirdly, the monotonic nature of this entailment relation allows that extra content can be added to a true explanatory claim and it will retain its truth, so if p explains q then $p \ \& \ r$ explain q also. Finally, DN_1 allows contraposed derivations using the metaphysical laws. This is the same problem that we addressed in §6.4.3. That is, while

$$(\text{Ontological Principle}) \quad \forall_{xx}(C_o(xx) \supset \exists y(y = O(xx)))$$

allows us to derive the existence of $O(xx)$ from the fact that the xx s satisfy C_o , it also, via *modus tollens*, allows us to derive from $\sim(\exists y(y = O(xx)))$, that $\sim C_o(xx)$. In other words,

this law allows the derivation of the existence of the composite from the fact that the parts stand in the right relation, *and* the derivation that it is *not* the case that the parts stand in the right relation from the fact that the composite does *not* exist. The latter derivation runs against our explanatory intuitions.

These four problems lead Wilsch to reject DN_1 as an adequate analysis of metaphysical explanation. As we shall see, his preferred theory is labyrinthine in comparison to both DN_1 and the theory we developed above. To deal with these problems, Wilsch proposes that we take seriously the idea that the laws “*apply* to some truths and *take them* to some other truths” (2016:9). That is, the job of the metaphysical laws is to guide the development of the world up through the levels of fundamentality, *and not in the other direction*. Wilsch thinks there is a helpful analogy here with the physical laws. Just as the physical laws take us from a state at one time to a state at the *subsequent* time, so the metaphysical laws only work *up* the axis of fundamentality. But that’s not a very helpful analogy for Wilsch. As we have repeatedly noted, there are deep symmetries in the physical laws which allow for DN derivations in either temporal direction.¹⁰⁹ Thus it would perhaps be better for Wilsch to say that *unlike* the natural laws, the metaphysical laws apply in only one direction. In this way, the analogy appears to undercut Wilsch’s position more than it provides support. Moreover, the idea here seems to be tantamount to injecting the priority intuitions into the bedrock of the theory.

Wilsch claims that *logic* offers a way to make sense of this idea of the metaphysical laws being unidirectional, and thus ruling out their use for reflexive, symmetrical and contraposed derivations. But he doesn’t have classical logic in mind, or any logics you might have encountered before. Rather, he handcrafts a custom logic, sensitive to the priority intuitions, to govern metaphysical explanation. He explains that “a law expressed by a sentence of the form ‘ $\forall x(\varphi(x) \supset \psi(x))$ ’ applies to a truth, p , just in case p witnesses the antecedent of ‘ $\forall x(\varphi(x) \supset \psi(x))$ ’. And p witnesses the antecedent of ‘ $\forall x(\varphi(x) \supset \psi(x))$ ’, if p is identical to the antecedent of a conditional that can be derived from ‘ $\forall x(\varphi(x) \supset \psi(x))$ ’ via \forall -elimination.” (2016:9).

¹⁰⁹ Indeed, rather ironically in this context, some, like Maudlin (2007) use something very much like diachronic grounding (cross-temporal ontological dependence, if you will) in order to achieve temporal asymmetry. The idea is that although the physical laws are time-reverse invariant, nevertheless there is a temporal direction because there is a direction to the ontological priority: later times depend on earlier times, but not the reverse.

This is just to say that a law applies to a truth p just in case p is one of the ways in which $\varphi(x)$ can be instantiated. In other words, a metaphysical law applies to a truth just in case that truth is an instance of the *antecedent* of a law with the universal quantifier stripped out. If p is an instance of the *consequent* condition then that law is not apt to be applied. The effect of this restriction is that the *modus tollens* argument form is not valid. In essence, the conditionals featuring in the metaphysical laws that underpin the entailment in metaphysical explanations can only be used for their implications in one direction.

Thus the non-classical notion of ‘grounding-entailment’ (\vdash_G) is defined: an entailment relation whereupon the only valid rules of inference are \forall -Elimination and *modus ponens*. Let’s play with Wilsch’s example, beginning with the putative law in the quote above:

$$(1) \quad \forall x(\varphi(x) \supset \psi(x))$$

From the general claim at (1), we can infer something about individual a , by \forall -elimination:

$$(2) \quad \varphi(a) \supset \psi(a)$$

Now, if we assume:

$$(3) \quad \varphi(a)$$

We can infer, via *modus ponens* from (2) and (3):

$$(4) \quad \psi(a)$$

However, if we assume:

$$(3^*) \quad \sim \psi(a)$$

We *cannot* infer, via *modus tollens* from (2) and (3*):

$$(4^*) \quad \sim \varphi(a)$$

because *modus tollens* is not a valid rule of inference according to grounding-entailment.

As \forall -elimination and *modus ponens* are the only valid rules, grounding-entailment will only allow reflexive explanation of p if it is a law that $p \supset p$. Thus, as this putative law presumably does not stem from any construction relation,¹¹⁰ grounding-entailment will not deliver any reflexive explanations. Furthermore, as long as there are not two laws of the form $p \supset q$ and $q \supset p$, there will not be symmetrical explanations. Wilsch says:

I don't believe that we need any additional justification for using these inference-rules in the analysis of the input/output-operations of the laws. Given that the laws are universally quantified conditionals, the two rules simply constitute the best fit. Characterizing lawful determination in terms of \forall -elimination and *modus ponens* is therefore not inadmissibly *ad hoc*. (2016:9).

However, recall Wilsch's claim that *logic* offered him a way of making sense of the unidirectionality of the metaphysical explanations. While grounding-entailment can do so, grounding-entailment is hardly a gift from logic. It is a logical operator Wilsch has built from scratch to serve his purposes. Moreover, the resultant logical system is very different from those to which we are accustomed—in what other context is *modus tollens* invalid?

Wilsch's updated account is:

(DN₂) $p_1 \dots p_n$ lawfully determine $q \leftrightarrow \text{def } p_1 \dots p_n \vdash_{G,L} q$

This is to say that $p_1 \dots p_n$ explain q just in case, in conjunction with the laws of metaphysics, they grounding-entail q . The restriction to grounding-entailment prevents contraposition, for contraposed arguments use *modus tollens* form.

Wilsch then proceeds to supplement DN₂ with the requirement of *minimality* (see §6.4.4) to prevent the explanatory relation from being monotonic. In this way, DN₂ avoids the problems of reflexivity, symmetry, irrelevance/monotonicity and contraposition that plagued DN₁. He goes on to define up a DN₃ which includes mechanisms for introducing names for new derivative entities, but that need not concern us here (for the interested reader, see his 2016:11-12). The take-home message is that Wilsch's DN account ends up satisfying all of the epistemic desiderata by making use of a custom-built entailment relation which prevents the unwanted derivations by heavily

¹¹⁰ Identity, presumably, is not a construction relation.

restricting the rules of inference, and using only those metaphysical laws that can be gleaned from the so-called construction relations.

Thus concludes our exposition of Wilsch's nomological account of ground. So, once again we can stop and take stock. Wilsch has built a DN theory of metaphysical explanation which is supposed to serve as a reductive base for grounding. Wilsch sells the account as an improvement over those accounts which take grounding as primitive.

On the one hand, it's not obvious that Wilsch's account is really an improvement over grounding-based theories. As, on his view, metaphysical explanations can be cashed out in terms of notions we already understood, such as laws and entailment, Wilsch argues that there is an advantage that stems from parsimony and elegance. Yet, grounding-entailment is surely a unique piece of ideology (*not* a notion we already understood), and the construction relations with their associated construction operations and metaphysical laws constitute primitive entities. Moreover, plausibly, primitivism about the construction relations is simply a manifestation of primitivism about grounding. So, Wilsch's theory and those theories that take grounding as a primitive are, from my perspective, equally bad.

On the other hand, whether Wilsch's view is better than one that takes grounding as primitive isn't really our target, here, so we can put it to the side. What is of interest is how the account compares to the other theories of metaphysical explanation. On this front, I contend that the account is on a par with the unfiltered grounding-based theory of metaphysical explanation. That's because, while it putatively satisfies all of the epistemic desiderata while retaining objectivity, the foundations of the view naturally lead to suspicions about illicit pre-filtering.

In Wilsch's case, the illicit pre-filtering is two-fold. Not only have the laws been gleaned from those construction relations that are believed to exist on the basis of the priority intuitions, but also, the logical notion of grounding-entailment has been custom-built to license only those derivations about which we have explanatory intuitions. So the whole logical system in which the theory is embedded—from the laws to the very notion of entailment—has an illicitly pre-filtered, anthropocentric tinge. In both of these ways, there are reasons to think that psychologistic filtering is sneaking in the back door of the theory.

6.7 Conclusion

The primary purpose of this chapter has been to evaluate how far an unfiltered DN theory of metaphysical explanation can go towards satisfying the epistemic desiderata while retaining objectivity. The answer to this question depends on what one thinks of Wilsch's theory. If you don't think that my arguments against his view hit home—that is, you don't think the view is guilty of illicit pre-filtering in the characterisation of the metaphysical laws and the notion of grounding-entailment—then the view is very attractive, and able to fulfil every desideratum.

If, however, you agree with me that Wilsch's view does not, in fact, secure objectivity due to the illicit pre-filtering, then the prospects for an unfiltered DN theory depend upon the prospects for developing a non-anthropocentric account of the difference between accidental and non-accidental modal correlations. If such an account can be provided, then the unfiltered DN theory can achieve all of the desiderata but asymmetry (and understanding, insofar as those derivations 'in the wrong direction' fail to increase our understanding). However, if such an account is not forthcoming, the view will, of the epistemic desiderata, only achieve irreflexivity: it will fail on relevance, understanding and asymmetry.

Of course, there remains the option of psychologically filtering the DN theory, sacrificing objectivity in order to secure the epistemic desiderata. Moreover, both the filtered and unfiltered variants of this view have the potential to be unified with the filtered and unfiltered scientific DN theories respectively, thus generating a more general theory of explanation. In the next chapter, I will consider two non-psychologically filtered variants on the DN framework: Strevens' (2008) Kairetic Theory and Kitcher's (1981, 1989) unification Theory. While the former is not promising, the latter has the potential to go further than the plain DN theory I built in this chapter, in securing the epistemic desiderata.

Chapter 7: Extended DN Theories

We have reached the point in the thesis where I am exploring avenues through which an objective, yet grounding-free, theory of metaphysical explanation might be built. In the previous chapter, I developed what one might call a ‘plain’ DN theory of metaphysical explanation. That is just to say that I built a theory that was analogous to Hempel and Oppenheim’s (1948) early presentation of a scientific DN theory. The unfiltered incarnation of that theory secured the desiderata of covering cases, objectivity and irreflexivity, but failed to secure asymmetry, relevance and understanding. The psychologically filtered incarnation secured every desideratum but objectivity. This dialectic is directly analogous to the one we find when we consider scientific DN theories. Moreover, in that context, other ways of filtering the DN derivations have been developed—ways that do not require thinking of explanation in a psychologistic way, or giving up on objectivity. Thus, I will now explore the prospects of developing metaphysical analogues of two of these extended DN theories: Strevens’ (2008) Kairetic theory and Kitcher’s (1981, 1989) unification theory. Once again, I want to emphasise that this chapter is *exploratory*, and that the theories I sketch below are inchoate.

Here is the plan. In §7.1 I’ll introduce the Kairetic theory, and show how it requires no tweaking in order to allow for metaphysical explanation, but adds little to the plain DN theory. §7.2 considers the motivations for Kitcher’s (1981, 1989) unification theory as he presents it in the scientific context, followed by a careful explication of the nuts and bolts of the theory. §7.3 applies Kitcher’s central idea of *general argument patterns* to metaphysical explanation. §7.4 shows how the unificationist machinery can be of assistance in ruling out unwanted symmetrical and irrelevant metaphysical DN derivations. Finally, §7.5 concludes.

7.1 Strevens’ Kairetic Theory

In recent years, the DN theory has undergone something of a revival in the form of Strevens’ (2008, §2.5.1) Kairetic account. Of note is that the Kairetic account, even without any tweaking, explicitly allows for metaphysical explanations alongside the scientific explanations traditionally dealt with by DN theories. This is attractive, as using a single account to deal with both kinds of explanation is desirable on the basis of parsimony and unification. As we saw in §2.5.1, the Kairetic account adds a further

condition of adequacy to the four outlined in §6.1 (five, including minimality. See §6.4.4). Namely, the explanatory DN derivations are those which represent an underlying asymmetric relation of determination, such that the relata of said relation are represented by the non-lawful premises of the argument and the conclusion, respectively. As such, the theory is a hybrid. In the scientific context, for example, it is a hybrid DN-Causal theory.

It is clear how the further requirement demanded by the Kairetic theory fortifies the scientific DN theory against some key objections. While the physical laws may be time-reverse invariant—and thus allow DN derivations in both temporal directions—as long as causal relations only run in one temporal direction, explanation can retain its temporal asymmetry. Likewise for laws that link apparently irrelevant states of affairs. If there is no causation, there is no explanation. In this way, the hybrid theory can do some heavy lifting to help the DN theory secure the desiderata. It is not without cost, however. One must believe in, and provide an account of, both laws of nature and causal relations in order to get the account up and running (though, as noted in Chapter 2, Strevens denies the need for an account of the latter).

Unfortunately, in the metaphysical context the Kairetic view is not only expensive, but also fails to do the heavy lifting. For, the view is that a putative DN metaphysical explanation only counts as such if there is an underlying determination relation linking that which is represented by the explanans to that which is represented by the explanandum. Unfortunately, this further condition of adequacy provides little assistance. Let's consider the various relations that might be represented. On the one hand, the Kairetic theorist might insist that explanatory derivations must represent underlying *grounding* relations. Yet, I have already argued against positing grounding, and the goal of the present chapter is to see how far we can get with an objective theory of metaphysical explanation that doesn't appeal to grounding. Thus, I won't explore this move here.

Alternatively, perhaps the underlying relations are those relations grounding is supposed to unify (Wilsch's (2015, 2016, §6.6) 'construction relations', or Bennett's (2011b) 'building relations'). Then, like Wilsch, we are in a position where we need to justify excluding relations like 'urelement formation' as underlying determination relations. It's not clear how we can do that without sacrificing objectivity by appealing directly to our explanatory intuitions.

Finally, perhaps (contra Strevens) the underlying relation in this case is not asymmetric, but non-symmetric. That is, perhaps what is required are underlying necessitation relations. In that case, nothing is gained by going Kairetic, as there are necessitation relations at play in every derivation that satisfies the plain DN theory's conditions of adequacy, so the putative extra condition would be idle. As such, while the Kairetic theory looks initially promising, it does not allow for a better theory of metaphysical explanation except insofar as our explanations are unified with causal explanations. Thus, the remainder of this chapter is devoted to exploring the unification theory.

7.2 Kitcher's Unification Theory

Kitcher's unification theory of explanation (1981, 1989) is best understood as an extension of the DN model, which emphasises the role of unification in making derivations explanatory. This theory is more promising than the Kairetic option. Kitcher contends that explanation is ultimately about unification. Thus he constrains the class of DN derivations that count as explanatory by allowing only those that best unify that which needs to be explained. Those derivations that feature in the best systematisation of the DN derivations are the explanatory derivations. In brief, the best system is that which uses the fewest argument patterns to generate the biggest conclusion set, while keeping the patterns stringent, such that derivations which instantiate the same argument pattern are genuinely similar. The resulting theory has the power to rule out some of the DN derivations which did not strike us as explanatory, without appealing to psychologicistic filtering.

The plan for this section is as follows: §7.2.1 considers the motivations behind Kitcher's theory, including some historical examples of scientific explanations being lauded as *good explanations* through achieving greater unification. Next, section §7.2.2 will provide a clear explication of the formal mechanics behind the theory. Once this is all explicated we will be well placed to build a metaphysical variant of the theory and see if it, in turn, can circumvent the metaphysical analogues of the symmetry and irrelevance worries.

7.2.1 Motivation and Examples in the History of Science

Before we dive into the nuts and bolts of the theory, we should take a moment to consider the role of unification in explanation more generally, and the viability of appealing to unification in developing a putatively objective theory of explanation. Kitcher finds motivation in Hempel's claim that what scientific explanation aims at is "not [an] intuitive and highly subjective kind of understanding, but an objective kind of insight that is achieved by a systematic unification, by exhibiting that phenomena as manifestations of common, underlying structures..." (1966:83). He was also influenced by Feigl (1970), who sees explanatory unification as a way of comprehending all the facts one is aware of in terms of a minimal collection of concepts and assumptions.

While according to the plain DN theory, to qualify as an explanation merely requires satisfying the conditions of adequacy, Kitcher places complex constraints on which arguments are accepted as explanations. Perhaps his most important insight is that whether a certain argument or derivation is explanatory cannot be assessed in isolation. It is highly dependent upon what other propositions one assents to and what other explanatory argument forms one uses. In short, not any old argument with conclusion C is an explanation of C. Rather, the explanatory arguments are those that best *unify* one's beliefs. Thus one must consider one's entire corpus of beliefs holistically before one can know what explains what. This loose talk will be made much tighter in the following section.

Kitcher claims that unification yields understanding because "By using a few patterns of argument in the derivation of many beliefs we minimize the number of *types* of premises we must take as underived. That is, we reduce, in so far as possible, the number of types of facts we must accept as brute." Furthermore, "a theory unifies our beliefs when it provides one (or more generally, a few) pattern(s) of argument which can be used in the derivation of a large number of sentences which we accept." (1989:514).

As a scientifically motivated view of explanation, Kitcher is at pains to show that his theory is continuous with scientific practice. More precisely, he notes various examples in the history of science where theories were preferred on the basis that they explained a great many different states of affairs using the same principles and processes. The emphasis is on *repeated patterns*. A Kitcherian reading of the history of science shows explanations of particular events improving through subsuming the particular case under

more and more general patterns. Here we will note the way in which explanations in physics and biology over the centuries seem to follow a path as Kitcher prescribes. In the next section, once the formalisms are in hand we will see how the stories told here can be schematised.

Kitcher (1981) tells the story of Galileo and the fusilier. The fusilier has observed that his gun fires the furthest distance over a flat plain if it is angled at 45 degrees to the horizontal, and wants an explanation of his observation. Rather than providing an explanation specific to this particular case, Galileo instead explains why an *idealised* projectile, projected with fixed velocity on a perfectly smooth plain and only subject to gravitational acceleration attains maximum range at 45 degrees. He thus omits various situation-specific details: air resistance, the nature of the projectile, the earth's curvature, the uneven ground: such things are of negligible effect and thus justifiably ignored. *This* explanation, Kitcher tells us, involves a 'kinematical' argument that demonstrates why, for a large class of cases, a 45-degree shot goes the furthest (via some principles or laws). The point of the story is to spark the intuition that a good explanation will show how the explanandum, being of a certain *type*, can be explained through a pattern which covers many similar cases.

While Galileo subsumes the fusilier's explanatory request within a more general explanatory pattern that covers, for all projectiles, the relationship between their angle of ascent and their distance covered, this can be seen as merely a partial unification. An even more generalised theory of the physics of motion was proposed by Newton. Indeed, Kitcher views Newtonian physics as a paragon of explanation by unification. The success of the *Principia* (1687, translated and republished in 1962) in showing how to derive the motions of bodies from knowledge of the forces acting upon them suggested the possibility of dealing with moving systems in a unified way. Kitcher argues that Newton's work was regarded as such an impressive explanation of the movements of bodies precisely in virtue of its ability to deal with such systems in a unified way.

This inspired scientists to more generally seek understanding of a minimal collection of basic laws, which, when applied to the basic constituents of the world, could derive all natural phenomena. Chemical reactions, for instance, could perhaps be understood in terms of rearrangement of miniscule elements of the system under the action of certain forces of cohesion and repulsion. Thus Newtonians even hypothesised

inter-atomic forces despite lacking direct evidence for them. Boscovich (1966) took this to the extreme, claiming that all of natural philosophy can be reduced to a single law of forces in nature!

The point of this is that the great explanatory power of Newtonian mechanics stems from its ability to derive, from a small stock of mechanical principles, descriptions of diverse phenomena. Instead of believing in different principles in each observed case, if we can apply the very same principles and get the correct prediction/description for diverse phenomena, we have found the kind of unifying pattern Kitcher advocates as key to explanation.

Another example, this time taken from biology, shows how a theory which unified highly diverse phenomena was judged to be highly explanatory. The example is, of course, Darwinian evolutionary theory. As with Galileo's explanation to the fusilier, Darwin (1859, republished 2008) seeks not merely to provide us with a one-off explanation of why a particular organism exhibits certain characteristics. Rather, he seeks to provide a general recipe, or pattern, which shows how structurally similar processes can give rise to such diversity in nature. In this way, explanations of very different adaptations can be provided within the same framework as a result of interactions between genetic variation, selective pressures and heredity. Indeed, Darwin used schematic representations of the kind Kitcher advocates, and argued that such a schema can in principle be instantiated by a derivation of a description of any species. The schema would include the principle of natural selection, perhaps some laws of variation and inheritance, and also space for the ancestral forms of the creature and the relevant environmental considerations.

The intuitive gloss suggested by these examples is that Kitcher evaluates whether a derivation is explanatory as a function of whether that derivation is an instance of an explanatory pattern. This pattern gets to be explanatory in virtue of the fact that, if other variables were plugged in, many other phenomena could be explained. This is precisely the sense in which explanations are said to *unify*. The idea is intuitively compelling, but we cannot leave this explication at the level of a gloss. Kitcher precisely formulates what it takes for two explanations to exemplify a common pattern. He also provides a way in which we can compare different attempts to explain the world (recall that all evaluation of explanations, for Kitcher, must be at a holistic level) as a function of the number of,

and complexity of, the patterns used as well as the number of conclusions derived. This may sound ambitious and complex, but once we have Kitcher's formalisms and terminology in hand, the theory is quite manageable and useful.

The idea, I take it, is that it is no coincidence that these unifying explanations were taken to be particularly good explanations. There is something in the nature of explanation that makes unification necessary for explanation. If this is a conceptual truth, and no anthropocentrism sneaks in when evaluating different attempts at unification, then this theory remains an objective theory of explanation. However, one might be sceptical of the idea that unification is conceptually tied to explanation: maybe creatures like us just happen to prefer unifying explanations. I won't defend either view here. I'll simply note that only those in the former camp can take Kitcher's theory as a promising objective theory of explanation. Those in the latter camp can merely think of the view as one way of spelling out how to go about psychologistic filtering. We now turn to the formal expression of the mechanics of the theory.

7.2.2 The Nuts and Bolts

As Kitcher makes use of a lot of formal machinery in order to make precise the notions of argument patterns and stringency, we cannot get by merely with a rough gloss of the view before we evaluate the prospects of a metaphysical variant. We need to get into the nuts and bolts. I would not delve into this if it was not relevant to our purposes here, but it is through the increased complexity of Kitcher's schemas (in comparison to the plain DN theory) that he purports to be able to overcome some objections. In §7.4.1 I will use this machinery to show, for instance, why a derivation from the existence of Socrates to his singleton is explanatory whilst the inverse derivation, though logically valid, fails to be explanatory. Thus, I shall focus upon building a clear picture of his view in the diachronic/scientific case in order to establish what resources we have to work with in building a metaphysical variant.

Recall that according to the DN view, explanations are arguments. The unification theory continues with this tradition, but refers to these explanatory arguments as derivations. A *derivation* in this context is an ordered pair: a set of statements to serve as premises, and a conclusion statement. We take the set of accepted statements/propositions to be consistent and deductively closed, and name the set K .

Think of K as a systematisation of (a consistent version of) the set of statements endorsed by the scientific community. There is a set of derivations which *best unifies* K . This is known as the *explanatory store over K* , or $E(K)$. To be an explanatory derivation is simply to be a member of $E(K)$. The subset of sentences in $E(K)$ that are conclusions of the derivations therein is known as $C(E(K))$.

Given a certain K , there will be many candidate $E(K)$ s just as there are many ways to systematise the same corpus of beliefs. Crucial to the unification theory is that not all of these systematisations are equal. There is a privileged set of derivations which *best unifies* K ,¹¹¹ and it is only these which qualify as explanations. In order to develop criteria that allow us to compare two attempts to unify K (that is, two $E(K)$ s) we require the notion of a *general argument pattern*, such that it is clear when two explanations share a pattern, and clear how many different patterns are being used to systematise a corpus of beliefs. But we can't jump straight in and start talking about general argument patterns. We must build up to it via some more basic notions.

General argument patterns are a kind of *schematic argument*, built from *schematic sentences*. Let's start with the sentences and build up, using an example from Kitcher's 1989. Consider the sentence:

Organisms homozygous for the sickling allele develop sickle-cell anaemia.

We can turn this sentence into a schematic sentence by replacing some variables with dummy letters as follows:

Organisms homozygous for A develop P .

Unlike the original, this schematic sentence is quite widely applicable, and (appropriately filled) tells us about a variety of relationships between homozygous genotypes and particular phenotypes. Yet the variables are restricted in the kinds of things which can validly fill them. A is to be filled with an allele, and P with the corresponding phenotype. No other substitutions are permitted. Thus, schematic sentences must be paired with instructions on how to restrict substitutions for the variables. These restrictions on substitutions are called *filling instructions*, and they ensure that the unificationist's patterns are appropriately constrained—and more tightly so than those employed by logicians.

¹¹¹ Though Kitcher allows for the possibility that more than one $E(K)$ does best.

A *schematic argument*, then, is a set of schematic sentence/filling instruction pairs like the above. A schematic argument is accompanied by a *classification* which describes the inferential characteristics of the argument. That is, the classification tells us which sentences are premises, which conclusions, and how we can infer some from others.

Now we are equipped to characterise general argument patterns. A *general argument pattern* is a schematic argument (complete with filling instructions for each schematic sentence) and a classification.

With all of these notions spelled out, it will be useful to see them in action in an example. Here is Kitcher's (1981:517, variables changed to suit formatting) example of a general argument pattern covering Newtonian derivations of the motion of a single body:

- (1) The force on a is b.
- (2) The acceleration of a is y.
- (3) Force = mass*acceleration.
- (4) (Mass of a)*(y) = b
- (5) $x = z$

The filling instructions tell us that all occurrences of 'a' are to be replaced by an expression referring to the body under investigation; occurrences of 'b' are to be replaced by an algebraic expression referring to a function of the variable coordinates and of time; 'y' is to be replaced by an expression which gives the acceleration of the body as a function of its coordinates and their time-derivatives (thus, in the case of a one-dimensional motion along the x-axis of a Cartesian coordinate system, 'y' would be replaced by the expression ' d^2x/dt^2 '); 'x' is to be replaced by an expression referring to the variable coordinates of the body, and 'z' is to be replaced by an explicit function of time, (thus the sentences which instantiate (5) reveal the dependence of the variable coordinates on time, and so provide specifications of the positions of the body in question throughout the motion). The classification of the argument tells us that (1)-(3) have the status of premises, that (4) is obtained from them by substituting identicals, and that (5) follows from (4) using algebraic manipulation and the techniques of the calculus. (Kitcher, 1981:517)

(1) through (5) are schematic sentences. Together, they constitute a schematic argument. The spiel following the argument includes both filling instructions for the schematic sentences and a classification for the argument. The entirety of the quote comprises a general argument pattern.

Our next question concerns when a particular derivation (one that mentions individuals rather than variables) *instantiates* a general argument pattern. Kitcher tells us that a particular derivation instantiates a general argument pattern if:

- (i) The sequence has the same number of terms as the schematic argument of the general argument pattern.
- (ii) Each sentence in the sequence is obtained from the corresponding schematic sentence in accordance with the appropriate set of filling instructions.
- (iii) It is possible to construct a chain of reasoning which assigns to each sentence the status accorded to the corresponding schematic sentence by the classification. (1981:517)

So, if we took our Newtonian pattern above, filled in each variable according to the filling instructions, and followed the logical steps outlined by the classification, we would generate an instance of that pattern. Of course, many different instances can be generated—one for the movement of each body we might be interested in—and thereby the pattern has the power to unify explanations of the movements of various bodies. This makes the pattern a good contender to be part of the most unifying E(K), though whether it makes the cut depends on all the other things that need explaining and the other general argument patterns that can be generated.

Thinking in terms of general argument patterns provides a useful framework within which to think about *similarities* between derivations. Derivations are similar to one another in virtue of instantiating a common general argument pattern. Any derivation is maximally similar to itself, and likely has some similarity of form to any derivation (perhaps they are all deductively valid, for example). Yet such similarity admits of degrees. A pair of derivations is *more similar* than another pair if the general argument pattern they both instantiate is more *stringent*. Stringency is determined by the extent to which we restrict the filling instructions and classification. Maximal stringency leads to a single-case ‘pattern’ (a general argument pattern so tightly constrained as to only cover

one particular derivation), whilst minimal stringency allows the other degenerate case of the all-inclusive pattern (a general argument pattern so loosely constrained as to cover all derivations).

Kitcher notes that his argument schemas differ from those employed by logicians. This is particularly clear given that Kitcher's schemas must be accompanied by filling instructions and classification. In this way, there are tighter restrictions on which derivations can fill a Kitcher pattern, as opposed to a logician's pattern.

Let's stop and take stock. To summarise thus far, we have multiple competing sets of derivations (explanatory stores; $E(K)$ s), each of which is viable for the derivation of members of K . The members of each candidate $E(K)$ instantiate some general argument patterns. Thus we can say that each $E(K)$ is 'backed' by a store of general argument patterns. There will be variety in the number of patterns a certain $E(K)$ makes use of, and in the stringency of those patterns. The 'backing' of a certain candidate $E(K)$ is its *generating set* or *base*. Thus various $E(K)$ s, each aiming to be the explanatory systematisation of K , can be compared on the basis of their respective generating sets.

For example, one candidate $E(K)$ might make use of 20 general argument patterns for the base of its conclusion set, whilst another might only use 15 general argument patterns to generate the same conclusions. The former uses more patterns, but the patterns are, let us suppose, more stringent. The latter uses fewer patterns, but the patterns are less stringent. With these notions in hand we can express the crux of the theory. Kitcher provides two competing criteria, such that the best systematisation of K will be the $E(K)$ which strikes the best balance between the two.

The first criterion is that of *paucity of patterns*. I will sometimes refer to this simply as 'paucity'. This criterion tells us that *more unification is achieved through deriving as many conclusions as possible from the fewest number of argument patterns*. Thus, a small generating set with a large conclusion set is better, all else being equal. The goal is to use few, powerful patterns to generate as many conclusions as possible. A motivation for this criterion is that a systematisation that uses two different argument patterns in similar cases where a single pattern could have done the job has failed to unify.

The second criterion is that of *stringency of patterns*. I will sometimes refer to this simply as 'stringency'. This criterion tells us that if we go overboard with paucity of patterns such that the general argument patterns in the generating set are insufficiently

stringent, the unification will not be genuine. Recall that stringency is determined by the strictness of the filling instructions and the classification. Thus the second criterion serves to constrain the first. If all that mattered was minimising the number of argument patterns, this would grant favour to $E(K)$ s which use a single, unilluminating pattern for every derivation, and thus achieve merely ‘spurious’ unification.

Kitcher’s (1981) example of spurious unification is:

God wills that ∂ ∂ is to be filled with any member of K

∂

This general argument pattern manages to derive many conclusions, but ultimately fails to unify. By ‘explaining’ everything using this single-pattern generating set, phenomena are being forcibly unified regardless of their (lack of) similarity. Like phenomena are not being appropriately grouped according to a genuinely common pattern, as the filling instructions for ∂ are not stringent enough.

In addition to the above criteria, Kitcher provides the following corollaries:

- A) Let Σ, Σ' be sets of arguments acceptable relative to K (i.e. potential $E(K)$ s) which meet the following conditions:
- i. the basis of Σ' is as good as the basis of Σ in terms of the criteria of stringency of patterns, paucity of patterns, presence of core patterns, etc.
 - ii. $C(\Sigma)$ is a proper subset of $C(\Sigma')$ (recall that $C(\Sigma)$ is the set of the conclusions of arguments in Σ).

Then $\Sigma \neq E(K)$.

- B) Let Σ, Σ' be sets of arguments acceptable relative to K (i.e. potential $E(K)$ s) which meet the following conditions:

- i. $C(\Sigma) = C(\Sigma')$.
- ii. The basis of Σ' is a proper subset of the basis of Σ .

Then $\Sigma \neq E(K)$.

(1981; slightly altered for readability)

Corollary B formalises the criterion of paucity of patterns: if two $E(K)$ s have equal sized conclusion sets we should prefer the one with a smaller generating set. Corollary A formalises the further criterion that all else being equal, an $E(K)$ with a larger conclusions

set is preferable, as it explains more phenomena.

Paucity and stringency give rise to a ‘Goldilocks’ trade-off, which concerns precisely how to strike the right balance between them. We want to derive as much as possible from the fewest number of patterns, yet those patterns must be stringent enough that instances of the same pattern are genuinely similar.

Kitcher doesn’t provide a general rule which governs the balancing of these conditions: “I shall not explore the ways in which tradeoffs among these factors might be made. I am prepared to allow for the possibility that, with respect to some possible corpora K , there might be genuine indeterminacy in deciding how to weigh relative stringency, paucity of patterns and range of conclusions against one another, with consequent indeterminacy about $E(K)$ ” (1989:435). Yet his machinery allows us to see what matters when determining which derivations are explanatory.¹¹²

This section has quite quickly covered a lot of technical material. The take-home message is that the unification view differs from the DN view in that the status of a particular derivation as explanatory (or not) cannot be assessed in isolation. Rather, it must be evaluated as part of a system of derivations. These systems can be compared on the basis of the number of general argument patterns they instantiate, the stringency of these patterns, and the conclusions generated. The best system will strike the right balance for this Goldilocks trade-off, such that few patterns are used, yet the patterns are such that their instantiations are genuinely similar. The derivations that are part of this best system are the explanations.

7.3 Metaphysical Unification

Just as the unification theory of scientific explanation is an extension of the DN theory of scientific explanation, so too is the unification theory of metaphysical explanation an extension of the DN theory of metaphysical explanation. The substantial

¹¹² It is also worth noting that Kitcher expresses concern regarding how to compare the numbers of patterns utilised in a particular generating set, given that small, core patterns may make an appearance in many other more complex patterns. For example, recall that the Newtonian schema above allows for the derivation of the motion of a single body. The motion of a particular body can serve as *part of* the explanation of a great many phenomena. These derivations may not share a common pattern overall, yet there is a sense in which they are not entirely distinct, as they share a sub-pattern, or a *core pattern*. This sharing of sub-patterns contributes to the unifying power of a generating set, but makes precise calculation of relative unifying power all the more difficult!

difference will be that constraints of unification will exclude as explanations some derivations which the plain DN theory allowed. Thus, as long as we don't end up ruling out derivations of intuitively explanatory cases, the unification theory inherits many of the features outlined by our desiderata. In particular, the view straightforwardly covers cases, secures irreflexivity, and ensures that a metaphysical explanans will entail its explanandum.

So, how will a unificationist variant of the metaphysical DN theory look? With Kitcher's machinery in hand, we can revisit some of our metaphysical DN derivations from Chapter 6 and build some general argument patterns. Indeed, in some cases we have already identified generic arguments which almost fit the bill. Recall:

- (1) Relation R obtains between the collection of entities $P_1...P_n$.
- [L2] Necessarily, for any entities $P_1...P_n$, if the P_n exist and relation R obtains between the P_n , then W , the object composed by the P_n , exists.

Therefore,

- (2) W , the object composed by the P_n , exists.

This argument can be schematised into a general argument pattern as follows:

- (1) Relation R obtains between the collection of entities $P_1...P_n$.
- [L2U] Necessarily, for any entities $P_1...P_n$, if the P_n exist and relation R obtains between the P_n , then W exists.

Therefore,

- (2) W exists.

*The filling instructions might tell us that the P_n are non-overlapping concrete objects, R tells us how the P_n must be arranged in order for there to exist a composite whole and W is the object composed by the P_n . The classification might tell us that (2) follows from (1) and [L2U] by *modus ponens*.*

Let's call this the *composition pattern*. It includes a schematic argument, filling instructions and a classification, and thus counts as a general argument pattern. The composition pattern looks like the kind of pattern we will want to have in our generating set, as it is powerful (in generating a lot of conclusions) and stringent (in limiting its

application to the genuinely similar explanations of the existence of composite entities in terms of their parts). This shows that where the plain DN theory was unproblematic, so too the unification theory will be unproblematic. Thus we will move quickly to some problematic cases for the plain DN theory that the unificationist theory can help address.

Consider now a pair of symmetrical ‘truthmaking derivations’ making use of [L5**]:

(1) Socrates exists.

[L5**] Necessarily, for any entity E , E exists just in case the proposition $\langle E \text{ exists} \rangle$ is true.

Therefore,

(2) $\langle \text{Socrates exists} \rangle$ is true.

(1) $\langle \text{Socrates exists} \rangle$ is true.

[L5**] Necessarily, for any entity E , E exists just in case the proposition $\langle E \text{ exists} \rangle$ is true.

Therefore,

(2) Socrates exists.

These arguments can be schematised, in turn, as follows (and made more general: the following patterns do not just concern whether objects exist). The former is an instance of the following *truthmaking pattern*:

(1) F .

[TMU] Necessarily, F is the case just in case the proposition $\langle F \rangle$ is true.

Therefore,

(2) $\langle F \rangle$ is true.

*The filling instructions tell us that any fact can be substituted for F . The classification tells us that (2) follows from (1) and [TMU] by *modus ponens*.*

The latter is an instance of the following *reverse truthmaking pattern*:

(1) $\langle F \rangle$ is true.

[TMU] Necessarily, F is the case just in case the proposition $\langle F \rangle$ is true.

Therefore,

(2) F .

The filling instructions tell us that any fact can be substituted for F . The classification tells us that (2) follows from (1) and [TMU] by modus ponens.

We won't try to find any asymmetries just yet. We will simply get the problematic cases spelled out first. Hence, consider again the symmetrical derivations between singletons and their urelements. This move is the same as the above, so we can skip straight to the general argument patterns. Here is the *singleton set formation pattern*:

(1) E exists.

[L4*] Necessarily, for any entity E , E exists just in case the singleton set $\{E\}$ exists.

Therefore,

(2) $\{E\}$ exists.

The filling instructions tell us that any entity can be substituted for E . The classification tells us that (2) follows from (1) and [L4] by modus ponens.*

Here is the *urelement formation pattern*:

(1) $\{E\}$ exists.

[L4*] Necessarily, for any entity E , E exists just in case the singleton set $\{E\}$ exists.

Therefore,

(2) E exists.

The filling instructions tell us that any entity can be substituted for E . The classification tells us that (2) follows from (1) and [L4] by modus ponens.*

The symmetry problem, as it arises in the context of the unification theory, concerns whether the best systematisation of our metaphysical explanations will include pairs of general argument patterns like the truthmaking pattern and the reverse truthmaking pattern, or the singleton set formation pattern and the urelement formation pattern. This is a question we will tackle in the next section, along with the symmetry-like cases from §6.4.3. For now, let's look at some patterns that give rise to apparently irrelevant derivations. Call the following the *spurious number pattern*.

(1) E exists.

[L7U] Necessarily, if E exists then N exists.

Therefore,

(2) N exists.

*The filling instructions tell us that any entity can be substituted for E , and any number can be substituted for N . The classification tells us that (2) follows from (1) and [L7U] by *modus ponens*.*

Call the following the *spurious universal pattern*.

(1) E exists.

[L8U] Necessarily, if E exists then U exists.

Therefore,

(2) U exists.

*The filling instructions tell us that any entity can be substituted for E , and any universal can be substituted for U . The classification tells us that (2) follows from (1) and [L8U] by *modus ponens*.*

Substituting an entity like my left foot for E in either the spurious number pattern or the spurious universal pattern generates a derivation where the premises are intuitively irrelevant to the conclusion.¹¹³ That we might end up with one of these spurious patterns in our generating set is the irrelevance problem, as that problem manifests in the context of the unification theory.

¹¹³ Unless there is a 'left foot' universal!

7.4 Paucity and Stringency to the Rescue?

We have seen above how instances of metaphysical explanation can be captured by general argument patterns. As such, thinking of metaphysical explanation through Kitcher's unificationist lens is viable. What remains to be seen is whether the best systematisation of our metaphysical explanations will make use of patterns that allow for symmetrical explanations or irrelevant explanations. If the unwanted patterns from §7.3 are part of the generating set we end up with, we have made no progress in solving these problems.

Recall the machinery at our disposal to help rule out these unwanted patterns: the best systematisation will use the fewest patterns to generate the largest conclusion set—but each pattern must be *stringent* such that its instances are genuinely similar.

7.4.1 Asymmetry

Here I propose that the Kitcherian apparatus might help defuse the symmetry problems. Let's tackle them in turn, starting with the reverse truthmaking pattern.

The first question to ask is whether making use of this pattern increases the size of our conclusion set. If adding the pattern to our generating set of argument patterns grants us no conclusions we could not already derive, we can call upon paucity considerations to rule it out.

The conclusions delivered by the reverse truthmaking include any obtaining fact. This should immediately make us suspicious about its stringency, and we will return to this thought. But first, do we need to use *this* pattern to derive these conclusions? For some, at least, it would seem not. For example, the pattern can derive Socrates' existence from the truth of <Socrates exists>. But consider that Socrates' existence can be derived via the composition pattern. More generally, all of the conclusions that the reverse truthmaking pattern delivers regarding the existence of composite objects are already accounted for by the composition pattern.¹¹⁴

Of course, the phrase 'already accounted for' is misleading here, just as it was when I used it back in §1.2. I don't want to imply that there are some patterns that we have, *at*

¹¹⁴ Note that it is only the existence of *composite concreta* that can be accounted for by the composition pattern. We will consider simples, abstracta, and other putative fundamentalia, below.

some earlier time, accepted into our generating set, and now we are thinking of which other patterns to add. Rather, the thought is that the composition pattern is, as noted in §7.3, powerful and stringent, and accounts for many apparent metaphysical explanations. As such, it seems like a very strong contender to be in the generating set, and that the reverse truthmaking pattern overlaps conclusions with the composition pattern tells against its inclusion.

However, as noted, the reverse truthmaking pattern doesn't just derive the existence of composite objects. It derives the existence of any obtaining fact. For instance, we can derive that the ball is red from the truth of <the ball is red>. The reverse truthmaking pattern is starting to look very powerful, indeed. It can even produce as an output anything it can take as an input (consider that <<P is true> is true> allows the derivation of <P is true>). Doesn't the ability to derive *anything* sound awfully like Kitcher's canonical example of spurious unification? What is the relevant difference between derivations from 'God believes that ∂ ' to ∂ and derivations from < ∂ is true> to ∂ ? As such, the reverse pattern appears to fail to unify genuinely similar phenomena.

So, perhaps we can rule it out. But perhaps this is too hasty. The reverse truthmaking pattern has the power to derive the existence of non-composite (i.e. simple) objects, which the composition pattern clearly cannot. Indeed, perhaps the reverse truthmaking pattern is the *only* pattern which can derive these conclusions,¹¹⁵ and does that not render it *indispensable*?

This is a very interesting question, and to answer it we must go back to our initial characterisation of Kitcher's view. In the metaphysical case, we can think of K as (a consistent version of) the set of statements endorsed by the metaphysical community. Once again, there is a set of derivations which best unifies K , known as $E(K)$. We are currently considering which patterns will be instantiated by the derivations in $E(K)$. But we should also recall that Kitcher tells us that the subset of sentences in $E(K)$ that are conclusions of the derivations therein is known as $C(E(K))$. That $C(E(K))$ is a *subset* of $E(K)$ is significant. Importantly, not every member of K need be a member of $C(E(K))$.

How does this link up with our question above? Well, statements about the existence of putatively fundamental existents like simples and numbers are just the kinds of statements we might expect to be part of K , but not part of $C(E(K))$. Plausibly, we

¹¹⁵ It is not, in fact, the only putative pattern with the power to derive these conclusions, as we shall see below. But we shall continue as though it were.

shouldn't be trying to create a systematisation of K that includes such claims in its conclusion set. Some things require no explanation, and there is no need to warp our explanatory store in such a way as to provide forced explanations of their existence. As such, that the reverse truthmaking pattern does provide putative explanations for such statements does not tell in its favour, or make it indispensable. So the stringency constraint rules out the reverse truthmaking pattern.

What, then, of the truthmaking pattern? The first thing to note is that *this* pattern *does* appear to increase the size of our conclusion set. There are no other patterns competing for the job of deriving conclusions about the alethic properties of propositions. Thus this pattern does not fall afoul of paucity. Yet, is it not similarly spurious to the reverse truthmaking pattern? It makes use of the very same metaphysical law, after all. In response to this, we can point out that, by running 'the other direction', there is a substantial difference in the relative stringency of the two patterns. The reverse truthmaking pattern takes something uniform (that certain propositions are true), and spits out conclusions about the existence of any fact that obtains. In contrast, the truthmaking pattern takes in a diverse range of statements (precisely the same diverse range that the other pattern spits out) and derives conclusions only relating to the truth of various propositions. As such, the truthmaking pattern has genuine similarity amongst its conclusion set in a way that the reverse truthmaking pattern does not. *This* similarity of conclusions is what motivates adding constraints of unification to explanatory theories.¹¹⁶

That the unificationist machinery can tell in favour of the truthmaking pattern over the reverse truthmaking pattern is a strong start. We now need to consider whether it can perform the same role with regards to the singleton set formation pattern and the urelement formation pattern. It can, and this can be shown relatively quickly, as similar considerations apply.

Of primary concern is that the urelement formation pattern fails the stringency requirement. We can use that pattern to derive that any entity exists on the basis of the existence of the singleton set that contains the entity. Like the reverse truthmaking pattern, this pattern generates a very large conclusion set, but the conclusions are not

¹¹⁶ Consider, for example, Kitcher's Newtonian pattern. It derives, for any body with any force acting upon it, the equation telling us where that body will be at what time. The similarity is in the type of conclusion the pattern derives, not in what kind of body is initially considered.

appropriately unified: the unification is spurious. The urelement formation pattern does allow us to derive conclusions that could not be derived without it (as there are putatively unexplained entities which are members of sets), but once again, as per our discussion of reverse truthmaking, being able to derive *these* conclusions is not a valuable contribution. Thus, this pattern is not indispensable and, while powerful, fails to be stringent and thus should not be part of our base.

The singleton set formation pattern, however, derives non-fundamental conclusions that cannot be derived elsewhere: namely, the existence of sets. Furthermore, the conclusions of this pattern are appropriately similar: they're all statements about the existence of sets. In the big picture, this pattern will probably be jettisoned in favour of a more general set formation pattern (one that allows the explanation of the existence of sets with more than one member), as using separate patterns to derive the existence of singleton and non-singleton sets will fail the requirement of paucity of patterns. But that's fine. The point is that the unificationist machinery tells in favour of patterns that derive sets from their members but against patterns that derive the existence of members from the existence of sets.

At this point I should note that the viability of my arguments, here, turns on whether it is legitimate to take some things as requiring no explanation. In particular, it turns on whether, by taking some specific things (such as simples) to require no metaphysical explanation, anthropocentric considerations are sneaking into our supposedly objective theory. I don't intend to answer this question, here. Rather, I am flagging it as a point with which future proponents and opponents of the unification theory of metaphysical explanation should engage.

We now move on from straightforward symmetry problems to talk about the symmetry-like problems discussed in §6.4.3. Recall that the problematic derivations were cases like the following:

(1) The champagne glass is fragile.

[L6*] Necessarily, for any entity E , E exists and possesses a categorical property of kind K , just in case E possesses the dispositional property of being fragile.

Therefore,

(2) The champagne glass has a categorical property of kind *K*.

This derivation can be schematised to form a general argument pattern:

(1) *E* is fragile.

[L6*] Necessarily, for any entity *E*, *E* exists and possesses a categorical property of kind *K*, just in case *E* possesses the dispositional property of being fragile.

Therefore,

(2) *E* has a categorical property of kind *K*.

The filling instructions tell us that any entity can be substituted for E, and that the K-properties are those categorical properties that realise the fragile disposition. The classification tells us that (2) follows from (1) and [L6] by modus ponens.*

Will this *disposition-categorical* pattern be part of our generating set? There are good reasons to think not. For we should only make use of this pattern if it can generate new conclusions. Given that the conclusions of this pattern tell us that some entity has a certain determinable property (like having a property of kind *K*), it is likely that the conclusion set will be a subset of the *determinate-determinable* pattern, which derives that objects have determinable properties on the basis of their determinate properties. It is a mere subset of the conclusion set of the *determinate-determinable* pattern because the *disposition-categorical* pattern will only derive determinable properties that are the basis of some disposition. Thus paucity of patterns tells us to jettison the *disposition-categorical* pattern in favour of the more powerful *determinate-determinable* pattern.

All in all, while I don't claim to have decisively solved the symmetry problem on behalf of the unificationist, I do think that I've shown that, unlike the plain DN theory, the view has the resources available to find such a solution.

7.4.2 Irrelevance

We now turn to the problem of irrelevance, and to consider whether constraints of unification can rule out unwanted irrelevant derivations. Recall (§6.4.4) that one way to rule out the seemingly irrelevant derivations is to call upon a distinction between accidental and non-accidental metaphysically necessary generalisations. However, I have

expressed pessimism regarding whether this can be done in a non-anthropocentric way. Here, we explore whether the unification theory can rule out intuitively irrelevant derivations without appealing to this distinction. Interestingly, Kitcher himself does not call upon this distinction in the scientific case. His view is that we discover what the laws are by looking at the generalisations that make an appearance in our best systematisation. We might think of this as an ‘explanations first’, rather than a ‘laws first’ approach.

Thus, let’s assume that all metaphysically necessary generalisations are candidate metaphysical laws. Then we will need to decide what to say about the following *necessary existent pattern*:

(1) E exists.

[L9] Necessarily, if E exists then N exists.

Therefore,

(2) N exists.

The *filling instructions* tell us that any entity can be substituted for E , and any necessarily existing entity can be substituted for N . The *classification* tells us that (2) follows from (1) and [L9] by *modus ponens*.

The necessary existent pattern is a general argument pattern that subsumes the spurious number pattern and the spurious universal pattern discussed in §7.3. It allows us to derive, from the existence of anything, the existence of any necessarily existing entity. The first thing to note is that some of the conclusions this pattern allows fall outside the category of statements we should expect to see in $C(E(K))$. That’s because many necessary existents are plausibly unexplained (numbers, and universals, for instance). While there will be those who think that the existence of universals depends upon the existence of their instances, these will surely not be the same people who think that universals are necessary existents, as for any universal there will be worlds where it is not instantiated. What about sets of numbers? These can already be derived via the set formation pattern, which is required to derive the existence of sets that don’t exist necessarily, and is more stringent than the necessary existent pattern.

So, what conclusions *can* the necessary existent pattern derive that might encourage us to include it in our generating set? Surely not all necessary states of affairs are unexplained (setting aside necessarily existing sets, the existence of which can be derived

from the existence of their members via the singleton set formation pattern). The truth of mathematical propositions, for instance, such as $\langle 2+2=4 \rangle$, are plausibly in need of explanation in terms of the existence of their constituent numbers.¹¹⁷ Tautologies, such as instances of $\langle \text{if } P \text{ then } P \rangle$ and $\langle P \text{ or not } P \rangle$ are necessarily true, and perhaps their truth is in need of explanation.

I think the right response for the unificationist here is to point out that there is a more stringent pattern than the necessary existent pattern which can derive these particular non-fundamental necessary conclusions. Consider the *analytic pattern*:

(1) A is an analytic proposition.¹¹⁸

[LA] Necessarily, if A is an analytic proposition, A is true.

Therefore,

(2) A is true.

The filling instructions tell us that A is an analytic proposition (including tautologies, mathematical propositions,¹¹⁹ etc). The classification tells us that (2) follows from (1) and [LA] via modus ponens.

The analytic pattern is more stringent than the necessary existent pattern. It has fewer instances (only one per analytic truth)¹²⁰, and can derive many of the conclusions of the latter pattern. Furthermore, the non-analytic conclusions that we cannot derive are precisely those we need not, namely those regarding the existence of necessary fundamental entities. Thus, a generating set that includes the analytic pattern is preferable to one that includes the necessary existent pattern.

All in all, I do not claim to have comprehensively resolved all of these problems on behalf of the unification theory of metaphysical explanation. However, it does seem that thinking in terms of a minimal set of stringent patterns can go a long way towards vindicating our intuitions about what metaphysically explains what. In particular, we can make headway on those symmetrical and irrelevant cases that we seemed stuck with on the plain DN theory. I hope to have shown that, for those who heavily weight the

¹¹⁷ Or perhaps such facts should be explained in terms of the entirety of the number-theoretic system, as mathematical structuralists think.

¹¹⁸ Where analytic propositions are those that are true in virtue of meaning alone.

¹¹⁹ Though not everyone will think that mathematical propositions are analytic.

¹²⁰ There are concerns about infinities, here. A pattern that has an instance for every analytic fact has an infinite number of instances, so it is hard to see how it could have fewer instances than some other pattern.

This can be made precise using a system of constructed infinities such as Conway's (1976) surreal numbers. 208

objectivity desideratum, the unificationist extension of the DN theory can go the furthest towards also securing the epistemic desiderata. Thus, it looks to be by far the most promising avenue through which to explore the possibility of an objective theory of what metaphysically explains what that accords with our intuitions about the epistemic features of such explanations.

7.5 Conclusion

There is a lot of work that remains to be done in developing a comprehensive unification theory of metaphysical explanation. However, these initial results are promising, and I contend that, for those who insist that a viable theory of metaphysical explanation must satisfy the ontic desideratum of objectivity, this is the place where effort should be expended. This marks the end of the investigation of this thesis, for Chapter 8 is devoted to weaving together all that I have argued so far, to show how the thesis has defended the master argument.

Chapter 8: Conclusion

I have argued that we should not posit grounding relations, for they are not indispensable to the best explanation of our observations. Recall the master argument of the thesis:

1. One ought (epistemically) to be ontologically committed to all and only those entities that are indispensable to the best explanation of our observations.
2. Grounding relations are not indispensable to the best explanation of our observations.
3. Therefore, we should not be ontologically committed to grounding relations.

In **Chapter 1** I pointed out the intuitive plausibility of premise 1 (the explanatory criterion for ontological commitment). Nonetheless, I do not take myself to have conclusively argued for this premise. Instead, I have taken it as an intuitive starting point, and assumed its truth throughout the thesis. Hence, the bulk of the thesis has been devoted to an extended comparison of the explanations offered by the Sophisticated Grounding Theory and the Grounding-free Theory. I have followed Colyvan (1999) in supposing that an entity, E , is dispensable to a theory, T , if there is available a distinct theory, T^* , that is empirically equivalent to T yet makes no mention of E , and T^* exemplifies theoretical virtues to a greater extent than T . Therefore, as the Sophisticated Grounding Theory and the Grounding-free Theory are empirically equivalent, Chapters 2 through 7 have defended premise 2 by showing that the Grounding-free Theory is more theoretically virtuous than its grounding-based competition.

Chapter 2 identified two classes of observations, the explanation of which might be thought to indispensably require grounding. These are our intuitions about which modal correlations are accidental and non-accidental, and our priority intuitions. The former were put to the side, narrowing our focus to potential explanations of the priority intuitions. Here I also distinguished between explanations that *vindicate* the priority intuitions and those that fail to do so, predicting that some readers will insist that the priority intuitions be explained in such a way as to vindicate as true many claims of the form ‘ x because y ’ uttered by contemporary metaphysicians. This vindication, I suggested, will take the form of a theory of metaphysical explanation, analogous to a theory of

scientific explanation. Such a theory will provide truth conditions for claims of the form ‘x *because* y’, and can be evaluated against the desiderata for a theory of explanation.

Chapter 3 introduced grounding, a primitive dependence relation I believe not to exist. After isolating the ‘relation’ (as opposed to the ‘sentential operator’) view of grounding as our target, I conceded that the Sophisticated Grounding Theory—which borrows some machinery from the psychological story I tell in Chapter 4—can explain our priority intuitions. I carefully reviewed the literature concerning the characteristics of grounding, and what is taken to ground what, showing how debates about grounding proceed via consultation with our priority intuitions. That is, consensus regarding the grounding structure of the world is achieved through introspection of what appears to be ontologically prior to what. I argued that this front-loading of the priority intuitions into the theory of grounding itself yields a theory of metaphysical explanation that, despite claims of unfiltered objectivity, is in fact *illicitly pre-filtered*. This was emphasised via an analogy with causation*, a relation that putatively obtains alongside all and only those instances of causation which we find explanatory. Just as a theory of scientific explanation built around causation* is implausible, so too is a theory of metaphysical explanation built around grounding. On the other hand, the filtered grounding-based theory of metaphysical explanation compares unfavourably to the filtered modal relations theory advanced in Chapter 5, for it secures the same desiderata but requires the additional posit of grounding, and faces some awkward epistemic questions.

In **Chapter 4**, I advanced my preferred explanation of the priority intuitions. I suggested that a pair of cognitive mechanisms—what I call the correlation detection mechanism and the causal detection mechanism—which evolved to track causal dependencies, are responsible for our judgements of priority. I noted that failing to detect causal relations can be very costly, while artificially imbuing the world with causal relations that do not, in fact, obtain is relatively low-cost. Thus the causal detection mechanism has a low threshold at which it triggers that there is a non-symmetric instance of causation: it *overgeneralises*. Moreover, it is plausible that the same mechanisms we use to detect diachronic dependencies would be co-opted to detect non-diachronic dependencies. As the cues to which the causal detection mechanism is sensitive are often not present in the non-diachronic context, its tendency to overgeneralise often delivers the verdict that there is a non-symmetric instance of a modal relation when the relevant relation obtains symmetrically. In addition, as these mechanisms evolved to detect an

asymmetric relation (causation) which is associated with explanation, we have the same intuitions in the non-diachronic context: that the apparent non-symmetry is due to an underlying asymmetric relation. In this way, the asymmetric and explanatory nature of our priority intuitions can be explained without an asymmetric ontological posit like grounding.

Moreover, I argued that the Grounding-free Theory is more theoretically virtuous than its grounding-based competitor. The first advantage is one of parsimony. As the Sophisticated Grounding Theory is committed to the existence of the set of psychological mechanisms I describe in Chapter 4 to explain our intuitions of diachronic priority—and, plausibly, also to explain the priority intuitions, by providing a mechanistic story about how our judgements track grounding relations—the Grounding-free Theory can generate the same predictions as the Sophisticated Grounding Theory while positing fewer entities. Secondly, grounding comes with ideological baggage, being governed, as it putatively is, by various principles. The Grounding-free Theory eschews the need for these principles and thus has less ideological complexity: it is more elegant. Thirdly, assuming that the Sophisticated Grounding Theory appeals to the psychological mechanisms I elucidate, the theories are equally unified, each accounting for our intuitions of non-diachronic priority by appealing to the functioning of precisely the same mechanisms that explain our intuitions of diachronic priority. Moreover, the grounding-free theories of metaphysical explanation promise unification with theories of scientific explanation, as I showed in Chapters 5, 6 and 7. As the Grounding-free Theory is empirically equivalent to, and more theoretically virtuous than, its rival, I conclude that grounding is not indispensable to the best explanation of our observations, and indeed should be dispensed with.

As noted, however, I foresee that without pairing my preferred explanation with a theory of metaphysical explanation which can vindicate the priority intuitions, many will find my explanation to be inadequate. For those who are friendly to a non-objective theory of explanation, in **Chapter 5** I developed the filtered modal relations theory. This theory built upon the work of Chapter 4 by showing how our own psychological states (along with the obtaining of modal relations) can serve as truthmakers for metaphysical explanations. The idea, in brief, is that the truth-conditions for claims of the form ‘x *because* y’ can be captured by a relativist semantics. My preferred account is Dispositional

Community Relativism, which has it that ‘x *because* y’ is true relative to a centred world $\langle w, t, i \rangle$ iff:

- (1) Most members of the community in which *i* is embedded are disposed to have the mental state according to which x *because* y, and
- (2) Either (i) y necessitates x or (ii) x supervenes on y.

Of course, other relativistic proposals are available for those who think that Dispositional Community Relativism fails to capture the right truth conditions. The point is that, for those friendly to a non-objective theory of metaphysical explanation, the application of psychologistic filtering to the modal relations can form the basis of such a theory. Moreover, this theory does equally well with the desiderata as the filtered grounding-based theory, securing every desideratum but objectivity.

The final chapters explored the prospects of developing objective yet grounding-free theories of metaphysical explanation. **Chapter 6** built a ‘plain’ DN theory, analogous to Hempel and Oppenheim’s (1948) DN theory of scientific explanation. The success I had with building this theory bodes well for the potential of metaphysical DN theories. While a psychologically filtered variant of the view does as well, according to the desiderata, as the filtered modal relations theory (securing all but objectivity), the objective variant succeeds in covering actual cases of metaphysical explanation and rules out reflexive explanations, but struggles with the epistemic desiderata of relevance, understanding and asymmetry. As such, its epistemic characteristics leave something to be desired.

Thus, **Chapter 7** built a metaphysical incarnation of Kitcher’s (1981, 1989) unification theory of scientific explanation. I showed how the Kitcherian apparatus has the potential to fulfil the epistemic desiderata where the plain DN theory could not. In particular, I showed how, if we take as given that some things are not in need of metaphysical explanation, the Kitcherian constraints of paucity and stringency can help rule out symmetrical and irrelevant explanations, by showing that these unwanted explanations instantiate general argument patterns that are not part of the most unifying generating set.

I do not claim to have comprehensively solved these problems. This is unsurprising, given that the jury is still out regarding whether the unification theory can secure these desiderata in the scientific context. Neither do I claim that the unification

theory of metaphysical explanation developed here is comprehensive. Nevertheless, I propose that, for those seeking an objective theory of metaphysical explanation, yet convinced by my anti-grounding arguments, the unification theory is a promising avenue via which such theories might be developed. Another avenue towards this goal might involve non-anthropocentrically distinguishing between accidental and non-accidental modal correlations, and plugging this distinction into the plain DN theory in order to rule out intuitively irrelevant derivations. I am less sanguine about this possibility.

However, I am optimistic about the future project of explaining our intuitions about accidental and non-accidental modal correlations in the context of the Grounding-free Theory. Indeed, as I noted in Chapter 1, the conclusions of this thesis are conditional on the success of this project, for otherwise I cannot claim that the Grounding-free Theory can explain the same observations as the Sophisticated Grounding Theory. If the theories are not empirically equivalent in this way, then there is no need to appeal to the theoretical virtues as a tie-breaker, and my arguments are entirely undermined. This is a project with which I am already engaged.

Finally, recall the big picture questions with which I introduced my project: questions about how to link up scientific investigation of macroscopic and microscopic phenomena, and about how to integrate the posits of the various scientific sub-domains into a unified ontological picture. I hope to have shown that these questions are not best answered by positing grounding relations between these entities, (or between the facts partially constituted by these entities) but by carefully investigating the modal relations that obtain between them. There is much work still to be done. But that is work for another day. I only hope to have said enough to motivate my reader to care about this future work, and to think that the work may bear fruit.

BIBLIOGRAPHY

- Aristotle (1984). 'Categories.' In J. Barnes (ed.), *The Complete Works of Aristotle, vol. 1*. Princeton: Princeton University Press. 3-24.
- Arkes, H. R. and Harkness, A. R. (1983). 'Estimates of contingency between two dichotomous variables.' *Journal of Experimental Psychology: General*, 112:117-135.
- Armstrong, D. M. (1997). *A World of States of Affairs*. Cambridge: Cambridge University Press.
- Armstrong, D. M. (2004). *Truth and Truthmakers*. Cambridge: Cambridge University Press.
- Atran, S. (2002). *In gods we trust: The evolutionary landscape of religion*. Oxford: Oxford University Press.
- Audi, P. (2012a). 'A Clarification and Defense of the Notion of Grounding.' In F. Correia and B. Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press.
- Audi, P. (2012b) 'Grounding: Toward a Theory of the In-Virtue-Of Relation.' *The Journal of Philosophy*, 109:685–711.
- Barnes, Elizabeth. (forthcoming). 'Symmetric Dependence.' In R. Bliss and G. Priest (eds.), *Reality and Its Structure*. Oxford: Oxford University Press.
- Barnes, Eric. (1992). 'Explanatory Unification and the Problem of Asymmetry.' *Philosophy of Science*, 59: 558–71.
- Baron, S. (2015) 'Et Tu, Brute?' In C. Daly (ed.), *The Palgrave Handbook of Philosophical Methods*. Palgrave MacMillan 258-281.
- Baron, S. (2016). 'Metaphysics as Fairness.' *Synthese*, 193 (7):2237-2259.
- Baron, S. (ms). 'Grounding Makes a Difference.'
- Barrett, J. L. (2004). *Why would anyone believe in God?* Plymouth: Alta Mira Press.
- Beebe, H. (2004). 'Causing and Nothingness.' In J. Collins, N. Hall and L. Paul (eds.), *Causation and Counterfactuals*. Cambridge, Mass: MIT Press.

- Bennett, K. (2011b). 'Construction Area (No Hard Hat Required).' *Philosophical Studies* 154: 79-104.
- Bird, A. (2005). 'Explanation and Metaphysics.' *Synthese*, 143: 89–107.
- Bohn, E. D. (2009). *Composition as Identity: A Study in Ontology and Philosophical Logic*. Doctoral Dissertation, University of Massachusetts Amherst.
- Bohn, E. D. (2014). 'From Hume's Dictum via Submergence to Composition as Identity or Mereological Nihilism.' *Pacific Philosophical Quarterly*, 95(3):336-355.
- Boscovich, R. J. (1966). *A Theory of Natural Philosophy* (trans. J. M. Child). Cambridge: M.I.T. Press.
- Boyer, P. (2001). *Religion explained*. London: Vintage.
- Braddon-Mitchell, D. and Jackson, F. (2007). *Philosophy of Mind and Cognition*. Malden, MA: Blackwell Publishing.
- Bromberger, S. (1965). 'An Approach to Explanation.' In R. J. Butler (ed.), *Analytical Philosophy*, Second Series. Oxford: Blackwell.
- Cameron, R. (2008). 'Turtles All the Way Down: Regress, Priority, and Fundamentality.' *The Philosophical Quarterly*, 58(230):1-14.
- Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: Clarendon Press.
- Chalmers, D. (2008). 'Varieties of Emergence.' In P. Clayton and P. Davies (eds.), *The Re-Emergence of Emergence: The Emergentist Hypothesis From Science to Religion*. Oxford: Oxford University Press.
- Chapman, L. J. and Chapman, J. P. (1967). 'Genesis of popular but erroneous psychodiagnostic observations.' *Journal of Abnormal Psychology*, 72:193-204.
- Chudnoff, E. (ms). 'Grounding and Entailment.'
- Colyvan, M. (1999). 'Contrastive empiricism and indispensability.' *Erkenntnis*, 51(2–3): 323–332.
- Colyvan, M. (2000). 'Conceptual Contingency and Abstract Existence.' *Philosophical Quarterly*, 50(198):87-91.

Colyvan, M. (2001). *The Indispensability of Mathematics*. New York: Oxford University Press.

Conway, J. (1976). *On Numbers and Games*. London: Academic Press.

Correia, F. (2005). *Existential Dependence and Cognate Notions*. Munich: Philosophia Verlag.

Correia, F. (2011). 'From Grounding to Truth-Making: Some Thoughts.' In A. Reboul (ed.), *Philosophical Papers Dedicated to Kevin Mulligan*. Geneva: Université de Genève.

Craig, W. L. and J. Sinclair. (2009). 'The Kalam Cosmological Argument.' In W. L. Craig, and J. P. Moreland (eds.), *The Blackwell Companion to Natural Theology*. Oxford: Blackwell.

Daly, C. (2012). 'Skepticism about Grounding.' In F. Correia and B. Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press.

Darwin, C. (2008). *On the Origin of Species*. Oxford: Oxford University Press.

Dasgupta, S. (2014). 'On the Plurality of Grounds.' *Philosophers' Imprint*, 14(20):1-28.

deRossett, L. (2013). 'Grounding Explanations.' *Philosophers' Imprint*, 13(7):1-26.

Duncan, M., Miller, K. & Norton, J. (ms). 'How to wear the crazy trousers: defending flatland.'

Elga, A. (2000). 'Statistical Mechanics and the Asymmetry of Counterfactual Dependence.' *Philosophy of Science*, 68(Supplement):313–324.

Feigl, H. (1970), 'The 'Orthodox' View of Theories: Remarks in Defense as well as Critique.' In M. Radner and S. Winokur (eds.), *Minnesota Studies in the Philosophy of Science, Volume IV*. Minneapolis: University of Minnesota Press.

Fine, K. (1994), 'Ontological Dependence.' *Proceedings of the Aristotelian Society*, 95:269–290.

Fine, K. (2010). 'Some Puzzles of Ground.' *Notre Dame Journal of Formal Logic*, 51:97-118.

Fine, K. (2012). 'Guide to Ground.' In F. Correia and B. Schnieder (eds.), *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge: Cambridge University Press.

FitzPatrick, W. J. (2015). 'Debunking Evolutionary Debunking of Ethical Realism.'

Philosophical Studies, 172(4):883-904.

- Gopnik, A., Glymour, C., Sobel, D., Schulz, L., Kushnir, T., and Danks, D. (2004). 'A theory of causal learning in children: Causal maps and Bayes nets.' *Psychological Review*, 111:1–30.
- Gould, S. J. (1991). 'Exaptation: A Crucial Tool for an Evolutionary Psychology.' *Journal of Social Issues*, 47(3): 43-65.
- Guthrie, S. (1993). *Faces in the clouds*. Oxford: Oxford University Press.
- Hagmayer, Y., Sloman, S., Lagnado, D., and Waldman, M. (2007). 'Causal Reasoning Through Intervention.' In A. Gopnik and L. Schultz (eds.) *Causal learning: Psychology, Philosophy and Computation*. Oxford: Oxford University Press.
- Hale, B. and Wright, C. (1992). 'Nominalism and the Contingency of Abstract Objects.' *The Journal of Philosophy* 89(3): 111-135.
- Harman, G. (1977). *The Nature of Morality*. New York: Oxford University Press.
- Hawking, S. and Mlodinow, L. (2010). *The Grand Design*. New York: Bantam Books.
- Hempel, C. G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. New York: Free Press.
- Hempel, C. G. (1966). *Philosophy of Natural Science*. Englewood Cliffs, N.J.: Prentice-Hall.
- Hempel, C. G. and Oppenheim, P. (1948). 'Studies in the Logic of Explanation.' *Philosophy of Science*, 15:135–175.
- Hitchcock, C. (1995). 'Discussion: Salmon on Explanatory Relevance.' *Philosophy of Science*, 62:304–20.
- Hornsby, J. (2005). 'Truth without Truth-making Entities.' In H. Beebe and J. Dodd (eds.) *Truth-makers: The Contemporary Debate*. Oxford: Clarendon Press.
- Jenkins, C. S. I. (2011). 'Is Metaphysical Dependence Irreflexive?' *The Monist*, 94(2):267–276.
- Joyce, R. (2006). *The Evolution of Morality*. Cambridge: MIT Press.

- Kim, J. (1993). *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Kitcher, P. (1981). 'Explanatory Unification.' *Philosophy of Science*, 48:507–531.
- Kitcher, P. (1989). 'Explanatory Unification and the Causal Structure of the World.' In P. Kitcher and W. Salmon (eds.) *Scientific Explanation*. Minneapolis: University of Minnesota Press.
- Koslicki, K., forthcoming, 'The Coarse-Grainedness of Grounding.' In K. Bennett and D. Zimmerman (eds.), *Oxford Studies in Metaphysics*. Oxford: Oxford University Press.
- Kovacs, D. (forthcoming). 'Grounding and the Argument from Explanatoriness.' *Philosophical Studies*.
- Kushnir, T., Gopnik, A., Lucas, C. and Schulz, L. (2010). 'Inferring Hidden Causal Structure.' *Cognitive Science*, 34(1):148-160
- Lagnado, D. A., and Sloman, S. A. (2004). 'The advantage of timely intervention.' *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30:856–876.
- Leuenberger, S. (2014). 'Grounding and Necessity.' *Inquiry*, 57: 151–174.
- Lewis, D. (1986a). *On the Plurality of Worlds*. Oxford: Blackwell.
- Lewis, D. (1986b) *Philosophical Papers vol. II*. Oxford: Oxford University Press.
- Liggins, David (2005). 'Truthmakers and explanation.' In H. Beebe and J. Dodd (eds.), *Truthmakers: The Contemporary Debate*. Oxford: Clarendon Press.
- Litland, J. E. (2013). 'On Some Counterexamples to the Transitivity of Grounding.' *Essays in Philosophy*, 14:19–32.
- Lowe, E. J. (2010). 'Ontological Dependence.' In E. N. Zalta (ed.), *The Stanford encyclopedia of philosophy* (Spring 2010 ed.).
<http://plato.stanford.edu/archives/spr2010/entries/dependence-ontological/>
- Maudlin, T. (2007). *The Metaphysics Within Physics*. Oxford: Oxford University Press.
- McFetridge, I. G. (1990). 'Truth, Correspondence, Explanation and Knowledge.' In J.

- Haldane and R. Scruton (eds.) *Logical Necessity and other essays*. London: Aristotelian Society.
- Michotte. A. (1956). *The perception of causality*. (trans. T. R. Miles). London: Methuen.
- Miller, K. (2012). 'Mathematical Contingentism.' *Erkenntnis*, 77(3):335-359.
- Mumford, S. (1998). *Disposition.*, Oxford: Oxford University Press.
- Newton, I. (1962). *The Mathematical Principles of Natural Philosophy*. (trans. A. Motte and F. Cajori). Berkeley: University of California Press.
- Nisbett, R. and Ross, L. (1980). *Human inference; Strategies and shortcomings of social judgement*. Englewood Cliffs, N.J.: Prentice-Hall.
- Orilia, F. (2009). 'Bradley's Regress and Ungrounded Dependence Chains: A Reply to Cameron.' *dialectica*, 63(3):333–341.
- Paseau, A. (2010). 'Defining Ultimate Ontological Basis and the Fundamental Layer.' *The Philosophical Quarterly*, 60(238):169-175.
- Plato (2002). *Five Dialogues: Euthyphro, Apology, Crito, Meno, Phaedo*. (trans. G. M. A. Grube). Second edition. Revised by J. M. Cooper. USA: Hackett Publishing.
- Price, H. (1996). *Time's Arrow & Archimedes' Point: New Directions for the Physics of Time*. Oxford: Oxford University Press.
- Puryear, S. (2014). 'Finitism and the Beginning of the Universe.' *Australasian Journal of Philosophy*, 92:619-629.
- Pyysiäinen, I. (2001). *How religion works: Towards a new cognitive science of religion*. Cognition and Culture Book Series, 1. Leiden: Brill.
- Pyysiäinen, I. & Anttonen, V. (2002). *Current approaches in the cognitive science of religion*. London: Continuum.
- Quine, W. V. (1948). 'On What There Is.' *The Review of Metaphysics*, 2(1):21–38.
- Raven, M. J. (2012). 'In defence of ground.' *Australasian Journal of Philosophy*, 90(4):687-701.

- Raven, M. J. (2015). 'Ground.' *Philosophy Compass*, 10(5):322-333.
- Raven, M. J. (forthcoming). 'New Work for a Theory of Ground.' *Inquiry*.
- Redelmeier, D. A. and Tversky, A. (1996). 'On the belief that arthritis pain is related to the weather.' *Proceedings of the National Academy of Science*, 93(7):2895-6.
- Rigdon, M., Ishii, K., Watabe, M., and Kitayama, S. (2009). 'Minimal Social Cues in the Dictator Game.' *Journal of Economic Psychology*, 30:358-367.
- Rodriguez-Pereyra, G. (2005). 'Why Truthmakers?' In H. Beebe and J. Dodd (eds.), *Truthmakers: The Contemporary Debate*. Oxford: Clarendon Press.
- Rodriguez-Pereyra, G. (2015). 'Grounding is Not a Strict Order.' *Journal of the American Philosophical Association*, 1(3):517-534.
- Rosen, G. (2010). 'Metaphysical Dependence: Grounding and Reduction.' In R. Hale and A. Hoffman (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford University Press.
- Ruben, D. (1990). *Explaining Explanation*. London: Routledge.
- Russell, B. (1903). *The Principles of Mathematics*. London: George Allen & Unwin.
- Salmon, W. (1971). 'Statistical Explanation.' In W. Salmon, (ed.) *Statistical Explanation and Statistical Relevance*. Pittsburgh: University of Pittsburgh Press.
- Salmon, W. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton: Princeton University Press.
- Sayre-McCord, G. (1988). 'Moral Theory and Explanatory Impotence.' *Midwest Studies in Philosophy*, Vol. XII. Minneapolis: University of Minnesota Press.
- Schafer, G. (1996). *The Art of Causal Conjecture*. Cambridge, Mass: MIT Press.
- Schaffer, J. (2005). 'Contrastive Causation?.' *Philosophical Review*, 114:327-358.
- Schaffer, J. (2009). 'On What Grounds What.' In D. Manley, D. Chalmers, and R. Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford: Oxford University Press.
- Schaffer, J. (2010). 'Monism: The Priority of the Whole.' *Philosophical Review*, 119(1):31–

- Schaffer, J. (2012). 'Grounding, Transitivity and Contrastivity.' In F. Correia and B. Schneider (eds.), *Grounding and Explanation*. Cambridge: Cambridge University Press.
- Schaffer, J. (2016). 'Grounding in the Image of Causation.' *Philosophical Studies*, 173(1):49-100.
- Schnieder, B. (2006). 'Truth-Making without Truth-Makers.' *Synthese*, 152:21–46.
- Shaklee, H. and Mims, M. (1982). 'Sources of error in judging event covariation: Effects of memory demands.' *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 8:208-224.
- Shaklee, H. and Tucker, D. (1980). 'A rule analysis of judgements of covariation between events.' *Memory and Cognition*, 8:459-467.
- Skiles, A. (2015). 'Against Grounding Necessitarianism.' *Erkenntnis*, 80(4):717-751.
- Sloman, S. (2005). *Causal Models: how people think about the world and its alternatives*. Oxford: Oxford University Press.
- Smedlund, J. (1963). 'The concept of correlation in adults', *Scandinavian Journal of Psychology*, 4:165-173.
- Street, S. (2006). 'A Darwinian Dilemma for Realist Theories of Value', *Philosophical Studies*, 127:109–66.
- Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Cambridge, M.A: Harvard University Press.
- Styvers, M., Tenenbaum, J. B., Wagenmakers, E. J. and Blum, B. (2003). 'Inferring causal networks from observations and interventions.' *Cognitive Science*, 27(3):453-489.
- Tallant, J. (2015). 'Ontological dependence in a spacetime world.' *Philosophical Studies*, 172(11):3101-3118.

- Trogon, K. (2013a). 'An Introduction to Grounding.' In M. Hoeltje, B. Schnieder, and A. Steinberg (eds.), *Varieties of Dependence: Ontological Dependence, Grounding, Supervenience, Response-Dependence*, Basic Philosophical Concepts, München: Philosophia.
- Trogon, K. (2013b). 'Grounding: Necessary or Contingent?' *Pacific Philosophical Quarterly*, 94:465–485.
- Trogon, K. (ms-a). 'Grounding and Explanation.'
- Trogon, K. (ms-b). 'Do grounds explain what they ground?'
- Trout, J. D. (2002). 'Scientific Explanation and the Sense of Understanding.' *Philosophy of Science*, 69(2):212-233.
- Van Fraassen, B. (1980). *The Scientific Image*. Oxford: Oxford University Press.
- Van Fraassen, B. (1989) *Laws and Symmetry*. Oxford: Clarendon Press.
- Waldmann, M. R. (1996). 'Knowledge-based causal induction.' In D. R. Shanks, K. J. Holyoak and D. L. Medin (eds), *The psychology of learning and motivation, Vol 34: Causal learning*. San Diego: Academic Press.
- Waldmann, M. R. and Hagmayer, Y. (2013). 'Causal Reasoning.' In D Reisman (ed.), *The Oxford Handbook of Cognitive Psychology*. Oxford: Oxford University Press.
- Wallace, M. (2011a). 'Composition as Identity: Part 1.' *Philosophy Compass*, 6(11):804-816.
- Wallace, M. (2011b). 'Composition as Identity: Part 2.' *Philosophy Compass*, 6(11):817-827.
- Williams, J. J. and Griffiths, T. L. (2013). 'Why are people so bad at detecting randomness? A statistical argument.' *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39(5):1473-90.
- Wilsch, T. (2015). 'The Nomological Account of Ground.' *Philosophical Studies*, 172:3293–3312.
- Wilsch, T. (2016). 'The Deductive-Nomological Account of Metaphysical Explanation.' *Australasian Journal of Philosophy*, 94(1):1-23.
- Wilson, A. (forthcoming). 'Metaphysical Causation.' *Noûs*.

Wilson, A. (ms). 'Classifying Dependencies.'

Wilson, J. (2014). 'No Work for a Theory of Grounding.' *Inquiry*, 57(5–6):1–45.

Woodward, J. (2003). *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

APPENDIX

Statement of originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes.

I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged. See below authorship attribution statement.

James Norton

Authorship attribution statement:

This thesis contains material from the following papers:

(1) Miller, K. and Norton, J., forthcoming. 'Grounding: it's (probably) all in the head', in *Philosophical Studies*.

(2) Miller, K. and Norton, J. (ms.) 'A Psychologistic Account of Metaphysical Explanation'.

(3) Baron, S. and Norton, J. (ms.) 'Studies in the Logic of Non-Causal Explanation'.

(1) forms the basis of Chapter 4, (2) forms the basis of Chapter 5, and (3) forms the basis of Chapter 6.

The order of authors on these papers is alphabetical. In each case, each author contributed equally to the work, both in terms of its conceptual inception, the development of the core argument of the paper, the literature review, and the writing up of the material. There is no sense in which any of the sections of the paper can be attributed to one of us, singly: all sections of the paper are the result of collaboration and joint writing. I have included on the following pages written permission from my co-authors to include this material in the thesis.

Nov 2, 2016

To whom it may concern,

This is to certify that James Norton has permission to use work jointly authored by each of us, in his PhD thesis.

The paper “Grounding, it’s (probably) all in the head” was written jointly, by both of us, and each of us contributed equally to the work, both in terms of its conceptual inception, the development of the core argument of the paper, the literature review, and the writing up of the material. There is no sense in which any of the sections of the paper can be attributed to one of us, singly: all sections of the paper are the result of collaboration and joint writing. As per convention in philosophy, the order of authors is alphabetical.

Chapter 4 of the thesis incorporates most of the material in that jointly written paper.

The paper “A psychologistic account of metaphysical explanation” was written jointly, by both of us, and each of us contributed equally to the work, both in terms of its conceptual inception, the core argument of the paper, the literature review, and the writing up of the material. There is no sense in which any of the sections of the paper can be attributed to one of us, singly: all sections of the paper are the result of collaboration and joint writing. As per convention in philosophy, the order of authors is alphabetical.

Chapter 5 of the thesis incorporates most of the material in that jointly written paper.

Kristie Miller

Joint Director, the Centre for Time

Department of Philosophy

The University of Sydney

Nov 2, 2016

To whom it may concern,

This is to certify that James Norton has permission to use work jointly authored by each of us, in his PhD thesis.

The paper “Studies in the Logic of Non-Causal Explanation” was written jointly, by both of us, and each of us contributed equally to the work, both in terms of its conceptual inception, the core argument of the paper, the literature review, and the writing up of the material. There is no sense in which any of the sections of the paper can be attributed to one of us, singly: all sections of the paper are the result of collaboration and joint writing. As per convention in philosophy, the order of authors is alphabetical. Chapter 6 incorporates some of the material in that jointly written paper.

Samuel Baron

Lecturer, Philosophy

University of Western Australia