Metaphysical Causation

Alastair Wilson
University of Birmingham & Monash University
a.j.wilson@bham.ac.uk

14th November 2013

Abstract: There is a systematic and suggestive analogy between grounding and causation. In my view, this is no coincidence: grounding and causation resemble one another because grounding is a type of causation. I defend the identification of grounding with metaphysical causation from some initial objections, and I draw on the causation literature to explore connections between grounding and counterfactuals. I outline a non-reductive counterfactual theory of grounding along interventionist lines, and use it to diagnose the prevalence of scepticism about grounding as deriving, at least in part, from scepticism about non-trivial counterpossible counterfactuals.

1. Introduction
2. Grounding & Causation
3. Causal Production & Causal Dependence
4. Dependence Counterfactuals
5. Metaphysical Causal Models
6. Summarizing the Grounding-Causation Analogy
7. Counterpossible Dependence
8. Conclusion

1. Introduction

> "Grounding is something like metaphysical causation."
>
> Schaffer (2012) p.122

I think Schaffer here understates the intimacy of the connection between grounding and causation. The main thesis of this paper is that grounding *just is* metaphysical causation. Call the claim that grounding is a type of causation ***G=MC***. According to G=MC, the grounding relation is a special case of the causal relation: when A grounds B, A is a (metaphysical) cause of B and B is a (metaphysical) effect of A.

Treating grounding as metaphysical causation has two major theoretical benefits:

- G=MC is ideologically parsimonious. If grounding is just metaphysical causation, then we don't need a separate theory of grounding invoking new primitive notions. Instead, our account of grounding will invoke only whatever ideology is employed by our best theory of causation[1].

- G=MC provides a straightforward explanation (or, alternatively, removes the need for any explanation) of why grounding claims are explanatory. Grounding explanations are a type of causal explanations, and we can account for their explanatory power in whatever way we usually account for the explanatory power of causal explanations.

Since we are working at a level of abstractness where we need to compare best total theories, the having of a theoretical benefit constitutes evidence in favour of a theory. These beneficial features can accordingly form the basis of a case for G=MC. But this paper also provides more indirect support for G=MC, by highlighting a systematic analogy between grounding and ordinary causation. These relations[2] have the same logical properties (which can be challenged in analogous ways); they have the same connections to explanation and to counterfactuals; and the same puzzle cases and theoretical issues arise when we try to give them a counterfactual analysis.[3] In my view, the best explanation of these persistent parallels is that grounding and ordinary causation are different species of the same genus.

Why prefer the thesis that grounding is metaphysical causation to the thesis that causation is physical grounding?[4] Because we have plausible candidates for

[1] In this respect, my proposal resembles those of Wilson (MS) and Hofweber (2009), who argue that grounding relations can be accounted for in terms of antecedently understood ideology such as counterfactual dependence or conceptual inclusion. Those inclined to think that 'grounds' is equivocal in the way Wilson and Hofweber suggest may still treat metaphysical causation as one of the disparate notions we draw together under the heading of grounding.

[2] Perhaps neither grounding nor causation is strictly speaking relational, but is instead best expressed with something like a sentence operator. I will ignore this complication here, as it does not affect my main arguments.

[3] Some of these parallels are also noted by Schaffer (2012).

[4] Thanks to Gonzalo Rodriguez-Pereyra for posing this question.

informative accounts of causation but we lack any good candidates for informative accounts of grounding. Conceiving of causation as a type of grounding, as Karen Bennett does (Bennett MS), infects causation with the persistent unclarity which surrounds grounding; but conceiving of grounding as a type of causation allows grounding theorists to make use of the concepts, distinctions and theoretical tools that have been elaborated in the extensive literature on causation. I hope that the benefits of my proposed approach will make themselves apparent over the course of this paper; of course, the proof of the pudding is in the eating.

Some signposting is in order. Section 2 briefly describes the target of the analysis - the grounding relation - and identifies some core examples on which to test the analysis. Section 3 discusses recent work on causation which distinguishes causal dependence from causal production, and argues that eliding this distinction may lead to unwarranted scepticism about G=MC. In section 4, I explore the fate of certain key counterfactuals associated with metaphysical causal dependence, and discuss some difficulties facing a counterfactual account of grounding. I argue that these problems are familiar from the metaphysics of causation and that they can be handled by more sophisticated counterfactual accounts of causal dependence. One interesting approach of this kind, interventionism, involves the use of causal models; in section 5 I present and discuss a number of metaphysical causal models which are analogues of familiar problem cases from the ordinary causation literature. Section 6 adds two further points of similarity, involving challenges to transitivity and symmetry, and summarizes the systematic analogy between grounding and causation. The counterfactuals implicit in metaphysical causal models will generally include counterpossible conditionals; section 7 explores the connection between the controversy over counterpossibles and the controversy over grounding and offers a diagnosis of the prevalence of scepticism about grounding. Section 8 is a conclusion.

2. Grounding & Causation

Most contemporary treatments (e.g. Rosen 2010) claim to take grounding as a primitive notion. It is not always entirely clear what this means; but, at least, it involves not giving any reductive analysis of the notion in independent terms. Friends of grounding instead typically use non-reductive methods to help non-initiates get a grip on their notion.

As a preliminary step, grounding theorists constrain the notion of grounding by specifying its formal properties. In the classification scheme of Fine (2012), here we will be concerned with the notion of *strict partial ground*, to be distinguished from *weak ground* (non-posteriority as opposed to priority) and from *whole ground* (a complete set of partial grounds). Thought of as a relation between facts, strict partial[5] grounding is usually taken to be a partial ordering: transitive, anti-symmetric, and irreflexive. This provides an initial plausibility check on G=MC, which it passes with no difficulty: the relation *is a cause of*[6] is likewise generally taken to be transitive, anti-symmetric, and irreflexive. We will see in section 5 that these logical properties can be challenged both for grounding and for ordinary causation, and in closely analogous ways.

Thinking of everyday physical causation and metaphysical causation as different species of the same genus - causation - has some immediate consequences for the logical relations between the two species. Assuming for the moment that the genus-level notion *is a cause of* is anti-symmetric, the holding of everyday physical causation in one direction excludes the holding of metaphysical causation in the other direction. If A is the everyday physical cause of B, then B cannot ground A, and vice versa. This requirement seems plausible, and it is predicted by G=MC; in the absence of counterexamples or of other explanations of the requirement[7], its plausibility provides defeasible support for G=MC.

Logical considerations don't take us all that far in understanding grounding. (Divisibility is a partial order on the natural numbers, but it is not the same relation as grounding.) Our grip on grounding is supposed to come in two other main ways: through the connection with explanation, and through examples.

Explanation provides a further significant parallel between causation and grounding. The two notions stand in the same distinctive relation to our

[5] For the remainder of the paper, I will omit the qualifiers 'strict' and 'partial'.

[6] I will not have anything special to say about the locution 'is *the* cause of'; in my view, this locution tends to muddle together a requirement of causal dependence with a distinct requirement of contextual or pragmatic salience. Nor will I discuss the related notion of 'actual causation'.

[7] Karen Bennett's view (set out in Bennett MS) that causation and grounding are both types of 'building relation' provides an alternative explanation of the requirement; my reasons for not going down this route were described in section 1.

practice of explanation: causal relations and relations of ground each figure in explanations, without being literally identical to those explanations. When we want to explain why the bottle broke, we cite the causes of its breaking; and when we want to explain why Singleton Socrates exists, we cite the existence of Socrates. This sort of distinction between causation and causal explanation is familiar from the philosophy of science.

Some grounding theorists distinguish the "worldly relation of grounding from the metaphysical relations between facts that it backs" (Schaffer 2012 p. 124). A similar distinction is can be made with respect to ordinary causation: we can distinguish 'worldly' relations of causation between concrete physical events from 'metaphysical' relations of causation between facts about those events. My proposal depends on the coherence of causal/grounding relations between facts, and it allows for (but does not require) 'worldly' relations in addition.

So far we have found nothing to distinguish ordinary causation from grounding: they have the same logical features (at least by and large) and they connect to explanation in the same ways. Further explication of grounding tends to go by way of example, and the recent literature contains a rich and diverse diet of proposed examples. Here is a representative sample:[8]

Singleton: The existence of Socrates grounds the existence of singleton Socrates.

Double-negation: The truth of P grounds the truth of ¬¬P.

Disjunction: The truth of P grounds the truth of PvQ.

Conjunction: The truth of P grounds the truth of P&Q.

Truthmaking: The existence of Socrates grounds the truth of 'Socrates Exists'.

[8] Too much should not be read into the names of these examples. As we're dealing here with the notion of strict partial ground, these true grounding claims may not fully characterize the metaphysics of the grounded entity. For example, a proper formulation of the moral theory of consequentialism would presumably need to specify that there can be no *other* grounds for the rightness of an act in addition to its consequences.

Mind/body: My being in brain state B grounds my being in mental state M.

Part/whole: The existence of my head grounds my existence.

Consequentialism: Act A's having the best consequences grounds A's being right.

Euthyphro: God's desiring that P grounds its being good that P.

Noether: The symmetry of the laws under time-translation grounds the fact that energy is conserved.

Since these cases are so different from one another, there is plenty of scope to deny that some are genuine instances of grounding. We could follow Ramsey (1927) in thinking ¬¬P just a notational variant on P, and deny Double-negation. Identity theorists deny Mind/Body. And of course, deontologists deny Consequentialism and non-theists deny Euthyphro. Accordingly, it isn't necessary (or desirable) that a theory of grounding should entail that each one of these examples is a genuine case of grounding. But a theory of grounding ought to underwrite these grounding claims at least in the context of the background assumptions which have typically motivated their defenders; otherwise, the theory could reasonably be accused of changing the subject.

Next, we can apply G=MC to the list of examples given above. If each grounding claim is true, then so is each of the following sentences *under a causal reading of 'because'*.

C-Singleton: Singleton Socrates exists because Socrates does.

C-Double-negation: ¬¬P is true because P is true.

C-Disjunction: PvQ is true because P is true.

C-Conjunction: P&Q is true because P is true.

C-Truthmaking: 'Socrates exists' is true because Socrates exists.

C-Mind/body: I'm in mental state M because I'm in brain state B.

C-Part/whole: I exist because my head exists.

C-Act-consequentialism: A is the right action because it has the best consequences.

C-Euthyphro: P is good because God desires it.

C-Noether: Energy is conserved because the laws are symmetric under time-translation.

Of course, G=MC cannot be supported merely by displaying a list of examples. Opponents of G=MC will no doubt respond either that each C-sentence is either false or else only true under an (alleged) non-causal reading of 'because'. Rather, the purpose of this list is to throw up a few immediate points for discussion.

First, there appear to be a wide variety of kinds of fact amongst the causal relata. Facts linked by causation according to the C-sentences include facts about concrete entities (Socrates), facts about abstract entities (Singleton Socrates), and facts about entities which are neither clearly abstract nor clearly concrete (God, laws of nature). We might reasonably doubt whether any plausible theory of causation will be able to encompass such a mixed bag of relata. This *challenge from relata heterogeneity* is addressed in the next section.

Second, some of the causal relata are abstract entities. It is widely, although not universally, held that abstracta do not enter into causal relations. Indeed, acausality is one of the options that David Lewis considers when seeking to characterize the abstract/concrete distinction (Lewis 1986). G=MC entails that abstracta (assuming that they ground or are grounded) do enter into causal relations; consequently G=MC appears not to do justice to the familiar thought that (at least some) abstract objects are outside the causal order.[9] This *challenge from abstract inactivity* is also addressed in the next section.

[9] As Nicholas Jones has pointed out to me, versions of this argument could be run with respect to specific kinds of abstract objects (such as sets), or with respect to some feature of some abstract objects (such as lack of spatial location). My response will also apply to these variant arguments.

Third, one relatively natural response to the above examples is to interpret them as metaphorical. It isn't literally the case that Socrates' existence causes the existence of Singleton Socrates, according to this response; rather, the nature of their relationship is in certain ways analogous to a causal relationship, which licenses the pretence that the one is the cause of the other. We might claim heuristic value for the pretence, while denying that it should be part of sober metaphysical theorizing. This *challenge from metaphor* is partially addressed in the next section, but my main response to it relies on the interventionist account of grounding which is sketched in sections 4 and 5.

3. Causal Production & Causal Dependence

In this section we will look at a distinction made in the recent literature on causation, before applying it in response to the challenges from relata heterogeneity, abstract inactivity, and metaphor. The distinction I have in mind is between causal production and causal dependence. The distinction is defended forcefully by Ned Hall (Hall 2004); other related distinctions are proposed in Sober (1985), Hitchcock (2003) and Strevens (forthcoming).[10] Hall argues that the following five claims about causation, while all apparently platitudinous, are not jointly satisfiable by any one notion of causation.

- **Transitivity**: If event[11] c is a cause of d, and d is a cause of e, then c is a cause of e.
- **Locality**: Causes are connected to their effects via spatiotemporally continuous sequences of causal intermediates.
- **Intrinsicness**: The causal structure of a process is determined by its intrinsic, non-causal character (together with the laws).
- **Dependence**: Counterfactual dependence between wholly distinct events is sufficient for causation

[10] (Strevens forthcoming) discusses Hall's distinction but prefers different terminology: influence vs. difference-making. Further varieties of 'causal pluralism' are surveyed by (Godfrey-Smith 2010). See also (Cartwright 2004) and (Psillos 2009).

[11] Although Hall formulates his principles in terms of events rather than of facts, he cannot be presupposing any overly restrictive view of the causal relata since he allows for omissions to be causes and effects.

- **Omissions**: Omissions—failures of events to occur—can both cause and be caused.

Hall's argument appeals to a group of examples which have played a central role in the philosophy of causation. There are intuitive cases of causation - including those known as *double prevention* and *causation by omission* - which seem to violate Transitivity, Locality and Intrinsicness. But these three theses are frequently relied on to deal with the threat from cases of *overdetermination*, including cases of *pre-emption*. (These problem cases are outlined in section 5, but for now we can set the details to one side.)

Hall's diagnosis is that there are two different concepts of causation in play, with Transitivity, Locality and Intrinsicness true of one (*causal production*) and Dependence and Omissions true of the other (*causal dependence*). On this picture, there might be a variety of ways in which causal dependence can obtain: it need not go via causal production.

We need not here explore the complexities of analyzing causal production. But to give the flavour of the idea, it may help to mention some specific proposals which are naturally seen as attempts to analyze production. There are the mark-transference theories of (Reichenbach 1958) and (Salmon 1984), and the conserved-quantity transference theories associated with (Fair 1979), (Skyrms 1980), (Dowe 1992) and (Salmon 1994). These proposals generally look for some specific feature of a process which renders it suitable to transmit information, and then dub processes with that feature *causal processes*. Production is then analyzed in terms of appropriate chains of causal processes. Production also resembles the notorious notion of 'biff' employed informally by David Armstrong, and discussed by Lewis (2000) and (Handfield *et al.* 2008).

The second concept of causation distinguished by Hall is the concept of causal dependence. Here Hall's proposal is very simple: causal dependence is just counterfactual dependence. But, as we shall see in the next section, this only holds given a specially-crafted theory of counterfactuals which excludes *backtracking* counterfactuals. This exclusion threatens the reductive ambitions of an analysis of causal dependence in terms of counterfactual dependence. But, we can set these complications aside for the time being, and regard causal dependence as involving specific patterns of counterfactual dependence - never mind what precise mechanisms (or lack thereof!) give rise to these patterns.

On the view I propose, metaphysical causation need not involve any process of causal production - no metaphysical biff! - but it does need to involve characteristically causal patterns of counterfactual dependence. We need to interpret G=MC accordingly: grounding is to be identified with metaphysical causal dependence rather than with metaphysical causal production.

We are now in a position to respond to the three challenges from the previous section. In each case, the challenge would be apt if G=MC were an identification of grounding with metaphysical causal production. But once we understand G=MC as identifying grounding with metaphysical causal dependence, then the challenge loses its force.

The challenge from relata heterogeneity is a threat to the identification of grounding with causal production. Production typically involves certain very specific kinds of causal processes, the sorts of processes which involve energy transfer and are studied by fundamental physics. Since fundamental physics concerns itself only with a sparse and highly-restricted subset of entities and properties, it's hard to see how facts as heterogenous as those in our examples could enter into a relation of causal production. But, with Hall's distinction on board, we can grant this point while still allowing facts about highly heterogenous entities to enter into relations of causal dependence. How this works will become clearer in the next two sections, as we settle on a specific theory of causal dependence and go on to apply it to metaphysical causation.

The challenge from abstract inactivity is likewise blunted by taking G=MC to concern causal dependence rather than causal production. As I see it, the motivation for thinking that abstract objects are acausal is that abstract objects don't appear to engage in the kinds of activity which can sustain causal production. More bluntly: abstract causal production is a crazy idea. Abstracta do not have mass or couple to quantum fields. Will we then propose new dynamical theories for abstract objects, positing metaphysical forces between them? No; a parallel physics of abstracta is a bad plan. But we can do full justice to this thought via a prohibition on abstract causal production, while still allowing for abstract causal dependence. This is because causal dependence need not rest on any productive connection between cause and effect. A classic example of this is causation by omission. But more generally, we can deny that dependence need go via production, undermining the challenge from abstract inactivity.

The challenge from metaphor requires a slightly different treatment. While it is certainly very plausible that talk of causal production by abstract entities is metaphorical in nature, it may seem plausible enough that talk of causal dependence is *also* metaphorical. Causation by omission, to which I just appealed in response to the challenge from abstract inactivity, is frequently explained away as metaphorical. For example, Liebesman (2011) proposes such a move as an alternative to Lewis' denial (Lewis 2004b) that causation is a relation. Our distinction between production and dependence therefore provides at most a partial response to the challenge from metaphor. I propose to meet this challenge in a different way, by first endorsing an account of causal dependence which undermines objections to a literal reading of causal dependence claims involving omissions and related cases, and by then applying this account to metaphysical causal dependence. That is the task of the next two sections.

4. Metaphysical Dependence Counterfactuals

The simplest counterfactual analysis of causation is the early theory of Lewis (1973). Lewis defines causation as the ancestral of counterfactual dependence, where counterfactual dependence of P on Q requires the truth of $\neg Q \ \Box\!\!\rightarrow \ \neg P$. (This account draws directly on one of Hume's 'two definitions of cause' (Hume 1748).)

Here are the relevant counterfactuals corresponding to our examples[12]:

CF-Singleton: If Socrates hadn't existed, nor would have Singleton Socrates.

CF-Double-negation: If P hadn't been true, nor would $\neg\neg P$ have been.

CF-Disjunction: If P hadn't been true, nor would $P \vee Q$ have been.

CF-Conjunction: If P hadn't been true, nor would $P \& Q$ have been.

[12] These examples are posed in the past tense (hadn't) instead of the present tense (were not to). I think this makes judgments clearer without affecting any substantive issues, and would invite readers who disagree to explain why the tense matters.

CF-Truthmaking: If Socrates hadn't existed, 'Socrates exists' would not have been true.

CF-Mind/body: If I hadn't been in brain state B, I wouldn't have been in mental state M.

CF-Part/whole: If my head hadn't existed, I wouldn't have existed.

CF-Act-consequentialism: If A hadn't had the best consequences, A wouldn't have been right.

CF-Euthyphro: If God hadn't desired that P, P wouldn't have been good.

CF-Noether: If the laws hadn't been symmetric under time-translation, then energy would not have been conserved.

Some of these counterfactuals seem fine: for example, CF-Singleton, CF-Double-negation, and CF-Noether. This suggests we are on the right track. But it looks like there are problems with others of them, of two different kinds:

- Some of the antecedents may be metaphysically impossible, in which case those counterfactuals are *counterpossibles*.

- Some of the counterfactuals seem to have the wrong truth conditions. Perhaps is P hadn't been true, Q would have been true, in which case PvQ would still have been true. Or perhaps if I hadn't been in brain state B, I might have been in a very similar state B*, in which case I'd still have been in mental state M. (See Yablo 2004 and Menzies & List 2009 for more discussion).

I'll postpone discussion of the first of these issues until section 7. The second issue will be dealt with via the interventionist approach described later in this section. First, we have a more urgent issue to confront. Even if the CF counterfactuals do hold, there might nevertheless fail to be metaphysical causal dependence as a result of the truth of *additional* counterfactuals.

Since grounding is usually taken to be anti-symmetric, if G=MC is correct then metaphysical causal dependence must likewise be anti-symmetric (at least

for the most part). Therefore, in addition to the holding of a given CF counterfactual, a simple counterfactual account will require the *failure* to hold of the corresponding RCF counterfactual:

RCF-Singleton: If Singleton Socrates hadn't existed, Socrates would not have existed either.

RCF-Double-negation: If ¬¬P hadn't been true, P would not have been true either.

RCF-Disjunction: If PvQ hadn't been true, P would not have been true either.

RCF-Conjunction: If P&Q had not been true, P would not have been true either.

RCF-Truthmaking: If 'Socrates exists' had been false, Socrates wouldn't have existed.

RCF-Mind/body: If I hadn't been in mental state M, I wouldn't have been in brain state B.

RCF-Part/whole: If I hadn't existed, my head wouldn't have existed.

RCF-Act-consequentialism: If A hadn't been right, it wouldn't have had the best consequences.

RCF-Euthyphro: If P hadn't been good, God wouldn't have desired it.

RCF-Noether: If energy hadn't been conserved, the laws would not have been symmetric under time-translation symmetry.

Unfortunately, many of these RCF counterfactuals seem to be as plausible, or nearly as plausible, as their CF counterparts. This looks like a challenge for defenders of G=MC; if the RCF counterfactuals are true, and if their truth suffices for causal dependence (as Hall's principle Dependence suggests it does), then G=MC delivers us widespread two-way metaphysical causal dependence.

That could be used as a *reductio* of the very idea of metaphysical causation, and accordingly (at least for grounding enthusiasts) as a *reductio* of G=MC.

However, that conclusion would be much too hasty. Similar problems with apparently-true reverse counterfactuals afflict counterfactual analyses of ordinary causation even in the simplest cases. Had the window not smashed, it would have been because no brick collided with it. But there would be no realistic chance of my missing if I threw, since I am standing right by the window. So, had the window not smashed, I would not have thrown the brick. It is a familiar point that the apparent truth of this latter counterfactual should not lead us to conclude that the smashing of the window caused me to throw the brick at it.

The standard way of dealing with this problem, for counterfactual analyses of causation, is to restrict the analysis so as to associate causal dependence only with a certain class of counterfactuals, a class which does not include the problematic smashing-to-throwing counterfactual. Lewis dubbed the problematic class *back-tracking* counterfactuals, and restricted his 1973 account so that only non-back-tracking counterfactuals could be sufficient for causal dependence. In combination with a specific semantic treatment of non-back-tracking counterfactuals, this account successfully excludes the most obvious problem cases.

The word 'back-tracking' doesn't properly capture what's wrong with the RCF reverse metaphysical dependence counterfactuals. Unlike the smashing-to-throwing counterfactual, the RCF counterfactuals don't track back in time from the (supposed) cause and then forward again to the (supposed) effect; they instead track *down* in the 'order of being' from the (supposed) cause and then back *up* to the (supposed) effect. So we might call them *down-trackers*, and then use the collective term *wrong-tracker* to cover both back-trackers and down-trackers.

If G=MC is on the right lines, it suggests that back-tracking and down-tracking are different ways of wrong-tracking; that there is a unified class of non-wrong-tracking, or *right-tracking*, counterfactuals which sustain genuine relationships of causal dependence. We can test this hypothesis by considering a syntactic feature associated with back-trackers, described by Lewis as follows:

> Back-tracking counterfactuals, used in a context that favors their truth, are marked by a syntactic peculiarity. They are the ones in which the usual subjunctive conditional constructions are readily replaced by more complicated constructions: "If it were that. .. then it would have to be that ... " or the like.
>
> Lewis (1979) p.458

This feature is also had by down-tracking counterfactuals. The RCF counterfactuals listed above are indeed more idiomatically posed with the more complicated forms Lewis refers to. If Socrates' singleton hadn't existed, then it *would have to have been that* Socrates didn't exist; if ¬¬P hadn't been true, P *couldn't have been* true either; if energy had not been conserved, the laws *would have had to have been* non-symmetric under time-translations. The CF counterfactuals, in contrast, are if anything less felicitous when posed in these more complicated forms and certainly do not gain in felicity to the same extent.

An adequate counterfactual analysis of causation must provide a natural, informative and non-*ad-hoc* characterization of right-tracking counterfactuals. In the case of ordinary causation, we could try to pick out right-trackers by reference to time variables somehow associated with the antecedent and the consequent; we simply specify that the antecedent-time must be earlier than the consequent-time. This move is transparently hopeless in the case of metaphysical causation. We (perhaps!) have a grasp on the location of an event in a time series that is independent of causal facts about it; but we have no grasp whatsoever of the level of a fact in the metaphysical order of being that is independent of our grasp on which facts it grounds or is grounded.

Lewis hoped to avoid making the temporal asymmetry of counterfactual dependence (and hence of causal dependence) into a necessary truth about causation (Lewis 1979/1986). Instead, he hoped to exclude back-trackers by appeal to contingent features of worlds like ours, which he thought would infect back-trackers with widespread indeterminacy (Lewis 1979/1986). *If I hadn't flicked the switch, the light would not have gone on.* This counterfactual is determinately true: the closest antecedent-worlds will all be pretty similar, and in none of them the light goes on anyway. But the reverse counterfactual, Lewis argued, is not determinately true. A large number of widely different alternative courses of events could have given rise to the light not going on; my not flicking, a power cut, a blown bulb, a loose connection. The closest antecedent worlds are varied, and there will be very little true at them all.

Will anything like this Lewisian indeterminacy-based manoeuvre work to distinguish right-tracking from wrong-tracking counterfactuals in full generality? No; even if it were successful in the case of ordinary causation (and it is not[13]), the manoeuvre would not carry over to the case of metaphysical causation. The asymmetry of traces in the actual world, as we have learned from thermal physics, is intimately tied to the monotonic increase in entropy in closed macroscopic systems. There is no analogue of entropy increase as we go up the order of being: no quantity which is determined in a lawlike way to be greater for a grounding entity than for the grounded entity. So there is no physical basis for any asymmetry of traces in the metaphysical order of being. Absent any other reason to believe reality has the relevant feature, the Lewisian indeterminacy-based manoeuvre does not get off the ground.

One possible response to the difficulties with characterizing right-tracking is to capitulate, and to give up the goal of analyzing causation in non-causal terms. We could characterize the right-tracking counterfactuals as those where the consequent is causally dependent on the antecedent. Any resulting counterfactual theory of causation would then be so uninformative that it could scarcely qualify as an analysis; but perhaps this is the best we can do. To 'take causation as primitive' is to stop here and to say no more; and giving up in this way on the project of the counterfactual analysis of causation does not undermine G=MC. It does not threaten the analogy between grounding and ordinary causation emphasized throughout the paper, and it does not vitiate the theoretical benefits of identifying grounding with metaphysical causation set out in section 1. However, I am mostly interested in more ambitious approaches in the remainder of this paper, in the hope that more informative accounts of causation can cast independent light on the nature of grounding.

One prominent approach of a more ambitious kind is the *interventionism* associated with Woodward (2003), Hitchcock (2001), and Pearl (2009). Unlike the Lewisian approach, interventionism does not comprise a reduction of causation to counterfactual dependence, but it is still a form of counterfactual theory since it involves a non-trivial 'systematic connection between causal claims and certain counterfactuals' (Woodward 2003, p. 70). As with the Lewisian counterfactual theory, the sort of counterfactual dependence is sufficient for causation must be restricted in order to prevent wrong-trackers

[13] Adam Elga has persuasively argued that the asymmetry-of-traces account is hopeless; see Elga (2001).

from giving rise to spurious causation. To encode the distinction between right-tracking and wrong-tracking counterfactuals, interventionists make use of causal models - ordered pairs <V,E> of sets V of variables and sets E of 'structural equations' constraining the possible values that each variable can take as a function of the values of the other variables. Roughly[14], in right-tracking counterfactuals the antecedent specifies that an 'intervention' sets some variables to some particular values and the consequent specifies that some other variables take some other particular values. For the standard philosophical account of interventionist counterfactuals, see Woodward (2003); for explorations of their semantics, see Briggs (2012) and Santorio (MS).

The notion of an intervention does a lot of work for interventionists. It effectively plays the role allotted to *small miracles* in the Lewisian semantics for right-tracking counterfactuals, the role of specifying that the antecedent be realized in a way which doesn't 'drag along' unwanted causal history. An intervention is a 'clean' alteration of the value of a particular variable which does not affect the values of upstream causal variables: for example, an intervention on the reading of a barometer leaves unchanged both the pressure in the room and the barometer's own causal origins. Here is Woodward's official definition of an intervention:

(**IV**) *I* is an intervention variable for *X* with respect to *Y* iff

> 1. *I* causes *X*;
> 2. *I* acts as a switch for all other variables that cause *X*. That is, certain values of *I* are such that when *I* attains those values, *X* ceases to depend on the values of other variables that cause *X* and instead depends only on the value taken by *I*;
> 3. Any directed path from *I* to *Y* goes through *X*. That is, *I* does not directly cause *Y* and is not a cause of any causes of *Y* that are distinct from *X* except, of course, for those causes of *Y*, if any, that are built into the *I* → *X* → → *Y* connection itself; that is, except for (a) any causes of *Y* that are effects of *X* (i.e., variables that are causally between *X* and *Y*) and (b) any causes of *Y* that are between *I* and *X* and have no effect on *Y* independently of *X*;
> 4. *I* is (statistically) independent of any variable *Z* that causes *Y* and that is on a directed path that does not go through X.
>
> Woodward (2003), p.98

[14] For the full details, see proposal M on p.59 of Woodward (2003).

It will immediately be apparent that this sort of characterization will not issue in a reductive theory of causation, since the notion of an intervention is explicitly causal[15]. Nonetheless, interventionists typically maintain that their account can still be informative because it shows us how various different causal claims are conceptually connected to one another. The approach will deliver verdicts about specific causal dependencies once we have specified a causal model, even though there is no algorithm for building causal models which does not itself appeal to causal judgments.

If this non-reductive approach to explicating causation is worthwhile in the case of ordinary causation, then it ought also to be worthwhile in the case of metaphysical causation. By applying the interventionist analysis to grounding construed as metaphysical causation, we might accordingly hope to derive some interesting and informative results about the relation of different grounding claims to one another. That will be my approach in the next section.

5. Metaphysical causal models

From an interventionist perspective, the counterfactual dependency judgments underlying claims of metaphysical causation will be underwritten (perhaps explicitly or unconsciously) by a particular *metaphysical causal model*. Such models may initially seem unfamiliar, so it will help to look at a range of examples. This section describes causal models for five disputed sorts of case from the causation literature, offers a selection of metaphysical causal models with the same structure, and discusses some interpretive problems which arise. The models presented will help us in at least three ways.

Firstly, the models will reflect a wide range of different applications of metaphysical causation. This will provide further illustrations of the flexibility of an approach to grounding involving G=MC.

[15] Reutlinger (2012) argues (his 'first argument') that the notion of an intervention can be dispensed with to yield a bare counterfactual theory which yields truth-conditions for causal claims equivalent to those yielded by Woodward's theory. I find this contention plausible, but it will only dispense with the causal ideology employed in conjuncts 1 and 2 of Woodward's definition. The bare counterfactual theory that Reutlinger recommends will still fail to be reductive as it will need to build in the explicitly causal constraints on right-tracking counterfactuals imposed by Woodward's conjuncts 3 and 4.

Secondly, the models correspond to metaphysical versions of a number of well-known puzzle cases from the causation literature. They include cases (omission, prevention, and double prevention) which motivate Hall's distinction between causal production and causal dependence, as well as cases (symmetric overdetermination and pre-emption) which have often been raised as counterexamples to simple counterfactual theories of causation but which are correctly handled by the interventionist approach. These cases illustrate that the same theoretical pressures arise both in the case of grounding and in the case of ordinary causation, further reinforcing the causation-grounding analogy.

Thirdly, these models and the interventionist counterfactuals they encode provide us with concrete examples which will be exploited later on in this section to explicate the concept of a metaphysical intervention, and in section 7 to frame the arguments concerning counterpossible dependence.

For each type of causal structure discussed below, I provide an ordinary causal model and two metaphysical causal models with the same structure. (I also provide one 'mixed' causal model combining the two types of causal connection.) Each model is represented visually by a 'neuron diagram', a popular and intuitive way of visualizing the dependency structure of causal models. Neurons are nodes representing particular facts - in the ordinary causation literature, this typically means facts concerning the occurrence or non-occurrence of particular events. Shaded neurons represent occurring events. Unshaded neurons represent non-occurring events: absences or omissions.[16] Lines between neurons represent 'stimulatory connections' between individual neurons; lines with circles at their ends represent 'inhibitory connections'.

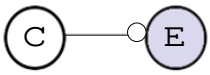
Neuron diagrams only represent some of the structure of a causal model, and have a number of limitations; one of these limitations is that they cannot represent variables which can take more than two possible values. This will not be a problem here, since I am specifying that each variable corresponds to the truth or falsity of a single fact, and am assuming a bivalent logic of facts. In addition, neuron diagrams cannot (without extension) represent cases in which variables depend on joint settings of multiple other variables. For example, they cannot represent cases where a neuron only fires if exactly one of two upstream neurons fire, though causal models exist for such scenarios. Fortunately, my

[16] The presence/absence distinction is clearest in the case of event causation, but in some other cases we can still make clear distinctions between positive and negative facts.

argument does not appeal to any causal models of that type. For further discussion of the uses and limits of neuron diagrams, see Hall & Paul (2013).

Our first kind of case, causation by omission, involves a dependence of the effect (here, the fact that the plant dies) on some other fact's not obtaining (here, on my failure to water the plant.) The plant dies *because* I don't water it.

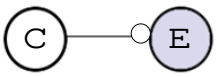
Causation by Omission



- C: I water the plant.
- E: The plant dies.

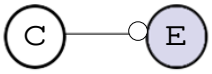
Such cases are prominent in the literature: in section 3, we saw Hall appeal to them in defending his distinction between dependence and production. And examples of grounding with the same structure are easy to find. Here are two: it is the case that all sets are pure *because* it is not the case that concreta exist, and it is the case that 'P' is false *because* it is not the case that P.

Grounding by Omission: Concreta



- C: Concreta exist.
- E: All sets are pure.

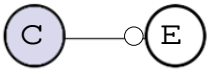
Grounding by Omission: Negation



- C: P
- E: 'P' is false.

The inverse of causation by omission is causation by prevention. Here the negative fact plays the role of the effect rather than the cause. Cases of prevention are just as serious a difficulty for production-style views of causation as cases of omission. A typical example is the plant which fails to flower because its soil dries out:

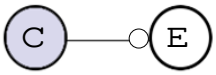
Causation by Prevention



- C: The soil dries out.
- E: The plant flowers.

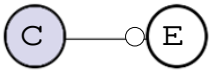
It is just as easy to generate prevention cases in the grounding context as in the context of ordinary causation. In the following two examples, the truth of the axioms of set theory prevents there from being a set of all sets, and the tightness of the binding of the electrons in a plastic rod prevents the rod from being conductive.

Grounding by Prevention: ZFC



- C: The axioms of ZFC are true.
- E: There is a set of all sets.

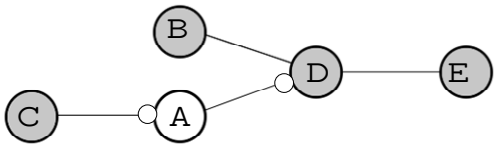
Grounding by Prevention: Conductivity



- C: The electrons in the plastic rod are tightly bound to their atoms.
- E: The plastic rod is conductive.

Problems for production-style theories of causation are made most most vivid in cases of double prevention. In an example from Hall (2004), a fighter plane escorting a friendly bomber shoots down an enemy who threatened the bomber. The fighter's shooting at the enemy is the cause of the eventual successful bombing trip:

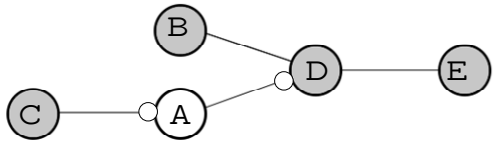
Causation by double prevention



- C: Escort shoots at enemy.
- E: Bomber bombs target
- A: Enemy shoots at Bomber.
- B: Bomber takes off from home base.
- D: Bomber arrives at target.

Double prevention cases can be constructed with relations of grounding simply by chaining together cases of grounding by prevention. In the first example, the switch being set to on causes the circuit to be live by preventing the switch it from being an insulator, which would have prevented the circuit from being closed.

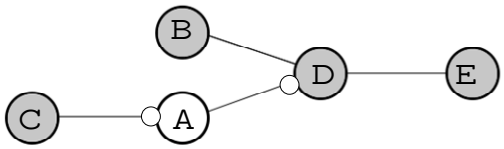
Grounding by double prevention: Circuit



- C: The switch is set to 'on'.
- E: The circuit is live.
- A: The switch is an insulator.
- B: The switch is wired to the battery.
- D: The circuit is closed.

For the second example, we assume a possible-worlds account of modality. Physicalism is true because no possible world contains zombies: the zombieless of all worlds prevents zombies from being metaphysically possible, which would have prevented the mental from supervening on the physical, which would would have rendered physicalism false:

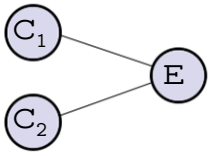
Grounding by double prevention: Physicalism



- C: No possible world contains zombies.
- E: Physicalism is correct.
- A: Zombies are metaphysically possible.
- B: Minds are constituted by matter.
- D: The mental supervenes on the physical.

Causation by omission and causation by prevention have seemed mysterious primarily because of the peculiar metaphysical status of omissions. In contrast, a different class of cases is interesting because the cases undermine simple versions of counterfactual analyses of causation. Under the heading of overdetermination cases, we can distinguish symmetrically overdetermined causation (where both causes, intuitively, 'take effect') from pre-emption (where one potential cause is prevented from taking effect by the action of other). A familiar and gruesome example of symmetric overdetermination is the firing squad:

Overdetermined causation

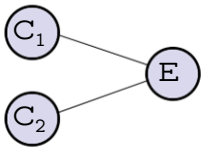


- C_1: A shoots the prisoner in the head.
- C_2: B shoots the prisoner in the head.

- E: The prisoner dies.

Symmetrically overdetermined grounding is commonplace. The presence of arsenic and the presence of strychnine are each sufficient for a potion to be poisonous, and the truth of P and the truth of Q are each sufficient for the truth of their disjunction.

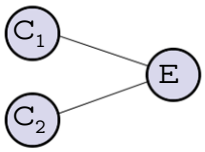
Overdetermined grounding: Poison



- C_1: The potion contains 1g arsenic.
- C_2: The potion contains 1g strychnine.

- E: The potion is poisonous.

Overdetermined grounding: Disjunction

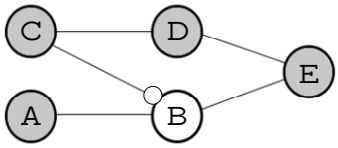


- C_1: P
- C_2: Q

- E: P $\vee$ Q

In early pre-emption cases, one of the potential causes of an effect is prevented by a fact which triggers a causal chain leading to the effect via a different route. Here, Kangaroo is prevented from eating a tasty-looking bush because Wombat eats it first:

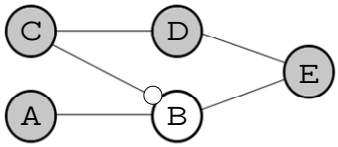
Causal early pre-emption



- C: Wombat bites into the plant.
- E: The plant is digested.
- A: Kangaroo sees the plant.
- B: Kangaroo eats the plant.
- D: Wombat swallows the plant.

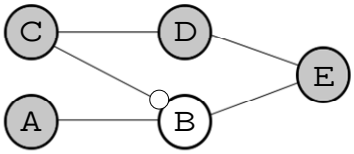
Cases of early grounding pre-emption tend to involve one principle trumping another. Here are two examples from different contexts: in the first, the presence and arrangement of all of my particles trumps the presence and arrangement of a subset of them when it comes to constituting a person, and in the second the circumstances according to which a cricket delivery counts as a 'No ball' trump the circumstances according to which it counts as a 'Wide'.

Grounding early pre-emption: Constitution



- C: My particles are arranged me-wise here.
- E: There is a person here.
- A: A subset of my particles are arranged me-without-a-little-finger-wise.
- B: There is a person with exactly 9 fingers here.
- D: There is a person with exactly 10 fingers here.

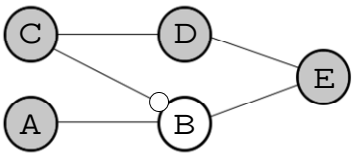
Grounding early pre-emption: Cricket



- C: The bowler over-steps the crease.
- E: The batsman's team is awarded one run.
- A: The ball passes 2m wide of the stumps.
- B: The ball should be called a Wide.
- D: The ball should be called a No ball.

One final kind of case deserves to be mentioned. The causal models described above all involve either ordinary causation or metaphysical causation; but we can also consider the results of mixing together the two sorts of causal link in a single model. Such models seem quite unproblematic; in the following example, the ordinary causal history of the cricket ball and the rules of cricket conspire to cause the fact that the batsman's team is all out.

Mixed early pre-emption



- C: The batsman catches the ball.
- E: The batsman's team is all out.
- A: The ball approaches the wicket.
- B: The ball strikes the wicket.
- D: The batsman should be given out Handled the ball.

How should we understand these metaphysical causal models? In particular, how can we make sense of an intervention on a variable like D in the Physicalism example, which has its value of metaphysical necessity? Intervening on non-contingent facts is metaphysically impossible. And how can we make sense of an intervention on a variable like D in the Constitution example, which has its value metaphysically necessitated by upstream variable C?

It has been well known at least since Kim (1973) that cases where facts are not fully distinct from one another (such as my existence, and the existence of my head) give rise to the sorts of counterfactual dependence which counterfactual theories of causation tend to count as sufficient for causal dependence. In order to rule out such cases, thinking of them as characteristically non-causal, interventionists have tended to apply a constraint such as the following to the variables within a single causal model.[17]

> *Independent Manipulability:* It is metaphysically possible that every proper subset of the variables in [a causal model] be set to every combination of their possible values by independent interventions
>
> Weslake (MS)

Independent Manipulability is not compatible with an interventionist treatment of metaphysical causation. Metaphysical causes often metaphysically necessitate their effects, and then the variables in the metaphysical causal models given above cannot be independently manipulated. Some combinations of their values are metaphysically impossible, and no metaphysically possible manipulations can give rise to a metaphysically impossible state of affairs.

At this point I want to suggest a modification of the interventionist picture which equips it to handle metaphysical causation, I suggest that's all that's required for a fact to be a possible 'metaphysical intervention' is that we be in a position to set some value of the variable *in thought* while holding fixed the values of other variables unchanged. On this proposal, it doesn't matter whether these joint settings of the variables are physically, metaphysically, or even logically possible, as long as they are thinkable. For example, we can't intervene to stop the mental from supervening on the physical, but we can (apparently) coherently suppose that supervenience fails; and although we cannot intervene to prevent a person from existing while holding fixed the arrangement of their particles, we can at least entertain the nihilist hypothesis that the particles be so-arranged without composing a person.

---

[17] Weslake (MS) and Woodward (MS) both use a constraint of this sort as part of their interventionist solutions to the causal exclusion problem. Although I cannot tackle the exclusion problem here, I think that any causal overdetermination resulting from thinking of grounding as metaphysical causation is altogether unworrying. Common-sense may tell us that events are not systematically overdetermined with respect to ordinary causation, but I see no reason to think that it tells us events are not systematically causally overdetermined with respect to causation in the broader sense.

Let's take stock. The interventionist[18] account of metaphysical causation under discussion connects causal relations with patterns of true counterfactuals which collectively constitute veridical metaphysical causal models. Counterfactuals with antecedents corresponding to combinations of metaphysical interventions on model variable settings, and with consequents corresponding to conjunctions of model variable settings, suffice for relations of metaphysical causal dependence between the variables in question. Other counterfactuals - including the intuitive class of wrongtrackers - don't meet this criterion and so don't suffice for causal dependence of any kind.

A principle like Independent Manipulability may still have a subsidiary role to play in the picture I am suggesting. Such a principle could be true of the ordinary and scientific causation that interventionists have typically modelled, yet fail for metaphysical causation. These considerations may in fact offer a natural way of distinguishing between different species of the genus 'causation': types of causation can be individuated by the 'outermost' sphere of possibilities required to count interventions on the model variables as compossible. For example, biological causation could be linked to patterns of interventionist counterfactuals where the antecedent intervention sets are biologically compossible, physical causation could be linked to patterns of interventionist counterfactuals where the antecedent intervention sets are physically compossible, and so on.

In section 7, I revisit the version of interventionism sketched in this section, using it to argue from a widely-held view about the semantics of counterfactual conditionals to the falsity of a broad range of grounding claims. But before doing so, in the next section I'll complete my case for G=MC by discussing some additional aspects of the grounding-causation analogy.

6. Summarizing the Grounding-Causation Analogy

One further aspect of the grounding-causation analogy remains to be mentioned. Orthodoxy has it that both grounding and causal dependence comprise partial orderings, having the logical properties of anti-symmetry,

[18] If metaphysically impossible interventions are allowed, the resulting view might be deemed not to deserve the name 'interventionism'; if so, perhaps we might substitute 'variationism'. I will continue to use the interventionist terminology, to avoid confusion.

irreflexivity and transitivity. However, it turns out that these logical properties can be challenged for both relations, and in exactly analogous ways.

Take transitivity first. The cases which seem to threaten the transitivity of ordinary causation are cases of pre-emption where the cause triggers *and then cuts* off an alternative causal pathway to the effect. Hall gives the example (Hall 2004) of a climber, who sees a boulder rolling towards her and ducks; the boulder passes harmlessly overhead and she walks on. Plausibly, the falling rock caused her ducking, and her ducking caused her survival, but the falling rock does not cause her survival. As Jonathan Schaffer has pointed out, we can generate structurally similar cases to challenge the transitivity of grounding. Schaffer (2012) discusses a case of a dented sphere O, arguing that the dent in O grounds O having determinate shape S*, and that O having S* grounds O being near-spherical, but that the dent does not ground O being near-spherical.

Not only are there analogous challenges to the transitivity of both ordinary causation and grounding, these challenges can be met in the same sorts of ways. Schaffer proposes a contrastive treatment of grounding as a diagnosis of the transitivity failure; this mirrors exactly his contrastive treatment of causation. Interventionism provides another way of implementing a constrastive treatment, and underwrites the possibility of transitivity failures both for metaphysical causation and for ordinary causation. Of course, friends of transitivity can bite the same bullets in both cases, either by rejecting one of the intuitive causal premises or by embracing the counterintuitive causal conclusion

The anti-symmetry (and consequently the irreflexivity) of causation has also been challenged. One of Lewis' motivations for not building the temporal asymmetry of causation directly into his 1979 analysis was a desire to allow for the coherence of backwards causation, such as might occur in cases of consistent time travel. For example, consider the case of the bootstrapping time-traveller: Old Tim travels back in time and gives the blue-print for a time-machine to Young Tim, who uses it to build a time-machine and later completes the loop. In recent work, the anti-symmetry of grounding has been challenged in a closely analogous way. Naomi Thompson (Thompson MS) and Elizabeth Barnes (Barnes MS) have given a number of candidate examples of grounding loops, concluding that grounding is non-symmetric. A particularly nice example of Thompson's is the following pair of propositions, where it seems that the truth of each is grounded in the truth of the other:

***P**:* "Q is true"
***Q**:* "P is true"

Again, the same sorts of response to these challenges to anti-symmetry are available in the causation and grounding cases. Probably the most popular responses will be either to reject all the purported cases of symmetric causation and symmetric grounding, or to endorse symmetry in either case only when restricted to some specific kinds of subject-matter.

A final (rather more nebulous) point of analogy between grounding and causation concerns their methodological status. Each notion has historically attracted suspicion from philosophers of empiricist inclinations concerned about the epistemological difficulties they present: consider Hume's argument that causation cannot be perceived, Lewis' campaign to account for everything in the scientific and manifest image in terms of his doctrine of Humean Supervenience, Sider's confession that "as a Humean I'm suspicious of metaphysical pushings and pullings" (Sider 2011 p.145) and Daly's recent arguments (Daly 2012) that the notion of grounding is 'unintelligible' or 'obscure' because it cannot be characterized in independent terms (Daly 2012). Relations of cause and ground are thought to lack clear content just insofar as they go beyond the uncontroversial notions (constant conjunction, supervenience) that they are supposed to explain. Here is not the place to properly evaluate this line of thought (although section 7 explores one way in which it could be developed); it will suffice for present purposes to note that it exists.

We are now in a position to draw the various strings of the grounding-causation analogy together and to sum up the case for G=MC. Grounding and causation are alike in the following respects:

- Both (strict partial) grounding and causation are ordinarily thought to form a partial order. (Section 2.)
- Both grounding and causation can be informatively cited in explanations, although neither relation is itself a relation of explaining. (Section 2.)
- Both grounding and causation are closely associated with distinctive patterns of one-way counterfactual dependence. (Section 4.)
- The projects of reducing each notion to counterfactuals face structurally similar problems with wrong-tracking counterfactuals. (Section 4.)

- Analogous puzzle cases can be constructed for counterfactual analyses of each notion. (Section 5.)
- A generalized interventionist approach can be applied to both notions, providing in each case an account that is non-reductive but potentially still informative. (Section 5.)
- Transitivity and symmetry can be challenged for each notion, through appeal to analogous kinds of cases, and structurally similar responses are available to these challenges. (Section 6.)
- Both notions seem 'spooky', the sort of thing that an austere Humean should not want in their picture of the fundamental world. (Section 6.)

My case for G=MC rests upon this systematic analogy, and upon the benefits of G=MC (described in section 1) with respect to ideological parsimony and to the grounding-explanation connection.

That's all I have to say in support of my central thesis that grounding is metaphysical causation. In the next section, I'll use G=MC to diagnose the prevalence of scepticism about grounding.

7. Counterpossible Dependence

In this section, I'll argue from G=MC and the view that counterpossible counterfactuals are vacuously true to the conclusion that grounding cannot play the central role in metaphysics which it has recently been assigned. Since I take counterpossible vacuity to be a very popular view, this argument provides a potential diagnosis of the resistance that grounding has faced in many quarters.

The problem is that grounding relations typically give rise to *counterpossible dependence*. According to G=MC, wherever we have a case of grounding we have a case of metaphysical causal dependence. Associated with causal dependence, via a causal modelling approach, are interventionist counterfactuals of the sort that were identified in section 5. What is characteristic of counterpossible dependence is that some of the interventionist counterfactuals involved are counterpossibles: they have metaphysically impossible antecedents.

At first glance, it may seem as though only some of the examples of grounding that we've been working with will involve counterpossible dependence. The CF counterfactuals of section 4 seem all to have

metaphysically contingent antecedents, with the possible exceptions of CF-Euthyphro (on some conceptions of God) and of CF-Noether (on some conceptions of laws of natures). In the causal models of section 5, we find a few more examples of counterpossible dependence:

- If the axioms of ZFC had not been true, then there would have been a set of all sets.
- If some possible world had contained zombies, then zombies would have been metaphysically possible.
- If it hadn't been the case that ¬P, then "P" would not have been false. *(For any P which is metaphysically impossible.)*

Still, though, it looks like all the true interventionist counterfactuals associated with several of the causal models will have metaphysically possible antecedents.

However, to stop here would be to neglect that causal dependence, in order to be (by-and-large) anti-symmetric, requires the *failure to hold* of further dependence counterfactuals. The causal models of section 5, given an interventionist reading, require the *falsity* of the following counterfactuals:

- If an intervention had been made *to prevent all sets from being pure while holding fixed the non-existence of concreta*, then there would have been concreta.

- If an intervention had been made *to prevent the water from having salinity zero, while holding fixed that the water contained no sodium ions*, then the water would have contained sodium ions.

- If an intervention had been made *to prevent "P" from being false, while holding fixed that ¬P*, then it would not have been the case that ¬P.

These counterfactuals are false, according to the causal models in question, since the variable settings described in the consequent are not obtained by applying the intervention described in the antecedent to the relevant causal model. The important point, for our purposes, is that the relevant interventions (italicized) are impossible ones and so these interventionist counterfactuals are counterpossibles.

The analogues of these interventionist counterfactuals for ordinary causation are counterfactuals such as "if an intervention had occurred to keep the plant alive, holding fixed that I did not water it, then I would have watered it". Interventionists have to rely on the falsity of these counterfactuals in order to obtain correct verdicts about the lack of causal dependence of my watering on the plant's survival. No worries, when ordinary causation is concerned: there are plenty of possible ways for the plant to stay alive even if I fail to water it. (Perhaps someone else waters it, or there's a fortuitous leak in the roof.) But when metaphysical causation is concerned, these counterfactuals will in general be counterpossible counterfactuals.

Counterpossible counterfactuals pose a difficult philosophical puzzle. Familiar semantic accounts of counterfactuals in terms of possible worlds break down when applied to counterpossible counterfactuals, for obvious reasons; and it's not obvious how to extend the treatment to a space of impossible worlds. In the light of such problems, a popular and strikingly simple response has been to declare all counterpossible conditionals trivially true. Our differential responses to counterpossibles can then be explained away on pragmatic grounds. Call this the *conservative* approach.

David Lewis was a conservative: he described himself as "fairly content to call counterfactuals with impossible antecedents be vacuously true" (Lewis 1973), noting that this approach is enforced (at least for inconsistent antecedents) by the combination of *ex falso quodlibet* and the attractive thesis that counterfactuals where the antecedent logically implies the consequent are automatically true. Stalnaker (1996) adopts a similar position, for similar reasons. Conservatism has also recently been fiercely defended by Timothy Williamson, who writes:

> The logic of quantifiers was confused and retarded for centuries by unwillingness to recognize vacuously true universal generalizations; we should not allow the logic of counterfactuals to be similarly confused by unwillingness to recognize vacuously true counterpossibles.
>
> Williamson (2008) p.175

From the conservative perspective, according to which all counterpossibles are trivially true, the interventionist counterfactuals associated with counterpossible dependence are all trivialized. If an intervention were to prevent

there from being any sets, while holding fixed the existence of Socrates, there would still be Socrates, right enough; but it's also true on this picture that, if an intervention were to prevent there from being any sets, while holding fixed the existence of Socrates, then Socrates would *not* exist. Conservatism about counterpossible counterfactuals undermines the differences in truth-value between interventionist counterfactuals which are essential to allow for non-trivial metaphysical causal models.

Take a step back for a minute, and revisit the familiar claim that purely modal analyses of the grounding relation are destined to fail. It is part of the contemporary folklore that grounding goes beyond a merely modal connection like one-way supervenience (Bennett & McLaughlin 2005). The classic examples which underwrite this folklore (many of which derive from the work of Kit Fine) are cases like Singleton. Singleton Socrates necessarily exists iff Socrates does; so no two worlds can differ with respect to whether Singleton Socrates exists without differing with respect to whether Socrates exists, and vice versa. So there is *two-way* supervenience between the existence of Socrates and the existence of Singleton Socrates. If the latter is grounded in the former, as intuition seems to tell us, then grounding is not one-way supervenience.

G=MC provides a way to revive the spirit, if not the letter, of modal analyses of grounding: instead of analyzing grounding, as the one-way supervenience analysis does, in terms of necessitated material conditions, we can analyze it in terms of subjunctive conditionals and use causal models to encode asymmetric patterns of counterfactual dependence. The trick is to use a theory of counterfactuals which allows for non-trivial counterpossible truth and falsity, and which can accordingly underwrite the needed variation in truth-value of the interventionist counterfactuals that are involved in metaphysical causal models.

I will use the term 'liberal' to cover those philosophers, such as Priest, Nolan, Fine, Goodman, and Brogaard & Salerno, who distinguish counterpossibles into true and false rather than merely into assertible and unassertible. Several advocates of this program (Priest 2005, Nolan 1997, Goodman 2004, Jago forthcoming) have developed a framework of *sui generis* impossible worlds to underwrite a familiar closeness-based semantics for assessing counterpossibles, while Restall 1997 proposes instead to to reduce impossible worlds to sets of possible worlds.

Accepting non-trivial counterpossibles opens the way for counterfactual-based treatments of the difficult cases - such as Singleton - which sank the one-way supervenience analysis of grounding[19]. Although grounding cannot be captured via necessitated strict conditionals, as in the supervenience approach, it can be captured in terms of interventionist counterfactuals instead. As I suggested above, this retains the spirit of the supervenience analysis[20]: the ideological resources appealed to are just those of our ordinary counterfactual thinking, so long as it is allowed to range beyond the limits of the possible.

At this point we come to a parting of the ways for advocates of G=MC. Consider the following *reductio* argument (similar arguments could be developed involving any one of the causal models of section 5):

1. G=MC is correct. (Premise.)
2. Interventionism is correct. (Premise)
3. The existence of Socrates grounds the existence of Singleton Socrates. (Premise.)
4. If G=MC and Interventionism are both correct, then if A grounds B then an intervention on B would alter the truth-value of A, but not vice versa. (Definitions of Interventionism, G=MC.)
5. It is false that if an intervention had been made to prevent Singleton Socrates from existing, while holding fixed the existence of Socrates, then Socrates would not have existed. (From 1, 2, 3, 4.)
6. "If an intervention had been made to prevent Singleton Socrates from existing, while holding fixed the existence of Socrates, then Socrates would not have existed" is a counterpossible counterfactual. (Premise.)
7. Not all counterpossible counterfactuals are trivially true. (From 5, 6.)
8. All counterpossible counterfactuals are trivially true. (Premise.)
9. *Reductio*. (From 7, 8.)

[19] After writing a first draft of this paper, I discovered that Krakauer (2012) develops an analysis of grounding which, like mine, makes use of counterpossible conditionals. However, Krakauer rejects G-MC and develops his analysis in a rather different way. A comparison of our approaches will have to await another occasion.

[20] It is interesting to compare my revival of the modal analysis of grounding in terms of counterfactuals with the revival of the modal account of essence by Brogaard & Salerno (2007). They similarly rely on non-trivial counterpossible counterfactuals to distinguish essential properties an object from properties merely necessitated by the object's existence. In light of the close connection between grounding and essence, counterfactual accounts of grounding and essence are natural companions.

If we want to combine G=MC with an interventionist approach to causation, then we cannot simultaneously uphold 3 and 8. Which should be given up?

3 is usually treated as basic data in the context of the grounding debate, but is much less widely accepted outside the circle of grounding enthusiasts (perhaps because grounding is not generally felt to be properly understood). 8 is rejected in many circles (especially amonst paraconsistent logicians and inflationary metaphysicians), but it is strongly endorsed by those philosophers (such as Lewis, Stalnaker, and Williamson) who prize straightforward and elegant connections between counterfactuals and metaphysical modality. In the light of this sociological divide, I predict that attitudes to the anti-triviality argument will largely divide along party lines.

Conservatives will be led by their denial of counterpossible triviality to deny that putative cases of grounding involve the characteristic patterns of counterfactual truth and falsity by which interventionists characterize causation. So conservative advocates of G=MC will be led by an interventionist picture to reject grounding as a useful notion in metaphysics. In contrast, liberal advocates of G=MC will be still able to countenance widespread metaphysical causation on an interventionist model. They can allow for non-trivial truth and falsity among the interventionist counterfactuals associated with all of our various candidate examples, vindicating grounding relations in these cases.

My own sympathies lie with the conservatives, although I won't defend that position here. Instead, I'll conclude this section by adducing a final piece of (circumstantial) evidence for G=MC. In my experience, philosophers do seem to cleave in relatively orderly fashion along the conservative-liberal distinction just outlined. Authors who are happy with talk of grounding also tend to be happy with non-trivial counterpossible counterfactuals (Kit Fine, Graham Priest and Daniel Nolan are paradigm examples), while others (amongst them David Lewis and Timothy Williamson) sternly reject both sorts of ideology. G=MC explains this sociological division: on the assumption that counterfactual dependence is sufficient for causation, the distinctive aspect of grounding is that it involves the counterpossible dependence which conservatives reject.

8. Conclusion

It is time to sum up. I have argued for G=MC on the basis of its ideological parsimony and its explanatory virtues, and on the basis of the close analogy between grounding and causation which has been charted over the course of this paper. G=MC makes sense of how we understand and assess grounding claims, and of the role we put them to in metaphysical theorizing. When combined with an interventionist approach to causation and with a semantics for counterfactuals which allows for non-trivial counterpossible truth and falsity, G=MC delivers sensible verdicts over a wide variety of cases.

G=MC also casts into sharp relief a divide that runs through contemporary metaphysics, between conservatives who reject counterpossible dependence and liberals who endorse it. Recognizing this divide provides us with a new handle on recent controversies over grounding. If G=MC is on the right lines, then the legitimacy of grounding talk stands or falls with the coherence of non-trival counterpossible truth and falsity.

References

Barnes, E. MS. "Symmetric Dependence".

Bennett, K. MS. *Putting Things Together*.

Bennett, K. & McLaughlin, B. 2005. "Supervenience". *Stanford Encyclopaedia of Philosophy*.

Briggs, R. 2012. "Interventionist Counterfactuals". *Philosophical Studies* 160(1): 139-166.

Brogaard, B. & Salerno, J. 2007. "Why Counterpossibles are Non-Trivial", *The Reasoner* 1(1): 5-6.

Cartwright, N. 2004. "Causation: One Word, Many Things", *Philosophy of Science*, 71: 805-819.

Daly, C. 2012. "Scepticism about grounding", in *Metaphysical Grounding: Understanding the Structure of Reality*, eds. F. Correia and B. Schnieder. Cambridge University Press

Dowe, P. 1992. "Wesley Salmon's Process Theory of Causality and the Conserved Quantity Theory", *Philosophy of Science* 59(2): 195-216.

Elga, A. 2001. "Statistical mechanics and the asymmetry of counterfactual dependence", *Philosophy of Science* (suppl. vol. 68, PSA 2000): S313-S324.

Fair, D. 1979. "Causation and the flow of energy", *Erkenntnis* 14(3): 219-250.

Fine, K. 1994. "Essence and Modality". *Philosophical Perspectives*, 8: 1–16.

Fine, K. 2001. "The Question of Realism", *Philosopher's Imprint*, 1(1): p.1-30.

Fine, K. 2012. "Guide to Ground", in *Metaphysical Grounding: Understanding the Structure of Reality*, eds. F. Correia and B. Schnieder, 122-38: Cambridge University Press.

Godfrey-Smith, P. 2010. “Causal Pluralism”, in H. Beebee, C. Hitchcock, and P. Menzies, (eds.), *Oxford Handbook of Causation*. Oxford University Press.

Handfield, T., Korb, K., Oppy, G. & Twardy, C. 2008. “The Metaphysics of Causal Models: Where's the Biff?”. *Erkenntnis* 68: 149-68.

Hall, N. 2004. “Two concepts of causation”, in *Causation and Counterfactuals*, eds. J. Collins, N. Hall & L.A. Paul. MIT Press.

Hall, N. & Paul, L.A. 2013. *Causation: A User's Guide*. Oxford University Press.

Hofweber, T. 2009. “Ambitious, Yet Modest, Metaphysics”, in *Metametaphysics*, eds. D. Chalmers, D. Manley, & R. Wasserman. Oxford University Press.

Hitchcock, C. 2001. “The Intransitivity of Causation Revealed in Equations and Graphs”, *Journal of Philosophy* 98: 273-299.

Hume, D. 1748. *An Enquiry Concerning Human Understanding*.

Jago, M. forthcoming. “Impossible Worlds”, to appear in *Nous*.

Krakauer, B. L. 2012. *Counterpossibles*. PhD Dissertation, University of Massachusetts, Amherst.

Lewis, D.K. 1973a. *Counterfactuals*. Blackwell.

Lewis, D.K. 1973b/1986. “Causation”, *Journal of Philosophy* 70(17): 556-567.

Lewis, D.K. 1979/1986. “Counterfactual Dependence and Time's Arrow”, *Nous* 13(4): 455-476.

Lewis, D.K. 1986. *On the Plurality of Worlds*. Blackwell.

Lewis, D.K. 2000. “Causation as Influence”, *Journal of Philosophy*.

Liebesman, D. 2011. "Causation and the Canberra Plan", *Pacific Philosophical Quarterly* 92(2): 232-242.

Menzies. P. & List, C. 2009. "Nonreductive Physicalism and the Limits of the Exclusion Principle". *Journal of Philosophy* 106 (9): 475-502.

Nolan, D. 1997. "Impossible Worlds: A Modest Approach". *Notre Dame Journal of Formal Logic* 38.4: 535-572.

Pearl, J. 2009. *Causality*, 2nd edition. Cambridge University Press.

Priest, G. 2005. *Towards Non-Being: The Logic and Metaphysics of Intentionality*. Oxford University Press.

Psillos, S. 2009. "Causal Pluralism", in R. Vanderbeeken & B. D'Hooghe (eds.) *Worldviews, Science and Us*, World Scientific Publishers.

Ramsey. F. 1927. "Facts and Propositions", *Proceedings of the Aristotelian Society Supplementary Volume VII*.

Reichenbach, H. 1958. *The Philosophy of Space and Time*. Dover.

Restall, G. 1997. "Ways Things Can't Be", *Notre Dame Journal of Formal Logic* 38.4: 583-596.

Rosen, G. 2010. "Metaphysical Dependence: Grounding and Reduction". In B. Hale & A. Hoffmann (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford University Press.

Salmon, W. 1984. *Scientific explanation and the causal structure of the world*. Princeton University Press

Salmon, W. 1994. "Causality Without Counterfactuals". *Philosophy of Science* 61(2): 297-312.

Santorio, P. MS. "New Grounds for the Semantics of Counterfactuals".

Schaffer, J. 2012. "Grounding, Transitivity, and Contrastivity", in *Metaphysical Grounding: Understanding the Structure of Reality*, eds. F. Correia and B. Schnieder, 122-38: Cambridge University Press.

Skyrms, B. 1980. *Causal Necessity*.

Sober, E. 1985. "Two Concepts of Cause", *Philosophy of Science* (supp. vol 2, PSA 1984), pp. 405-424.

Stalnaker, R. 1996. "Impossibilities", *Philosophical Topics* 38.4: 583-596

Strevens, M. forthcoming. "Causality Reunified", *Erkenntnis*.

Thompson, N. MS. "Metaphysical Interdependence".

Weslake, B. MS. "Exclusion excluded".

Williamson, T. 2008. *The Philosophy of Philosophy*. Blackwell.

Wilson, J. MS. "No Work for a Theory of Grounding".

Woodward, J. 2003. *Making Things Happen*. Oxford University Press.

Woodward, J. MS. "Interventionism and Causal Exclusion".

Yablo, S. 2004. "Advertisement for a Sketch of an Outline of a Prototheory of Causation", in *Causation and Counterfactuals*, eds. J. Collins, N. Hall & L.A. Paul. MIT Press.