

# Grounding in the image of causation

Jonathan Schaffer

Published online: 9 January 2015  
© Springer Science+Business Media Dordrecht 2015

**Abstract** Grounding is often glossed as metaphysical causation, yet no current theory of grounding looks remotely like a plausible treatment of causation. I propose to take the analogy between grounding and causation seriously, by providing an account of grounding in the image of causation, on the template of structural equation models for causation.

**Keywords** Metaphysics · Grounding · Fundamentality · Explanation

Thus all the other things are either said of the primary substances as subjects or in them as subjects. So if the primary substances did not exist it would be impossible for any of the other things to exist. (Aristotle 1984a, p. 5)

[P]hilosophy is often interested in questions of explanation—of what accounts for what—and it is largely through the employment of the notion of ontological ground that such questions are to be pursued. Ground, if you like, stands to philosophy as cause stands to science. (Fine 2012, p. 40)

Koko the gorilla was raised to understand spoken English and to communicate by sign. She is able to understand thousands of words and to produce thousands of signs, sometimes in novel and creative patterns. Koko is real. She's not a fictional gorilla or some other mythic beast. Indeed she was born in 1971, lives in Woodside, California, and stars in various *YouTube* videos.

Koko is merely non-fundamental. She is a gorilla, and like all biological organisms, she is presumably dependent on the physical. Perhaps the fundamental entities are spacetime with certain physical fields. However exactly that story goes,

---

J. Schaffer (✉)  
Department of Philosophy, Rutgers University, 106 Somerset Street, 5th Floor,  
New Brunswick, NJ 08901-4800, USA  
e-mail: Jonathan.schaffer@rutgers.edu

it seems clear that if and when final physics is done, Koko the gorilla won't be part of that story. Rather she must somehow be understood as grounded in spacetime with its fields, or whatever does turn up in the physical story.

In general, one can ask whether a purported entity exists. If the answer is yes (as with Koko), then one can ask how this entity exists, in the sense of asking how it is generated from what is fundamental. To answer this explanatory question, one needs to know what is fundamental (perhaps spacetime with its fields), and how such fundamental entities ground the rest. The concepts of grounding and fundamentality thus play a central role in metaphysical inquiry. One should clarify these concepts if one can.

Following Fine (2001; see also Schaffer 2009; Rosen 2010), there has been an explosion of interest in grounding and fundamentality. Yet I think that the core structure of these concepts has been misunderstood. Grounding is often glossed as “metaphysical causation” (Sider 2011, p. 145; also Schaffer 2012, p. 122; Wilson manuscript). Indeed the analogies run deep: both feel like relations of generation, both look something like partial orders, and both can back explanation. Yet no current treatment of grounding looks even remotely like a plausible treatment of causation.

I propose to take the analogy between grounding and causation seriously, by providing an account of grounding in the image of causation, on the template of *structural equation models* for causation (Pearl 2000; Spirtes et al. 2000). So I offer structural equation models for grounding. I think that this approach proves able to:

- enfold grounding within a precise formalism,
- handle the paradigm cases,
- sustain the analogy with causation,
- unify the input to explanation,
- reveal shortcomings in current treatments of grounding, and
- show grounding to be an intelligible and unified posit.

But I hope that exploring this proposal may prove interesting regardless of its ultimate fate. For even if it should fail, that would presumably reveal *some hitherto unknown difference* between grounding and causation.

## 1 The concept of grounding

### 1.1 Communicating a concept

I wish to communicate a concept. It is not a new concept. It is the concept that Plato (1961a, p. 178) brings to prominence in the *Euthyphro*, asking: “Is what is holy holy because the gods approve it, or do they approve it because it is holy?” It is the concept which returns in the metaphor of the cave in the *Republic* (1961b, p. 744), where the form of the good is declared the ground of the other forms: “[T]he objects of knowledge not only receive from the presence of the good their being known, but their very existence and essence is derived to them from it, . . .” It is the concept that

Aristotle (1984b, p. 1609)—in both the *Categories* and the *Metaphysics*—codifies as *priority in nature*: “Some things then are called prior and posterior... in respect of nature and substance, i.e. those which can be without other things, while the others cannot be without them,...” Aristotle (1984b, p. 1688) then conceives of metaphysics as the study of the substances, which are what come first in nature: “Substance is the subject of our inquiry; for the principles and causes we are seeking are those of substances. For if the universe is of the nature of a whole, substance is its first part,...”

Nor is the concept I wish to communicate a lost concept. It is the concept at work in contemporary discussions of *physicalism*. Thus Loewer (2001, p. 39) characterizes physicalism as the view that “the fundamental properties and facts are physical and everything else obtains *in virtue of* them.” It is the concept at work in contemporary discussions of *truthmaking*, as revealed in Armstrong’s (1997, p. 115) rhetorical question: “Must there not be something about the world that makes it to be the case, that serves as an ontological ground, for this truth?” And it is the concept which has now become central to discussions of the nature of *metaphysics* itself. As I (Schaffer 2009, p. 379) have suggested: “[M]etaphysics as I understand it is about what grounds what. It is about the structure of the world. It is about what is fundamental, and what derives from it.”

But the concept I wish to communicate is beset by controversy. Some claim not to grasp it at all. For instance Hofweber (2009, p. 260), in the course of worrying about the viability of metaphysics as a “legitimate philosophical discipline,” goes on (2009, p. 268) to dismiss the notion as “esoteric.” Other say the concept can be grasped but in many ways. Indeed Koslicki (forthcoming) sees little value in a single general concept, saying instead that “by treating a collection of phenomena which is in fact heterogeneous as though it were homogeneous, we have, if anything, taken a dialectical step backward.” And still others—including Fine (2012) and Rosen (2010)—do claim to have a clear grasp of the concept, but I think they have misunderstood the notion (and presumably they would say the same of me).

The concept I wish to communicate is of course *grounding*, but saying the name hardly seems to help. There is a background question lurking as to how one can communicate a concept at all. Some hope for definitions, but I think that these are virtually never available, and that almost all of our concepts are in fact grasped by us in some other way. In any case I have no definition to offer. I think that the best one can do is:

- provide a wide range of paradigm cases to illustrate diverse applications,
- offer some orienting glosses and analogies as further guidance, and
- enfold the concept within a formalism to detail the internal structure of the concept, and its external connections to other concepts.

In what remains I aim to communicate the concept of grounding in these three ways. The reader who thinks that more is needed should take this as an invitation to say what more is needed generally, for a concept to be communicated at all.

## 1.2 Paradigm cases

Perhaps the best way to communicate a concept is to provide a wide range of paradigm cases to illustrate diverse applications, and trust that the reader may naturally grasp the common idea. So I begin with examples. I should clarify from the start that these examples may be individually contentious in various respects. (If there were any uncontroversial examples there would be no controversy over the legitimacy of the concept.) But the reader need not agree with all—or even any—of the illustrative cases. She need only understand the meaning of the claims they involve and grasp their common theme.

So to begin with, the classic illustration of grounding—which Aristotle uses in the *Categories* as his example of priority in nature—is that of the dependence of truth on being. As Aristotle (1984a, p. 22) puts it:

[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.

For instance, consider Socrates and the proposition <Socrates exists>. Intuitively, Socrates, by existing, *makes* the proposition true. There is an *asymmetry*: it is not as if the proposition, by being true, is making Socrates exist. And this backs an *explanation*: <Socrates exists> is true because Socrates exists.<sup>1</sup>

A second cluster of cases—invoked in my opening example of Koko the gorilla, and at use in contemporary discussions of physicalism (Sect. 1.1)—is that of the dependence of the higher-level on the lower-level. So consider the physical state of Socrates and his mental state. For the physicalist, the physical state *realizes* the mental state. There is an *asymmetry*: it is not as if the mental state is realizing the physical state. And this backs an *explanation*: Socrates is in this mental state because he is in that physical state.<sup>2</sup>

A third cluster of cases—emphasized in Fine (2012) and Rosen (2010)—is that of the dependence of complex truths on simpler truths. In this vein, consider the atomic

<sup>1</sup> A nice feature of this example is that the asymmetry involved looks *hyperintensional*, at least in the sense that any metaphysical possibility in which Socrates exists is a possibility in which <Socrates exists> is true, and vice versa. So it seems that merely intensional notions like supervenience cannot make sense of the direction of dependence.

<sup>2</sup> Everyone should agree that the physical and mental states both exist, but one should still want to distinguish the physicalist view that the mental depends on the physical, from the dualist view that both are independent fundamental features of nature. There is also the idealist view that the physical depends on the mental, and the neutral monist view that both are dependent aspects of something even more fundamental. The main dispute in the metaphysics of mind is about what grounds what.

propositions  $\langle p \rangle$  and  $\langle q \rangle$  alongside the molecular proposition  $\langle p \ \& \ q \rangle$ . As built into the standard semantic clause for ‘&’, the valuations of  $\langle p \rangle$  and of  $\langle q \rangle$  *determine* the valuation of  $\langle p \ \& \ q \rangle$ . There is an *asymmetry*: it is not as if the valuation of  $\langle p \ \& \ q \rangle$  is determining the valuations of  $\langle p \rangle$  and of  $\langle q \rangle$ . And this backs an *explanation*:  $\langle p \ \& \ q \rangle$  is true because  $\langle p \rangle$  is true and  $\langle q \rangle$  is true.

A fourth style of example—tracing back to Aristotle’s (1984a) contention in the *Categories* that qualities depend on the substances they qualify, rearing in Descartes’s (1985, p. 196) view of properties as modes of substances, and resurfacing in Armstrong’s early views on the primacy of things—is that of the dependence of modes on the things they modify. Thus Armstrong distinguishes the qualified thing (“thick particular”) from its substratum (“thin particular”, the “this”) and its modes (“such”), and (1975, p. 148) writes:

There are particulars and there are properties—but they are nothing apart from each other. What is capable of independent existence, and so is the true substance of the world, are particulars-having-certain-properties: this-suches, as Aristotle would have said.

For instance, consider the apple and its red and round aspects. Intuitively, the apple, understood as a qualified thick particular (a this-such), *supports* its red aspect and its round aspect as dependent abstractions. There is an *asymmetry*: it is not as if the red aspect and the round aspect are supporting the apple, at least on this picture. And this backs an *explanation*: the red aspect of the apple exists and the round aspect of the apple exists because of how the apple is. (Those who don’t like this view may still like something analogous for complex modes. For instance, one might think of hue, saturation, and brightness as dependent abstractions of color.)

A fifth example—which Fine uses as one of his main examples of ontological dependence—is that of the dependence of non-empty sets on their members. As Fine (1995, p. 271) puts it:

Consider Socrates and the set whose sole member is Socrates. Then, necessarily, if Socrates exists so does the set. But we do not want to say, on that account, that Socrates depends upon the set, that what Socrates is depends upon what the set is. Indeed, we are inclined to say the opposite, that the set depends upon its member.

Given the iterative conception of the non-empty set as formed from its members,  $\{\text{Socrates}\}$  is *formed* from Socrates. There is an *asymmetry*: it is not as if Socrates is formed from  $\{\text{Socrates}\}$ . And this backs an *explanation*:  $\{\text{Socrates}\}$  exists because Socrates exists.

I am suggesting that these examples share a common genus. They are all species of grounding. This is not to deny that there are many differences between the examples, but only to say that there is a seeming commonality as well concerning such factors as a kind of “causative” feel of making or shaping or generating, an intuitive asymmetry in direction, and an explanatory claim in the background.

Here are some further illustrations (cf. Correia 2005, pp. 49–50), tabulated without further commentary:

Grounds	Grounded
Natural features of the situation	Normative features
Usage in the community	Meaning features
Mathematical structure (given structuralism)	Numbers
Regularities in nature (given Humeanism)	Causal relations
Categorical features (given quidditism)	Dispositional features
Determinate features of a thing	Determinable features
Material host	Hole
Integrated whole (given monism)	Fragmentary parts

I remind the reader that she need not agree with the illustrative cases, but only understand the claims involved. For instance, she need not agree with Bradley's (1978, p. 521; also Schaffer 2009) monistic claim that "everything less than the Universe is an abstraction from the whole." She may prefer Russell's (2003, p. 92) pluralistic claim that "the existence of the complex depends on the existence of the simple, and not vice versa." For present purposes I only ask that she understand the concept at use in the debate.

### 1.3 The analogy with causation: content and internal structure

Having provided a range of paradigm cases, I turn to offering some orienting glosses and analogies as further guidance. The obvious analogy—already emerging in Sect. 1.2—is with causation. Indeed the idea that grounding is something like "metaphysical causation" is expressed by many theorists including Sider (2011, p. 145), Schaffer (2012, p. 122), and Wilson (manuscript). In this vein, Wilson (manuscript)—whose views are very similar to my own—goes so far as to say that "there is a systematic and suggestive analogy between grounding and causation" and recommends a view on which "grounding *just is* metaphysical causation." Bennett (2011a, pp. 93–4) even hints at the converse idea that causation just is a kind of building relation, with "later stages of the world [viewed] as being built from earlier ones."<sup>3</sup>

The analogy with causation encompasses content, internal structure, and also external connections to surrounding concepts. At the level of content, both relations feel—in an admittedly elusive way—like relations that may be aptly described in terms of "generation" or "production". As discussed in Sect. 1.2, it is apt to use causative verbs like 'making' in glossing grounding relations. Likewise it is apt to invoke general notions like "dependence" in glossing both causal and grounding relations. I do not expect this sort of consideration to convince grounding skeptics (convincing skeptics is a hopeless task). But I do think it fair to consider this common feel as an initial sign for what turns out to be a very systematic analogy.

<sup>3</sup> Compare Maudlin (2007, p. 182) on "our initial picture of the world" as including the following idea: "The universe, as well as the smaller parts of it, is *made*: it is an ongoing enterprise, generated from a beginning and guided towards its future by physical law."

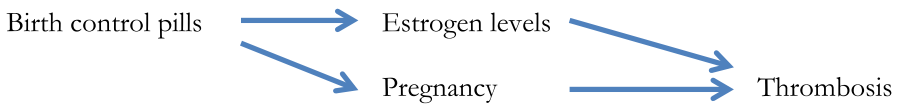
In terms of internal structure, the primary parallel is that both grounding and causation look something like partial orders. Both are standardly regimented as irreflexive, asymmetric, and transitive binary relations.<sup>4</sup> Or if one prefers the more general contrastive regimentation of both grounding and causation (Schaffer 2005, 2012), then both may be regimented as obeying the following natural contrastive counterparts to irreflexivity, asymmetry, and transitivity:

- Differential Irreflexivity:*  $\sim R \langle x, x^* \rangle, \langle x, x^* \rangle,$
- Differential Asymmetry:*  $R \langle x, x^* \rangle, \langle y, y^* \rangle \rightarrow \sim R \langle y, y^* \rangle, \langle x, x^* \rangle,$  and
- Differential Transitivity:*  $(R \langle x, x^* \rangle, \langle y, y^* \rangle \ \& \ R \langle y, y^* \rangle, \langle z, z^* \rangle) \rightarrow R \langle x, x^* \rangle, \langle z, z^* \rangle.$

If one thinks of ‘*x* rather than *x\**’ as a single unit—call this a *difference*—then these principles are just irreflexivity, asymmetry, and transitivity applied at the level of differences. Such contrastive relations can then be seen as relations of “making a difference,” and the idea of wiggling the cause/ground to wiggle the effect/grounded has a natural interpretation insofar as it is now specified what one is wiggling between.

There are a great many diverse partial ordering relations so the fact that grounding and causation are analogous in this respect is merely a further hint. But the analogies run far deeper. With both relations one can naturally distinguish application at the level of *type* and *token*. For instance, this short circuit caused that fire, and generally short circuits can cause fires. Likewise this H, H, and O arrangement grounds that drop of water, and generally H, H, and O arrangements ground water.

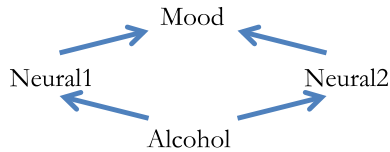
More tellingly, with both relations one naturally finds a *component* versus *net* distinction. For instance, birth control pills are a mixed cause of thrombosis: they contribute to thrombosis by increasing estrogen levels, but they also prevent thrombosis by preventing pregnancy. Overall it turns out that the net effect is preventative since the preventative component is stronger. The way to think about this is to think about birth control pills as contributing to thrombosis risk *by multiple paths* (these are the components), and to think of the net effect by summing these components:



Likewise, it is plausible—especially given neural interconnectivity—that a single chemical event in the brain such as the introduction of alcohol can (at a single fixed time) be a mixed ground of a given psychological state such as mood: the chemical might play a role in grounding multiple neural states, which might in turn play a

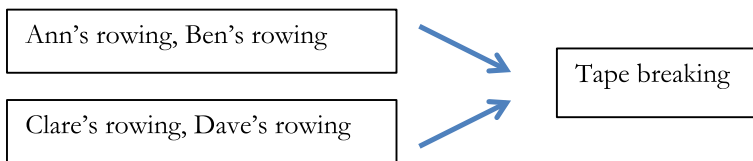
<sup>4</sup> Interestingly, there are challenges to irreflexivity, asymmetry, and transitivity for both relations, and these challenges tend to run parallel (A. Wilson manuscript). For instance, McDermott (1995), Hall (2000), and Hitchcock (2001) argue against the transitivity of causation, and Schaffer (2012) presents parallel arguments against the transitivity of grounding, while recommending a parallel contrastive resolution in both cases.

mixed role in grounding the person's psychological state. Overall it may turn out that one component is stronger, as with alcohol and mood for which the depressive component is stronger. So a similar picture in which contribution is factorized by path seems called for:

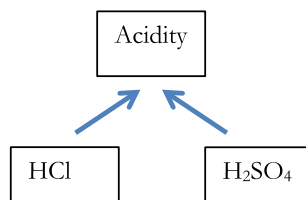


Or consider the way that  $F_a$  and  $\sim F_b$  are mixed (and equally weighed) components in the truth-value determination of 'Most(x) Fx', and imagine a case in which  $F_a$  and  $\sim F_b$  themselves have a common ground.

Moreover with both relations there is a natural tripartite distinction between *incomplete* (*/partial*), *complete* (*/full*), and *total* factors. On the causal side, if Ann and Ben jointly row a first boat across the finish line, and Clare and Dave jointly row a second boat across the finish line at the same time, so as to overdetermine the breaking of the tape at the finish line, then—boxing together complete cause packages—one can think of the causes of the breaking of the tape at the finish line in three distinct ways:



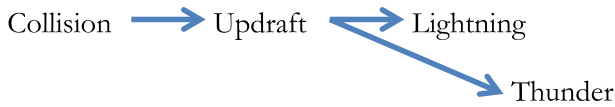
Ann's rowing and Ben's rowing together (as boxed) form a complete cause of the tape breaking, Ann's rowing (by itself) is an incomplete cause, and all the rowings together form the total cause. Likewise on the grounding side, if a molecule of sulfuric acid ( $H_2SO_4$ ) is paired with a molecule of hydrochloric acid (HCl), so as to overdetermine the acidity of the compound, then one can think of the grounds for the acidity as per:



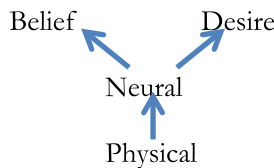
The atoms of the  $H_2SO_4$  molecule together (as boxed) form a complete ground of the acidity, the S atom is an incomplete ground, and all the atoms together form the total ground.



Finally, with both relations there is a natural idea of *screening-off*. On the causal side, suppose that the collision of the air masses causes an updraft in the sky, which in turn causes both lightning and thunder:



Given the presence of the updraft, the thunder becomes plausibly independent of the collision and the lightning. So long as the updraft is present then there will be lightning whether or not there was a collision and whether or not there will be thunder. Likewise, suppose that the physical state grounds the neural state, and that the neural state in turn grounds both a given belief and a given desire (all at a single fixed time):



Holding fixed the neural state, the belief becomes plausibly independent of the physical state and the desire. So long as the neural state is present then there is the belief whether or not there is the physical basis and whether or not there is the desire. With screening-off comes the idea of a “path analysis” of determination. Just as the collision of the air masses produces the lightning via the updraft, so the physical basis grounds the belief state via the neural state.

#### 1.4 The analogy with causation: external connections

Beyond analogies of content and internal structure, a third and final point of support for the analogy between grounding and causation comes via common external conceptual liaisons. Both notions seem similarly embedded in parallel surrounding networks of notions including law, necessity, conditionals, and—perhaps most interestingly of all—explanation (Sect. 1.2). As far as the connection to law, it is widely accepted that if there is a particular case of token causation, there is some background causal generalization—and ultimately behind it some law of nature—connecting events of the relevant types. Similarly for grounding. As Audi (2012a, p. 693; see also Rosen 2010, p. 131 on Formality) notes: “[T]he fact that my shirt is maroon grounds the fact that it is red. It is purely incidental to this relation which particular has the properties in question... Anything maroon is red, and indeed, anything maroon is red *in virtue of* being maroon.” Indeed in both cases there is some reason to think that the relevant generalizations are not merely accidental but ultimately backed by underlying *formative principles*—whether laws of nature or laws of metaphysics—which unify the patterns.

With the connection to necessity, there is a natural notion of *causal necessity*, which one can think of as necessity holding fixed the laws of nature. As van Fraassen (1989, p. 28) aptly summarizes: “Wood burns when heated, because wood must burn when heated. And it must burn because of the laws which govern the behavior of the chemical elements of which wood and the surrounding air are composed.” Analogously there is a sister notion of *metaphysical necessity*, which one can think of as necessity holding fixed the laws of metaphysics (the most general principles of grounding), which—following Rosen (2006, p. 35)—may be understood as specifying “the categories of basic constituents and the rules for their combination. They determine how non-basic entities are generated from or ‘grounded’ in the basic array.”<sup>5</sup>

With conditionals, it is widely granted that there are tight connections between causation and counterfactuals. In the one direction, counterfactual dependence is at least an excellent heuristic for token causation (Sect. 2.3). In the other direction, the laws of nature associated with causal necessity support counterfactuals. To continue van Fraassen’s example with the wood, it is true that if this sample of wood were heated, then it would burn. The same pattern of connection with conditionals surfaces with grounding. It is a sign of token grounding that Koko the gorilla’s biological state counterfactually depends on (and more generally, covaries with) her underlying physical state. And moreover, the laws of grounding associated with metaphysical necessity support counterfactuals. If Audi’s maroon shirt were navy, then it would be blue.<sup>6</sup>

Lastly—but in my view most tellingly of all—causation and grounding are among the elite handful of relations able to back explanation. One way to explain why Koko the gorilla is hungry is to tell a causal story about her meager breakfast earlier. Another way to explain why Koko the gorilla is hungry is to tell a grounding story about her neural state now. Indeed Fine (2001, pp. 615–616) characterizes ground as “an explanatory connection” and indeed as “the tightest such connection” and “the ultimate form of explanation,” saying that “if the truth that P is grounded in other truths, then they *account* for its truth; P’s being the case holds *in virtue of* the other truths’ being the case.” Audi (2012b, p. 104)—in a section entitled “the argument for grounding”—goes so far as to say: “The reason we must countenance grounding is that it is indispensable to certain important explanations.”

I have reviewed the content, internal structure, and external connection aspects of the analogy between grounding and causation, in order to make the *prima facie* case

<sup>5</sup> There are a range of more precise principles which might be adopted connecting causation/grounding to a necessity claim of the appropriate sort (Trogon 2013a). For instance, some say that it is necessary that if the causes/grounds are present, then the effect/grounded is present and caused/grounded in that way. The strongest principle I would myself endorse is a *global supervenience principle* for total causation/grounding relations: if the causal/metaphysical laws are deterministic, then there are no two causally/metaphysically compossible worlds with the same total causes/grounds but different effects/groundeds. (No difference in total causes/grounds, no difference in effects/groundeds.)

<sup>6</sup> As Wilson (manuscript) notes, there is an elegant generalization of Lewis’s notion of the counterfactual asymmetry to “right-tracking counterfactuals.” Back-tracking in time and down-tracking in levels are both ways of wrong-tracking.

for taking the analogy seriously at the level of formalism. Here are the points of analogy seen so far:

- both are generative relations;
- both are partial orders, admitting a type/token distinction, a component/net distinction, an incomplete/complete/total distinction, and screening-off relations; and
- both are backed by non-accidental generalizations, delimit a specific form of necessity, are supportive of and diagnosable by counterfactuals, and can back explanation.

This is already extremely telling. I aim to add: both are best formalized via structural equation models. Indeed, given the depth of the analogy it is actually a substantive and difficult question as to whether there is any more than a nominal distinction between grounding and causation, and I am open to (though ultimately dubious of: Sect. 4.5) the idea that there is a single unified relation of generation called ‘causation’ when it drives the world through time, and ‘grounding’ when it drives the world up levels.

## 2 Structural equation models for causation

So far I have completed two of the three tasks I take to be required for communicating a concept (Sect. 1.1): I have provided a range of paradigm cases (Sect. 1.2), and offered orienting glosses and analogies (Sects. 1.3–1.4) The sole remaining task is to wrap the concept into a formalism, so as to detail the internal structure of the concept and its external connections to other concepts. The way forward is already indicated by the depth of the analogy with causation: treat grounding in the image of causation.

### 2.1 The rise of structural equation models

To provide an account of grounding in the image of causation requires first providing a sketch of causation to serve as a template. Fortunately, there is an emerging consensus over the leading formalism for causation. I have in mind structural equation models, in the form developed by Pearl (2000) and Spirtes et al. (2000).<sup>7</sup> In this vein Hitchcock (2001, p. 273; cf. Paul and Hall 2013, p. 18) writes: “We live in exciting times. By ‘we’ I mean philosophers studying the nature of causation. The past decade or so has witnessed a flurry of philosophical activity aimed at cracking this nut, and, surprisingly, real progress has been made.”

Structural equation models of causation boast several connected advantages, of which I mention three. The first advantage comes in the relative level of precision achieved in these frameworks (Sect. 2.2). Structural equation models are sufficiently

<sup>7</sup> The history of this approach traces back at least to Wright’s (1934) path analysis, Wold’s (1964) work on linear recursive equations, and Jöreskog’s (1973) incorporation of both latent variables and factor analysis into the computer program LISREL.

precise that you can implement every last detail on a computer. Even the very best alternatives—such as Lewis’s (1986a) counterfactual dependence account—are in the end irremediably vague about crucial details. (The reader who doubts this is invited to try to program a computer to decide, for instance, whether a given case exhibits Lewisian counterfactual dependence, not to mention whether it exhibits Lewisian quasi-dependence.)

The second advantage comes in the range of causal notions covered by these frameworks (Sect. 2.3). Structural equation models permit integrated treatments of *type causation* (fires cause smoke) and *token causation* (this fire caused that smoke), *component influence* (taking birth control pills can cause thrombosis in this way) and *net influence* (overall taking birth control pills prevents thrombosis), *screening-off* relations, as well as notions of *interventions* and *counterfactuals*. No other alternative integrates anywhere near such a wide range of causal notions (much less in such a precise and elegant way).

The third advantage comes in the epistemological breakthroughs that these frameworks have enabled. It used to be a platitude in statistics that one cannot infer causation from correlation. This is now refuted. Structural equation models come with precise—and indeed freely downloadable—*discovery algorithms* that allow one, given certain plausible assumptions, to estimate causal structure from sufficiently rich correlational structure over three or more variables.

By way of illustration, imagine that you are interested in the question of why some academics publish more than others.<sup>8</sup> You suspect that intrinsic ability and quality of graduate training may play a role, but you are also interested in whether factors such as gender might play a role, and of course you are interested in quantifying how much of a role these and other factors are playing. You have the data: for over a thousand professors you know their *Stanford-Binet V* score (as proxy for intrinsic ability), *Philosophical Gourmet Report* rank of graduate program (as proxy for quality of training), gender, and number of publications. Among the questions you are interested are:

*Q1*: What if any role does gender play in publication output, and

*Q2*: What if any role did Professor Marple’s gender play in her having 27 publications to her name?

Question *Q1* concerns type causation, and so in addressing it I offer the reader the following two theoretical options:

- input your data into TETRAD (or some other causal discovery algorithm), and receive a precise and empirically reliable estimate of direction and strength of causal influence; or
- consult your intuitions, read your tarot cards, and/or hazard a guess.

These are the only options I can offer because there is no philosophical account of causation (beyond probabilistic accounts, which structural equation models

<sup>8</sup> This illustration is loosely based on Spirtes et al. (2000) discussion of Rodgers and Maranto’s (1989) work on the causes of academic productivity.

supersede) that even treats the type-level case at all. The philosophical accounts focus almost exclusively on the token-level, as if that might exhaust the topic.<sup>9</sup>

Question *Q2* concerns token causation, and so in addressing it I offer the reader the following options, and bid her the best of luck if she opts for either of the first two:

- imagine the nearest possible world in which the event of Professor Marple’s having that gender did not occur due to a “local miracle”; imaginatively introspect as to whether the event of Professor Marple’s having 27 publications would still occur;
- trace the paths of energy–momentum flow from Professor Marple’s gender to her having 27 publications; or
- use the type-level picture you got from using a causal discovery algorithm in addressing *Q1*, assign values to your variables for ability, training, and gender as befits the case of Professor Marple, and then look at what would happen to your number of publications variable if you kept ability and training fixed but re-assigned the gender variable to male instead of female.

In this example, the precision, integration of type and token notions, and epistemological breakthroughs that structural equation models offer come together.

I am *not* saying that structural equation models are any sort of magical panacea, or that all philosophical questions are thereby dissolved. Indeed several of the classical puzzles (such as preemption cases) remain puzzles, many interesting metaphysical questions remain, and the formalism itself is subject to ongoing development and revision. I am saying that as of now this approach provides the leading formalism, by a wide margin. At least as of now, there is no other serious game in town. I am saying that one may learn something from studying this formalism and understanding how it works as well as it does.

## 2.2 The formalism

There are a number of ways of characterizing structural equation models formally, but I in all essentials follow the conceptually illuminating tripartite structure presented in Halpern (2000). I also simplify the presentation by only considering models with finitely many variables, and by assuming determinism.<sup>10</sup>

<sup>9</sup> The philosophical fixation on token causation is surprising, and seems at best a historical accident. In fact Pearl (2000, ch. 10)—in the concluding chapter of his book—is perhaps the first theorist in the structural equations tradition to even consider token causation in any detail (previous theorists had focused primarily on type-level concepts). Amusingly Pearl (2000, p. 328) concludes his discussion of token causation with an acknowledgment to Don Michie “who spent many e-mail messages trying to convince me that the problem is not trivial...”

<sup>10</sup> See Halpern and Pearl (2005) for a presentation of the formalism without these restrictions. Note that the restriction to the deterministic case is reasonable insofar as one wants a template for grounding, since “indeterministic grounding” seems impossible. Grounding seems to imply supervenience: fix the grounds and one fixes the grounded. The status of the grounded thus cannot be open to chance. By way of illustration, it seems impossible that, given a fixed physical ground, the biological status of the system remains open to chance.

To begin with, one represents *the system under study*. One selects distinct variables to represent distinct features of the world. These variables are born divided into “exogenous” variables representing independent conditions, and “endogenous” variables representing dependent conditions. These variables also come born situated in a space of incompatible values representing contrasts. Thus one introduces *the signature*, understood as a triple  $S = \langle U, V, R \rangle$ .  $U$  is a finite set of exogenous variables,  $V$  is a finite set of endogenous variables, and  $R$  is a function mapping every variable  $X \in U \cup V$  to an at-least-two-membered set of allotted values. For instance, if one is studying a rock being thrown through a window, one might work with the very simple signature  $SI = \langle UI = \{Throw\}, VI = \{Shatter\}, RI \rangle$ , where  $RI$  maps *Throw* to  $\{0, 1\}$  (contrasting the rock’s being thrown with its being dropped to the ground) and maps *Shatter* to  $\{0, 1\}$  (contrasting the window’s being shattered with its remaining intact).<sup>11</sup>

Secondly, one adds in *the dynamics*. One codifies functions to say how each endogenous variable is to be evaluated, on the basis of the values of other variables. Thus one introduces *the linkage*, which is a pair  $L = \langle S, E \rangle$  where  $S$  is a signature as just characterized, and  $E$  is a set of structural equations. For every endogenous variable  $V \in V$ ,  $E$  must include an equation  $E \in E$  such that  $E$  outputs a value  $v$  to  $V$  on the basis of values allotted to certain other variables, which thereby count as  $V$ ’s *parents*.  $E$  is also subject to a global acyclicity (“no loops”) constraint: no variable can stand in the ancestral of the parenthood relation to itself. In the case of the rock being thrown through the window with the signature  $SI$  just described, a natural linkage is  $LI = \langle SI, EI \rangle$ , where  $EI$  is simply  $\{Shatter \leftarrow Throw\}$  (outputting a 0 for *Shatter* given a 0 for *Throw*, and a 1 for *Shatter* given a 1 for *Throw*).<sup>12</sup>

Thirdly and finally, one still needs to say *what actually happened*. One sets a function saying how each exogenous variable is to be evaluated. Thus one adds in *the assignment*, which is a pair  $M = \langle L, A \rangle$  where  $L$  is a linkage as just characterized, and  $A$  is the smallest function mapping every exogenous variable  $U \in U$  to exactly one value. In the case of the rock being thrown through the window, one needs to say whether or not the rock was thrown. So one just adds  $MI = \langle LI, AI \rangle$ , where  $AI$  is the smallest function mapping *Throw* to 1:  $\{\langle Throw, 1 \rangle\}$ .<sup>13</sup> So here is a simple structural equation model of a rock thrown through a window:

<sup>11</sup> Briggs (2012, p. 142) speaks of the range function as providing “answers to the question posed by the variable.” In this useful way of speaking, one can think of *Throw* as posing the *yes/no* question of whether or not the rock was thrown. It is important to appreciate that one sometimes wants non-binary and even continuous valued variables, for instance if one is posing the question of how massive the rock is, or how forceful the shattering is.

<sup>12</sup> Notational convention: I am using the schema ‘ $\Phi \leftarrow \Psi s$ ’ to notate the idea of the value of one variable (schematically: ‘ $\Phi$ ’) being determined by the values of some plurality of parent variables (schematically: ‘ $\Psi s$ ’). One sometimes sees ‘=’ used instead (‘ $\Phi = \Psi s$ ’), followed by a caveat that the determination in question is not really the symmetric relation of identity.

<sup>13</sup> I am using the schema ‘ $\{\langle \Phi_1, \varphi_1 \rangle \dots \langle \Phi_n, \varphi_n \rangle\}$ ’ to notate the assignment function in extension, as a set of ordered pairs ( $\Phi_1 - \Phi_n$  are the exogenous variables, and  $\varphi_1 - \varphi_n$  are their respective assigned values). One sometimes sees ‘=’ used here as well (‘ $\Phi_1 = \varphi_1, \dots, \Phi_n = \varphi_n$ ’).

$$\begin{aligned}
 SI &= \langle \{Throw\}, \{Shatter\}, RI \rangle, \text{ where } RI \text{ maps both } Throw \text{ and } Shatter \text{ to } \{0, 1\} \\
 LI &= \langle SI, \{Shatter \leftarrow Throw\} \rangle \\
 MI &= \langle LI, \{ \langle Throw, 1 \rangle \} \rangle
 \end{aligned}$$

Structural equation models permit a useful graph theoretic visualization, via the following recipe: draw each variable as a vertex, and for every case of parenthood draw a directed edge from parent to child. The graph for the simple model of a rock being thrown through the window is thus:<sup>14</sup>



Directed acyclic graphs are partial ordering structures, so in that sense a structural equation model induces a partial order over its variables. But structural equation models actually provide a much richer structure than the graphs alone: the model-graph relation is many-one. All models with the same cardinality of variables and pattern of parenthood relations induce the same graph. These graphs are thus helpful but impoverished visualizations, omitting such matters as the contrasts allotted to the variables, the forms of the functions for evaluating the endogenous variables, and the values assigned to the exogenous variables.

I am saying that our best current treatments of causation are embedded within the formalism just outlined, and recommending a treatment of grounding in its image. Abstracting away from the formal details, the core aspects of conceptual structure are as follows:

- independent and dependent conditions are distinguished from the start,
- each condition is situated within a space of contrasts,
- each dependent condition is associated with a function evaluating it on the basis of the values of its parents (subject to global acyclicity constraints), and
- each independent condition is assigned a value.

Structural equation models work as well as they do by integrating these core ideas. That is, these ideas guide the leading approach to understanding the internal structure and external conceptual connections of causation. Note that—despite the fact that this formalism is sometimes labeled as ‘causal modeling’—there is nothing inherently causal in either the formalism or the core ideas it integrates. Indeed I think that this formalism is excellent for modeling directed dependency relations generally, and so equally fitting for modeling both causal and metaphysical dependencies.

### 2.3 Counterfactuals and token causation

With the structural equations formalism to hand, one can do a lot of work. For instance, structural equation models permit a precise treatment of the notion of an “intervention” alongside a precise semantics for “counterfactuals,” or at least for a

<sup>14</sup> The reader familiar with Lewis style neuron diagrams (Lewis 1986a) must beware: these graphs look similar to neuron diagrams but richer and more systematic. Hitchcock (2007a, b) provides a useful comparison.

certain kind of conditional whose antecedents and consequents specify situations corresponding to conjunctions of values of the model's variables. To evaluate such counterfactuals in a given model  $M$ , one considers a modified (or "mutilated") counterpart  $M^*$  that stipulates the new values of the variables as per the antecedent. More precisely, one may consider a counterfactual of the following schematic form, assessed relative to a given model  $M$ :

If  $\Phi 1 = \phi 1$  and  $\Phi 2 = \phi 2$  and  $\Phi 3 = \phi 3 \dots$  then  $\Psi 1 = \psi 1$  and  $\Psi 2 = \psi 2$  and  $\Psi 3 = \psi 3 \dots$

(Capital letters refer to variables, and lowercase letters to their associated values. So ' $\Phi 1 = \phi 1$ ' says that schematic variable  $\Phi 1$  takes schematic value  $\phi 1$ .) To assess whether this counterfactual is true in  $M$ , first modify  $M$  into  $M^*$  via the following recipe (while doing nothing further):

- *Cut any incoming links* For all variables  $\Phi j$  in the antecedent such that  $\Phi j \in V$ , (i) delete  $\Phi j$  from  $V$  to obtain  $V^*$ , (ii) insert  $\Phi j$  into  $U$  to obtain  $U^*$ , and (iii) delete the equation in  $E$  with  $\Phi j$  on the left to obtain  $E^*$ ;
- *Re-assign the stipulated values* For all variables  $\Phi j$  in the antecedent (all of which are now in  $U^*$ ), modify the assignment  $A$  into  $A^*$  by assigning  $\Phi j$  to the value  $\phi j$  specified in the antecedent; and
- *Obtain the modified model*  $M^* = \langle L^*, A^* \rangle$ , where  $L^* = \langle S^*, E^* \rangle$ , and  $S^* = \langle U^*, V^*, R \rangle$ .

The counterfactual is true in  $M$  if and only if the consequent (' $\Psi 1 = \psi 1$  and  $\Psi 2 = \psi 2$  and  $\Psi 3 = \psi 3 \dots$ ') holds in  $M^*$ . Effectively one has modified the model in order to surgically "intervene" on the variables in the antecedent, by first converting them into independent conditions and then hand-setting their values.<sup>15</sup>

It is worth noting that this semantics for counterfactuals is obviously related to, but may still differ from, the natural language semantics for 'if ... then ...' constructions in English, and any intensional treatment in terms of "nearby possible worlds". Indeed this semantics is—in ways crucial to the extension of the system to grounding—tolerant of hyperintensional distinctions (Sect. 3.2). It is just a bit of mathematics and does not at any point appeal to intensions. (Of course if one thinks hyperintensional distinctions are in the end senseless then one will think that this bit of mathematics is tolerant of the senseless, but one will presumably also think that grounding is senseless. My point is just that this semantics *tolerates* hyperintensional distinctions, for those who wish to draw them.)

These intervention counterfactuals then allow for a useful "counterfactual dependence" test for token causation relative to a model  $M$ . Indeed, within a restricted simple range of cases, there is wide agreement that counterfactual dependence tracks token causation (as befits the many successes of Lewis's 1986a account, and as befits the *sine qua non* test used in the law). More specifically, in a given model  $M$ :

<sup>15</sup> Notice that  $M^*$  is a minimal mutilation of  $M$ : all other equations in  $E$  are left intact (beyond those for the variables explicitly set in the antecedent of the conditional). This encodes a substantive claim to the effect that causal influence is "modular" (Woodward 2003, p. 339; cf. Cartwright 2007).



*Counterfactual dependence test for causation* If  $X$  and  $Y$  are binary variables, there is a direct  $X \rightarrow Y$  path, and no other distinct path to  $Y$ , then:  $X = x$  is a token cause of  $Y = y$  if and only if  $X = x$ ,  $Y = y$ , and if  $X \neq x$  then  $Y \neq y$ .

Moreover the restriction to binary variables can naturally be lifted by introducing a second layer of contrastivity, beyond the presence of the range function in the signature, into the structure of token causation itself (see generally Hitchcock 1996; Maslen 2004; Schaffer 2005; Northcott 2008):

*Contrastive counterfactual test for causation* If there is a direct  $X \rightarrow Y$  path, and no other distinct path to  $Y$ , then:  $X = x$  rather than  $x^*$  is a token cause of  $Y = y$  rather than  $y^*$  if and only if  $X = x$ ,  $Y = y$ , and if  $X = x^*$  then  $Y = y^*$ .

When ' $X \neq x$ ' and ' $Y \neq y$ ' no longer have unique meanings, one simply stipulates which of the many alternatives to  $X = x$  and to  $Y = y$  are at issue. For instance, if one is interested in question  $Q2$  about the role of Professor Marple's gender in her having 27 publications to her name, but one is working in a model in which the gender variable is not treated as a binary male/female variable but as a ternary variable allowing a "genderqueer" option, one might find that Marple's being female rather than male impacted her productivity in one way, but that Marple's being female rather than genderqueer had quite different effects.

More complicated cases with intermediate variables between  $X$  and  $Y$ , and/or multiple pathways into  $Y$ , remain subject to controversy. Indeed there is ongoing debate as to how best to read token causal relations off structural equation models, so as to handle preemption cases and other standard problem cases.<sup>16</sup> But so long as one focuses on the simple cases, there is a straightforward and informative working test of token-level causation to be had, in terms of *counterfactual covariation*: wiggle the cause, and the effect wiggles.

In summary, with the formalism of structural equation models in hand, one can go on to do crucial work for causation including:

- offering a precise treatment of the notion of an intervention,
- writing a precise semantics for a certain kind of counterfactual, and
- formulating counterfactual tests for token causation.

<sup>16</sup> Hitchcock (2001), Woodward (2003), Halpern and Pearl (2005), and Weslake (forthcoming) offer various accounts of token causation couched within a structural equations framework. There is also the view that token causal claims *cannot* be read off structural equation models as I have presented them, but require an added gizmo that distinguishes "default" from "deviant" values for variables (Menzies 2007; Hitchcock 2007a, b; Hall 2007; Halpern 2008). I am unmoved (see Blanchard and Schaffer forthcoming), but for present purposes I would just insist on treating causation and grounding in parallel ways, and so would recommend that friends of default-relative causation also endorse default-relative grounding (some analogous arguments are available in both cases).

I aim to show that this formalism can be adapted for grounding, and that one can do comparable work.<sup>17</sup>

## 2.4 Reduction and apt representation

Before turning to grounding, I pause to discuss two of the many philosophical concerns that might arise with structural equation models for causation. The first concern is that the whole framework looks *circular*. It seems to invoke causal notions from the start, such as the initial division of variables into “independent” and “dependent” conditions and the parent–child asymmetry built into the structural equations. Where is the reductive analysis to complete the schema ‘*c* causes *e* if and only if ...’?

It should be evident that this is a *conceptually nonreductive* framework (Woodward 2003, p. 106; Halpern and Pearl 2005, p. 849; Menzies 2009, p. 357). Various causal notions are baked in from the start. After all, structural equation models were developed by computer scientists and social scientists to provide empirical answers to questions like, “Does this chemical cause cancer?” These people usually take themselves to already have a decent working understanding of the relevant concepts, and they typically have little interest in reductive analysis or other matters of metaphysical moment. They are mainly trying to develop empirical tools for inferring causation. That said, as with the best nonreductive treatments, one sees an illuminating interconnection of various notions, such as type and token causation, but also component and net influence (where component influence is influence along a path), as well as interventions and counterfactuals (where an intervention is understood via modified models, and counterfactuals are understood via interventions).

Moreover, I think that the quest for a reductive conceptual analysis was hopeless anyway. If induction is any guide, reductive conceptual analyses are hopeless for virtually any interesting philosophical concept. And if Hume is right about the centrality of causation to our conceptual scheme, it is especially unlikely that one can ever get “underneath” the concept of causation.<sup>18</sup>

What is lost with a nonreductive treatment of causation? Nothing important, as far as I can see. There is nothing *metaphysically scary* in a nonreductive approach.

<sup>17</sup> I am *only* saying that one can do work comparable to those points just bulleted. There are others uses to which the formalism gets put which may well only be apt for causation, including most prominently causal discovery. The usual discovery algorithms presuppose a background empirically given probability distribution, and build on some assumptions (the Markov condition, Minimality, and Faithfulness: see Spirtes et al. 2000, pp. 53–56) that are not just specially plausible for the causal case, but moreover only apply under indeterminism. (Deterministic systems tend to violate Faithfulness.) Though see Glymour (2007) and Baumgartner (2009) for discussions of causal discovery under determinism. For present purposes I do not try to port the discovery algorithms across to the case of grounding, and must leave the epistemology of grounding for separate discussion. (A formalism can have overall good-making features that are not usable in every application.)

<sup>18</sup> In this vein Hume (1778, p. 662) memorably concludes the *Treatise* by identifying our concepts of resemblance, contiguity, and causation as “*to us* the cement of the universe.” Similarly Carroll (1994, p. 118) comments: “With regard to our total conceptual apparatus, causation is at the center of the center.”

From the fact that the concept is not reductively analyzable, it does not follow that the causal relation in the world can float free of the occurrent pattern of events.<sup>19</sup> Nor is there anything *epistemically scary*: indeed the causal discovery algorithms associated with structural equation models furnish as precise and well-understood an epistemology as one could hope for. Nor is there anything *semantically scary*: if someone wants to know what the word ‘causation’ means, the answer is that it means causation. This is not unhelpful, since we all in fact possess this concept (contrast: for many it would be unhelpful to be told that ‘rarefaction’ means rarefaction). If someone honestly claims not to understand the concept of causation until she sees a reductive analysis, then I think she must be self-deceived.

Note that—extending the point about there being nothing metaphysically scary—a conceptually nonreductive treatment is compatible with all sorts of views about the metaphysics, from the idea that causation is a primitive and fundamental constituent of reality, to the idea that causation is a real but derivative phenomenon, through to the idea that causation is a crude concept corresponding to nothing. My preferred background picture is that conceptual reductions are almost always failures. On the conceptual side one should expect a rich tapestry of interwoven notions. But on the metaphysical side fundamental reality is sparse. I think that causation is at best an approximately real but derivative phenomenon, derivable from the laws of nature. But my present focus is on the conceptual side. I am saying that structural equation models provide a fitting picture of how causal notions are laced together and threaded through the larger tapestry.

The second concern is more interesting, and often overlooked in technical discussions of the formalism. The concern is that these models are in the end just mathematical *representations*. Whenever one indulges in representation the question arises whether the representation is apt. (This is a very general issue that arises throughout the sciences, and in any domain in which models are employed. Structural equation models are but one type of model.) So one must specify *aptness constraints* on the models. Standard aptness constraints on structural equation models include the following (Blanchard and Schaffer forthcoming: Sect. 1.3):

- the values of distinct variables should all represent distinct events,<sup>20</sup>
- the values of each individual variable should represent incompatible contrast events,
- the variables should represent sufficiently many events to capture the structure of the situation,

<sup>19</sup> *Analogy*: Williamson (2000) defends a *knowledge-first* approach to epistemology, in which the concept of knowledge is treated as primitive and irreducible. But knowledge-first epistemology is clearly compatible with the supervenience of knowledge facts on the physical facts, and no sensible person should be scared that the knowledge-first view leads to metaphysical dualism (fundamental mental states) for positing unanalyzable mental concepts.

<sup>20</sup> This corresponds to the metaphysical image of causation as a relation between distinct events (see generally Lewis 1986a). Though note that there are few metaphysical constraints on what can count as an “event” in the system. Indeed one wants to be able to have variables representing such matters as number of publications, and allowed to take values like 27 (a scattered and disjunctive affair) or even 0 (an absence). After all, having 27 publications can cause one to be promoted, while having 0 publications can cause one to be fired.

- the counterfactuals encoded in the model's equations should be true,<sup>21</sup> and
- the assignment should represent the initial conditions correctly.

(This list is not intended to be exhaustive, but just to provide a useful handful of heuristics.)

But the concern only deepens, insofar as it is very doubtful that these or any other reasonable aptness constraints will pin down a uniquely correct model for every situation. It is hard to see an objectively unique determinant of exactly which events must be included, and which contrasts must be allotted.<sup>22</sup> This is a concern because multiple apt models of a given situation may *disagree* over causation. A general issue lurks: when multiple apt representations of the world disagree, what is the best thing to say about the world?

To illustrate the difficulty, suppose that at the end of the day there are exactly two apt models of a given situation, both of which use the variables  $X$ ,  $Y$ , and  $Z$  to respectively represent the events  $c$ ,  $d$ , and  $e$ . And suppose that one of these two apt models has  $X = x$  as the actual cause of  $Z = z$  and represents no other actual causation, while the other has  $Y = y$  as the actual cause of  $Z = z$  and represents no other actual causation. What should one say about the causal relationships in the world? If one adopts the existential requirement that there be an apt causal model representing causation (Hitchcock 2001), then one concludes that both  $c$  and  $d$  both cause  $e$ . This is surprising: one thing that all of the apt causal models agreed on was that  $e$  has exactly one cause and not two. If one adopts the universal requirement that every apt causal model represents causation, then one concludes instead that  $e$  is uncaused. Again this is surprising: all of the apt causal models agreed that  $e$  has exactly one cause and not none. One might also consider further options including supervaluating over apt models, and saying that whether one event causes another is *relative to a representation*, so that the causal facts include the fact that  $c$  causes  $e$  relative to the first model, and the fact that  $c$  does not cause  $e$  relative to the second model (Halpern and Pearl 2005, p. 845). For many metaphysicians who think of causation as an objective feature of the natural world, such representation-relativity may seem shocking.<sup>23</sup>

I have sketched some options but have no fixed solution on offer. For present purposes perhaps the best thing to say is simply that structural equation models are no worse in this respect than any other modeling technique, and that various modeling techniques are used throughout the sciences. The problem of apt representations is

<sup>21</sup> The extent to which this is an objective constraint corresponds to the extent to which the truth of counterfactuals is an objective matter. For further exploration of the semantics for counterfactuals embedded in structural equation models see Shulz (2011) and Briggs (2012).

<sup>22</sup> As Halpern and Hitchcock (2010, p. 394) put the point: "A modeler has considerable leeway in choosing which variables to include in a model. Nature does not provide a uniquely correct set of variables." Nature does not seem to provide a uniquely correct set of contrasts either (Schaffer 2010b, pp. 268–269).

<sup>23</sup> Perhaps this should not be so shocking. To the extent that there are significant objective constraints on aptness, and to the extent that all of the apt models tend to agree on clearcut cases, some lingering representation relativity around the margins might be appropriate for the specific concept of token causation. Indeed, as Halpern and Hitchcock (2010, p. 384) note, "the experimental evidence certainly suggests that people's views of causality are subjective, ..."

everyone's problem. Something general needs to be said. Whatever that might be, consider it said for this specific case.

### 3 Structural equation models for grounding

With structural equations models for causation (Sect. 2) as a template, it remains to display a formalism for grounding. Doing so completes the sole remaining task from Sect. 1.1, that of folding the concept of grounding into a formalism, so as to detail the internal structure of the concept and its external connections to other concepts. This choice of formalism is motivated by the depth of the analogy with causation (Sects. 1.3–1.4) and the success and richness of structural equation models for causation (Sect. 2).

#### 3.1 The formalism itself

The formalism itself may be ported directly: it needs no adjustment whatsoever. It is an excellent formalism for modeling directed dependency relations generally, as can be seen by the way in which it proves equally fitting for modeling both causal and metaphysical dependencies.

So to build a structural equation model for grounding, one may proceed in the same three-step way as with causation, starting with representing the system under study mathematically via variables. Just as in the causal case, the models come with a built-in distinction between the most independent conditions of the system under study and the relatively dependent conditions (these are the “initial conditions” on a causal interpretation, and the “fundamental conditions” on a metaphysical interpretation). And just as in the causal case, the models come with a built-in space of alternatives for each condition (these are “the contrasts”). Accordingly one starts with a signature  $S = \langle U, V, R \rangle$ . As before,  $U$  is a finite set of exogenous variables,  $V$  is a finite set of endogenous variables, and  $R$  is a function mapping every variable  $X \in U \cup V$  to an at-least-two-membered set of allotted values. For instance, if one is studying a shirt which is red because maroon, one might begin with the very simple signature  $S2 = \langle U2 = \{Determinate\}, V2 = \{Determinable\}, R2 \rangle$ , where  $R2$  maps *Determinate* to  $\{0, 1\}$  contrasting the shirt's being maroon (1) with its being navy (0), and maps *Determinable* to  $\{0, 1\}$  contrasting the shirt's being red (1) with its being blue (0).

Just as in the causal case, the models will then incorporate dependency functions evaluating dependent conditions on the basis of their parents (these are “the dynamics” on a causal interpretation, and “the bridge principles” on a metaphysical interpretation). So again one introduces *the linkage*, which is a pair  $L = \langle S, E \rangle$  where  $S$  is a signature as just characterized, and  $E$  is a set of structural equations. As before,  $E$  has to cover every endogenous variable, and cannot permit loops. In the case of the shirt which is red because maroon, a natural linkage is  $L2 = \langle S2, E2 \rangle$ , where  $E2$  is  $\{Determinable \leftarrow Determinate\}$  (outputting a 0 for *Determinable* given a 0 for *Determinate*, and a 1 for *Determinable* given a 1 for *Determinate*). This encodes how the shirt's determinate shade sets its determinable color.

Just as in the causal case, one now just needs to say *how things actually are*. So again one adds in *the assignment*, which is a pair  $M = \langle L, A \rangle$  where  $L$  is a linkage as just

characterized, and  $A$  is the smallest function mapping every exogenous variable  $V \in U$  to a value. With the shirt, one needs to say which determinate shade the shirt has (maroon and navy being the options under consideration). So one just adds  $M2 = \langle L2, A2 \rangle$ , where  $A2$  is the smallest function mapping *Determinate* to 1. Putting this all together:

$S2 = \langle \{Determinate\}, \{Determinable\}, R2 \rangle$ , where  $R2$  maps *Determinate* and *Determinable* to  $\{0, 1\}$

$L2 = \langle S2, \{Determinable \leftarrow Determinate\} \rangle$

$M2 = \langle L2, \{ \langle Determinate, 1 \rangle \} \rangle$

So one can model the case of the shirt which is red because maroon in a way isomorphic to the case of the rock being thrown through the window (Sect. 2.2). Graphically this is just:

*Determinate*  *Determinable*

Mathematically nothing has changed, beyond the purely decorative matter of the labels on the variables.

The point of porting the formalism directly is not that the formalism is a magical panacea, but merely that it implements some core ideas that prove useful (Sect. 2.2), namely:

- independent and dependent conditions are distinguished from the start,
- each condition is situated within a space of contrasts,
- each dependent condition is associated with a function evaluating it on the basis of the values of its parents (subject to global acyclicity constraints), and
- each independent condition is assigned a value.

Formal details aside, these are the core ideas guiding the leading approach to understanding the internal structure and external conceptual connections of causation. They are equally useful as a guide to grounding.

### 3.2 Surrounding notions: interventions, counterfactuals, grounding

What does need adjustment are some of the background understandings of surrounding notions. As mentioned in Sect. 2.3, the crucial work in the causal case includes:

- offering a precise treatment of the notion of an intervention,
- writing a precise semantics for a certain kind of counterfactual, and
- formulating counterfactual tests for token causation.

It needs to be shown that this work can still be done on a grounding interpretation.

The mathematical definitions of intervention and the associated semantics for counterfactuals may remain unchanged, though of course the idea of cutting incoming links and hand-assigning the values now corresponds to treating the relevant condition as metaphysically fundamental rather than causally initial. And the idea that one can “surgically intervene” requires a kind of modularity condition on grounding, which

corresponds to a free recombination assumption for the more fundamental: one can adjust one of the more fundamental parameters while leaving the others as is. But it is worth mentioning that, on a grounding interpretation, the formalism must tolerate a range of connected and potentially worrisome phenomena including *counterlogical interventions*, *countermetaphysical antecedents*, and *hyperintentional connections*.

To illustrate some of these issues, suppose that the pure sets exist with metaphysical necessity, and consider a variant of Fine's example (Sect. 1.2) of the dependence of  $\{\text{Socrates}\}$  on Socrates, concerning the analogous dependence among the pure sets of  $\{\emptyset\}$  on  $\emptyset$ . One very simple way to model this is via another model with two connected binary variables, with *Empty* representing the empty set's existing (1) versus not existing (0), and *Singleton* representing it's singleton's existing (1) versus not existing (0):

$$\mathbf{S3} = \langle \{\text{Empty}\}, \{\text{Singleton}\}, \mathbf{R3} \rangle, \text{ where } \mathbf{R3} \text{ maps both } \text{Empty} \text{ and } \text{Singleton} \text{ to } \{0, 1\}$$

$$\mathbf{L3} = \langle \mathbf{S3}, \{\text{Singleton} \leftarrow \text{Empty}\} \rangle$$

$$\mathbf{M3} = \langle \mathbf{L3}, \{\langle \text{Empty}, 1 \rangle\} \rangle$$

So far nothing seems new. Mathematically this is still the same structure as seen with the rock being thrown through the window, and the shirt which is red because maroon. Graphically this is yet again just:



But given the role of interventions in the system, the model allows one to consider an intervention setting *Singleton* to 0. This leads to the following modified model:

$$\mathbf{S3}^* = \langle \{\text{Empty}, \text{Singleton}\}, \{\}, \mathbf{R3} \rangle, \text{ where } \mathbf{R3} \text{ maps both } \text{Empty} \text{ and } \text{Singleton} \text{ to } \{0, 1\}$$

$$\mathbf{L3}^* = \langle \mathbf{S3}^*, \{\} \rangle$$

$$\mathbf{M3}^* = \langle \mathbf{L3}^*, \{\langle \text{Empty}, 1 \rangle, \langle \text{Singleton}, 0 \rangle\} \rangle$$

This modified model may seem shocking: it represents matters as if  $\emptyset$  and  $\{\emptyset\}$  were independent fundamental conditions, with only the former existing. This is a countermetaphysical scenario. (In the logic cases discussed in Sect. 3.4, one can even get counterlogical modified models, such as a model with the atomic propositions  $p$  and  $q$  being true but with the disjunctive proposition  $pvq$  set to false by a counterlogical intervention.)

There is nothing formally problematic in terms of the internal mathematics. The mathematics doesn't "know" if an intervention is countermetaphysical or counterlogical. It just sees adjusted values to variables and adjusted functions, which it solves as before. But there may be interpretive worries about what these mean. In this vein Woodward (2008, p. 224; also Woodward 2003, pp. 112–116) considers what happens "if we are asked to consider hypothetical interventions that make it the case that  $2 + 2 \neq 4$  or that the same object is at the same time both pure gold and pure aluminum or that transform human beings into houseflies." He says that the meanings of these interventions are "unclear or at least have no legitimate role

in empirical inquiry.” His underlying worry (2008, pp. 226–227) is that the contrast situation is unclear.<sup>24</sup> In a related vein, Spirtes and Scheines (2004, pp. 840–845) consider “ambiguous manipulations” such as manipulations of total cholesterol level, assessed for impact on heart disease. Assuming that total cholesterol level is HDL + LDL, and that HDL incites while LDL prevents heart disease, intervening to halve total cholesterol seems crucially underspecified. One needs to know what proportion of the resultant total cholesterol is HDL and what LDL in order to assess what if any impact on heart disease can be expected.

I have two replies to these worries, the first of which is that these worries arise as much on a causal interpretation of the formalism as on a grounding interpretation of the formalism (as made clear by the examples just given). *The proper solution to these worries, wherever they arise, is to make the contrast situation clear.* Accordingly one ought to add, as an additional constraint on model aptness (Sect. 2.4) sistered to the constraint that the values of each individual variable should represent incompatible contrasts, the further constraint that:

- the values of each individual variable should represent relevantly specific contrasts.

Specificity is presumably going to need to be assessed case-by-case, relative to the outcomes of interest. For instance, if one has a “total cholesterol monitor” that reliably indicates total cholesterol levels and one wishes to model the relationship between total cholesterol levels and monitor readings, then one can simply work with contrasts specifying different levels of total cholesterol. But if one wishes to model the relationship between total cholesterol level and heart disease (given that HDL and LDL contribute differently), then one needs to work with more specific contrasts. (This is connected to the aptness constraint that the counterfactuals encoded in the model’s equations must be true. The contrasts need to be relevantly specific enough for the counterfactuals to have determinate truth values.)

My second reply to these worries is to offer an opening for grounding skeptics (Sect. 4.4). If there is something specifically incomprehensible or unspecifiable or otherwise unclear about countermetaphysical scenarios, then there will be something specifically inapt about structural equation models on a grounding interpretation, at least in the application to some of the paradigm cases. I myself think that countermetaphysical and counterlogical scenarios were already independently needed (see generally Nolan 1997, pp. 536–541). Assuming just for the sake of the example that classical logic is the one true logic, one still wants to assign the right truth-values to counterfactuals such as: ‘if intuitionist logic were the one true

<sup>24</sup> As Woodward (2003, pp. 115–117) elaborates the point, even common and seemingly well-confirmed causal claims such as ‘being female causes one to be discriminated against in hiring and/or salary’ still need clarification, since there are several different things one might mean by the notion of an “intervention” on the value of a female/male variable, and since these differences might well be relevantly different in causal impact. For instance, one might have in mind a “far-backtracking” intervention in which the female candidate was born male, or one might have in mind a “near-backtracking” intervention in which the female candidate was subjected to a very recent sex change operation and an accompanying barrage of hormonal treatments. These distinct interventions can plausibly be expected to have *opposite* effects on how the candidate would be treated in a real-world hiring situation, given the various prejudices now existing.



logic, then the law of excluded middle would not hold,’ and ‘if intuitionist logic were the one true logic, then the law of excluded middle would still hold.’ It seems to me—given what intuitionistic logic says—that the former is true and the latter false, which requires a non-vacuous treatment of counterlogicals. (Likewise if one wants to model the content of the beliefs of an intuitionistic logician in terms of a set of points in a space, one needs counterlogical points.) But for present purposes I only mean to say that the use of structural equation models for grounding does in some cases require countermetaphysical and counterlogical scenarios. I thus offer the grounding skeptic an opportunity to make a more precise complaint than “whadda you mean by ‘grounding’?” at a level where one can engage in a clearer and already explored issue about countermetaphysicals and counterlogicals. It is a good-making feature of a framework for grounding if it clarifies the opportunities for skeptical reaction.<sup>25</sup>

Relatedly, the system must not only tolerate countermetaphysical (and counterlogical) interventions, it must also tolerate countermetaphysical antecedents. It must allow a non-vacuous treatment of the hypothesis that the empty set does not exist, even on the assumption currently in play that the empty set exists with metaphysical necessity. Mathematically, the formula for evaluating counterfactuals in this system can already do what it must. Given that **R3** is as it is, and that **L3** is as it is, there is a precise answer in **M3** as to what value *Singleton* would take if *Empty* = 0. Since *Empty* is already exogenous one need not cut any incoming links (there are none). One need only revise the assignment as **M3\*\*** = <**L3**, {<*Empty*, 0>}, and plug in the equation *Singleton* ← *Empty* to derive *Singleton* = 0 in **M3\*\***.

Finally, the semantics also must make *hyperintensional* distinctions, in seeing an asymmetric dependence of { $\emptyset$ } on  $\emptyset$ , even on the assumption currently in play that both exist at every metaphysical possibility.<sup>26</sup> Fortunately, the formalism can already do what it must in this respect as well. (Again the formalism just “sees” variables and functions relating their values.) Whether or not these entities are coexistent with metaphysical necessity, one can represent them separately simply by modeling them with separate variables (*Singleton* and *Empty*), and one can encode substantive claims of directed dependency between such entities, simply by modeling them as linked by an equation in which the dependent term appears on the left hand side (*Singleton* ← *Empty*).

<sup>25</sup> A different concern: The framework itself looks to build in some metaphysical assumptions such as the directedness of ground and the contrastivity of various alternatives, as well as logical assumptions such as the evaluation of functional expressions. Can there be mutilated models in which these very framework-structuring assumptions break down? For instance, does it make sense to evaluate the output of a given function in a scenario in which that very function is imagined to break down? I think one can distinguish the logic of the modeling language from the logic of the scenario being modeled, but there are difficult issues lurking.

<sup>26</sup> Whether this is true hyperintensionality depends on the range of possible worlds one countenances. If metaphysical possibility is the widest sense of possibility then this is true hyperintensionality. But if—as I think—metaphysical possibility is itself a restricted sense of possibility (perhaps restricted to worlds with common laws of metaphysics, just as nomological possibility is usually thought to be restricted to worlds with common laws of nature) then there may be room for an intensional distinction between  $\emptyset$  and { $\emptyset$ } after all, at “worlds” with different set theoretic principles. In the main text I will continue to label this “hyperintensionality” to accord with the extant literature.

With counterfactuals in hand, *Counterfactual dependence test* and *Contrastive counterfactual test* pass through simply by changing ‘cause’ to ‘ground’:

*Counterfactual dependence test for grounding*: If  $X$  and  $Y$  are binary variables, there is a direct  $X \rightarrow Y$  path, and no other distinct path to  $Y$ , then:  $X = x$  is a token *ground* of  $Y = y$  if and only if  $X = x$ ,  $Y = y$ , and if  $X \neq x$  then  $Y \neq y$ .

*Contrastive counterfactual test for grounding*: If there is a direct  $X \rightarrow Y$  path, and no other distinct path to  $Y$ , then:  $X = x$  rather than  $x^*$  is a token *ground* of  $Y = y$  rather than  $y^*$  if and only if  $X = x$ ,  $Y = y$ , and if  $X = x^*$  then  $Y = y^*$ .

So there is a straightforward and informative parallel working test of token grounding to be had, in terms of counterfactual covariation: wiggle the ground, and the grounded wiggles. This allows one to conclude, in the model of the shirt which is red because maroon, and the model of the singleton which exists because its member exists, that the model is capturing natural intuitions about what grounds what. (In Sect. 3.4 I consider the wider range of paradigm cases mentioned in Sect. 1.2. But these are the first two tastes of success.)

### 3.3 Surrounding notions: reduction and representation

The structural equations framework is *conceptually nonreductive* (Sect. 2.4), and remains so as applied to grounding. The division into independent and dependent variables, and the structural equations encoding directed dependencies, are already laden with ground-theoretic notions. Though many have wanted a reductive conceptual analysis of causation (Paul and Hall 2013, pp. 7–8; cf. Anscombe 1975), with grounding most have been content to accept the notion as primitive (Fine 2001, p. 1; Schaffer 2009, p. 364; Rosen 2010, p. 113). The friend of grounding-as-a-primitive should only hope to see it enfolded within such a rich nonreductive framework.

Again note that a conceptually nonreductive treatment is compatible with all sorts of views about the metaphysics, from the idea that grounding is a primitive and fundamental constituent of reality, to the idea that grounding is a real but derivative phenomenon, through to the idea that grounding is a crude concept corresponding to nothing. My preferred view on the metaphysics is that grounding is a real but derivative phenomenon, derivable from the laws of metaphysics. But again my present focus is on the conceptual side. I am saying that structural equation models provide an excellent picture of how ground-theoretic notions are laced together and threaded through the larger tapestry, on a pattern that closely parallels causal notions.

A final point of adjustment concerns the background constraints on model aptness, some of which were reflecting the idea of causation as a relation between distinct events. Just looking at the partial list of aptness constraints specified, what one wants is something like the following heuristics (I’ve italicized the bits changed from the causal case in Sect. 2.4, and included the bit about specificity of contrasts from Sect. 3.2):

- the values of distinct variables should all represent *non-identical entities*,

- the values of each individual variable should represent incompatible contrast *entities*,
- the values of each individual variable should represent relevantly specific contrasts,
- the variables should represent sufficiently many *entities* to capture the structure of the situation,
- the counterfactuals encoded in the model's equations should be true, and
- the assignment should represent the *fundamental* conditions correctly.

This represents three adjustments. The final adjustment is to replace 'initial conditions' with 'fundamental conditions', reflecting the shift from causation to grounding, and not calling for further comment.

The second adjustment is to replace 'events' with the more general 'entities.' This calls for comment insofar as it is controversial whether to think of grounding as a relation between entities (Schaffer 2009), or as a relation just between facts (Rosen 2010), or instead as a sentential operator (Fine 2012).<sup>27</sup> But it is clear what thinking of grounding in the image of causation demands: one can also regiment causal claims via a relational predicate over facts or a sentential operator, but with causation the standard view is that the most metaphysically perspicuous regimentation is via a relational predicate whose relata are *events* or some other concrete elements of physical reality. The relation between facts or sentential operation in the neighborhood is not causation but *explanation*. The frame 'S because T, U, V, ...' is at best a first-pass regimentation of explanatory patterns, and has no plausibility as a regimentation of the causal relations that back these explanatory patterns. To treat grounding as a relation between facts or sentential operation is to conflate the image of causation with the image of causal explanation (Sects. 4.1–4.3).

That said, those who would insist that grounding can only relate facts are welcome to replace 'entities' with 'facts'. I think they are confused, but not in ways that preclude them from stating reasonable aptness conditions for structural equation models for grounding.<sup>28</sup> Those who would insist that grounding is best treated as a sentential operator may try to speak of distinctness conditions over something like the subject matters of the sentences involved, but I leave it to them to work out the details.

The third and perhaps most interesting adjustment to the aptness constraints comes in replacing the "distinctness" requirement with a weaker "non-identity"

<sup>27</sup> This is the distinction between what Correia and Schnieder (2012, pp. 10–11; see also Correia 2005, p. 48) label "the predicational view" on which grounding claims are best regimented by the relational predicate 'is grounded in' flanked by singular terms, and "the operational view" on which grounding claims are best regimented by a sentential operator 'because' (understood as taking a metaphysical reading).

<sup>28</sup> As Handfield et al. (2008, p. 151) rightly note, the structural equation formalism itself is neutral on the metaphysical nature of the relata: "The formal structure of a causal model does not require a variable to represent events, event-types, propositions, states-of-affairs, or any other particular members of the ontological zoo. Causal models, then, provide us with a blank screen onto which we can project our metaphysical claims." More precisely, when looking at a mathematical statement assigning a value to a variable (e.g. ' $X = x$ ') all that is required is that the variable possess a range of values, and that the values contrast with each other. Any entities that can be contrasted are thus formally eligible as relata.

requirement. This is weaker: distinctness (in the sense relevant for causation) is supposed to cover not just identity but also conditions such as mereological overlap. For instance, the performance of the first four acts does not cause the performance of the whole five act play, even though it is true that if the first four acts had not been performed then the whole five act play would not have been performed. Explanation: the performance of the first four acts and the performance of the whole five act play are non-identical but still non-distinct events. They are non-distinct due to overlap.

Indeed it was never very clear what exactly ‘distinct’ was supposed to mean in the causal case. I suggest a ground-theoretic reading, on which distinct entities are *neither identical nor connected by grounding* (neither grounds the other, nor do they have a common ground). Metaphysically, distinct entities are wholly separable portions of reality, with no common roots.<sup>29</sup> On this ground-theoretic treatment of distinctness, given entities  $a$  and  $b$ , there are then three options: identity ( $a = b$ ), grounding-connection ( $a$  grounds  $b$ ,  $b$  grounds  $a$ , or  $a \neq b$  but and there is some entity  $c$  which grounds both  $a$  and  $b$ ), or distinctness. (On this approach, any two non-identical fundamental entities count thereby as distinct. This connects the idea that there are no necessary connections between distinct existences, with the idea that fundamental entities are “degrees of freedom” and indeed set the parameters of possibility.<sup>30</sup>)

So given that distinctness rules out there being a grounding connection, one does not want to insist that the variables in a structural equations model for grounding must represent that which is distinct, on pain of disallowing any connections at all in the grounding models! Accordingly, the condition for model aptness is simply that the values of distinct variables should represent non-identical entities. This might seem like a mere tweak, but actually I think it points to a subtle but deep difference between causation and grounding (Sect. 4.5): causation is an external relation linking distinct portions of reality, while grounding is an internal relation operating within a given portion of reality.

### 3.4 Handling paradigm cases

With the formalism and associated notions in hand, I am ready to return to the paradigm cases of grounding (Sect. 1.2), in order to show that structural equation

<sup>29</sup> Wilson (2010, p. 601) suggests that two notions of distinctness have been conflated in the literature: nonidentity, and the capacity for either entity to exist without the other. I am suggesting a third sense of the notion in grounding-theoretic terms. This notion of distinctness may not be so far from what Hume himself (1978, p. 634) had in mind: “Whatever is distinct, is distinguishable; and whatever is distinguishable, is separable by the thought or imagination. All perceptions are distinct. They are, therefore, distinguishable and separable, and may be conceiv’d as separately existent, and may exist separately, without any contradiction or absurdity.”

<sup>30</sup> Conversely, if everything is grounded in a common fundamental entity (as on the monistic hypothesis of Schaffer (2010a) in which everything is grounded in the cosmos as a whole, but also as on many theistic views in which everything is grounded in God), then nothing counts as ultimately distinct and there cannot be true causal connections in nature. On such a view, there is no fundamental causation, but rather merely a derivative and approximate relation between derivative events that are only approximately distinct.

models are capable of handling every case. By “handling every case” I mean providing plausibly apt models that, in conjunction with the counterfactual dependence test (or plausible extensions thereof) deliver the intuitive verdict about what grounds what.

I should emphasize from the start that I am *not* trying to show that merely by adopting the structural equations formalism, one gets the right answers to grounding questions for free. On the contrary: to get the right answers to grounding questions one has to put the right structural equations into the model, which encode the form and direction of dependency. This is also true for structural equation models for causation. And an analogous point is true for other frameworks for thinking about grounding: the frameworks themselves are relatively neutral. They provide the language in which various claims about grounding may be made, and through which various controversies about grounding may be pursued.

Nor am I trying to provide a *perfectly neutral framework* for representing any metaphysical hypothesis. Neutrality is a false idol: state any substantive principle whatsoever and it will conflict with some madcap metaphysical hypothesis. Some metaphysical hypotheses deserve to be ruled out, and it is the job of a framework to say as much. Which metaphysical hypotheses are to be ruled out, and which should still count as “in the game” is itself a matter of controversy. Perhaps the best one can do is to aim for a framework that *coheres* with one’s own judgments as to which first-order hypotheses should be in the game.<sup>31</sup>

With that in mind, I begin by recalling that I have provided plausibly apt models of the determinate-determinable case (Sect. 3.1) which—in conjunction with the counterfactual dependence test—delivered the intuitive verdicts that the shirt’s being maroon (rather than navy) grounds its being red (rather than blue). And I have provided plausibly apt models of the member-singleton case (Sect. 3.2) which—in conjunction with the counterfactual dependence test—delivered the intuitive verdicts that the empty set’s existing (rather than not) grounds its singleton’s existing (rather than not). These were the first two “tastes of success” in Sect. 3.2.

For brevity I do not consider every one of the baker’s dozen of (clusters of) cases mentioned in Sect. 1.2, but focus on a handful, leaving the rest to the reader. My goal is to write out plausibly apt models which encode the patterns of counterfactual covariation indicative of grounding. I do not argue that the models are plausible or walk through the derivation of the counterfactuals. I simply present the models along with the induced graph as an aid to visualization.

<sup>31</sup> For instance, Fine (2001, p. 26) wants to accommodate the “neo-expressivist” view that ethical claims are *ungrounded* but still *unreal*, and so rejects the identification of the real with the ungrounded, positing separate primitive notions of *grounding* and *reality* instead. Sider (2011, pp. 125–126) claims that his framework—with a primitive ‘structural’ operator mapping terms to truth when they carve at the joints—is better able to accommodate this view. I think that the ethical view in question is incoherent, and that it is a virtue of a framework to say as much. (I think that the neo-expressivist whom Fine and Sider are trying to accommodate is better understood as saying that ethical claims are grounded in a distinctive way, partly via our conative states. She thinks that the moral facts are mind-dependent). For present purposes I am saying that Fine and Sider are within their rights to accommodate the theory in their frameworks (they think it coherent), and that I am within my right to exclude it from my framework (I think it incoherent). There is not even a neutral notion of neutrality.

Starting with truthmaking, the same simple two-connected-binary-variable structure is already sufficient for this case:

$S4 = \langle \{Socrates\}, \{Proposition\}, R4 \rangle$ , where  $R4$  maps both *Socrates* and *Proposition* to  $\{0, 1\}$

$L4 = \langle S4, \{Proposition \leftarrow Socrates\} \rangle$

$M4 = \langle L4, \langle Socrates, 1 \rangle \rangle$

*Socrates*  $\longrightarrow$  *Proposition*

Using *Socrates* = 1 to model the existence of Socrates and *Socrates* = 0 to model nonexistence, and using *Proposition* = 1 to model the truth of the proposition <Socrates exists>, and *Proposition* = 0 to model falsehood, one arrives at a plausibly apt model in which the truth of the proposition <Socrates exists> counterfactually depends on the existence of Socrates. This exhibits the dependence of truth on being. Indeed the same structure can model how the truth of the proposition <Socrates is snub-nosed> depends on the nasal features of Socrates. Simply reinterpret *Socrates* = 1 to model Socrates being snub-nosed, and *Socrates* = 0 to model aquilinity, and reinterpret *Proposition* = 1 to model the truth of the proposition <Socrates is snub-nosed>, and *Proposition* = 0 to model falsehood.

Turning to levels, in order to illustrate something with a bit more complexity, consider a model of atoms forming carbon monoxide (CO) and salt (NaCl) molecules, and in turn forming a chemical compound (CO + NaCl). So, using obvious labels for the variables, and using 1 values to represent existence and 0 nonexistence, let the signature be:

$S5 = \langle \{C, O, Na, Cl\}, \{CO, NaCl, CO + NaCl\}, R5 \rangle$ , where  $R5$  maps all variables to  $\{0, 1\}$

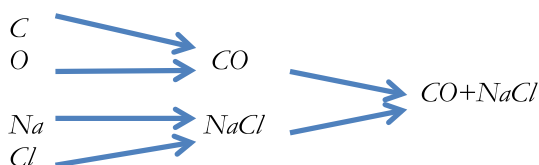
For the linkage one wants to set *CO* to 1 if and only if both *C* and *O* are at 1, and likewise one wants to set *NaCl* to 1 if and only if both *Na* and *Cl* are at 1, and moreover one wants to set the variable for the total compound *CO + NaCl* to 1 if and only if both *CO* and *NaCl* are at 1:

$L5 = \langle S5, \{CO \leftarrow \min(C,O), NaCl \leftarrow \min(Na,Cl), CO + NaCl \leftarrow \min(CO, NaCl)\} \rangle$

Finally I am imagining that actually all of the atoms are present, so the assignment should be:

$M5 = \langle L5, \langle C, 1 \rangle, \langle O, 1 \rangle, \langle Na, 1 \rangle, \langle Cl, 1 \rangle \rangle$

As a useful visualization, the induced directed acyclic graph is:



This befits the dependence of molecule on atom, compound on molecule, and compound back on atom. Strictly speaking this is a case with indirect links and multiple in-paths in which the counterfactual dependence test becomes risky. But every plausible extension of the test agrees on the grounding results.

The logic cases work as follows. The structure of conjunction-type dependence is already exhibited in the *min*-functions employed in **L5**. So here is how to encode the dependence of  $\langle p \ \& \ q \rangle$  on distinct atomic propositions  $\langle p \rangle$  and  $\langle q \rangle$ , using 0 values for falsity and 1 values for truth:

$S6 = \langle \{P, Q\}, \{Conj\}, R6 \rangle$ , where  $R6$  maps all variables to  $\{0, 1\}$

$L6 = \langle S6, \{Conj \leftarrow \min(P, Q)\} \rangle$

$M6 = \langle L6, \{ \langle P, 1 \rangle, \langle Q, 1 \rangle \} \rangle$



The structure of disjunction-type dependence with distinct atomic disjuncts may be obtained by relabeling ‘*Conj*’ as ‘*Disj*’ and switching from a *min* to a *max* function ( $Disj \leftarrow \max(P, Q)$ ). In this case one loses counterfactual dependence due to grounding overdetermination. But virtually every plausible extension of the counterfactual dependence test agrees on the grounding result in this case as well. So one gets the result that each true conjunct is a token ground of the true conjunction, and each true disjunct is likewise a token ground of the true disjunction.<sup>32</sup>

For (classical) negation one can revert to a two-connected-binary-variable structure, but with a *flip* function, as per: ‘ $NotP \leftarrow 1 - P$ ’. (The intended model should be obvious). Simple counterfactual dependence returns, delivering the intuitive verdict that the value of the negation is grounded in the value of its negatum.

The quantifier cases may simply be treated as (potentially infinitary) conjunctions/disjunctions, but in a way that calls for at least two comments. First,

<sup>32</sup> Note that the system distinguishes conjunctive-type dependence from disjunctive-type dependence, not by needing to take a primitive plural notion of complete ground (so as to say that  $p, q$  is the only complete ground for the conjunction, but that  $p$  itself is a complete ground for the disjunction), but rather by articulating the *form* of the grounding connection in different ways. With 0 representing falsity and 1 truth, conjunction is a *min* function but disjunction is a *max* function.

I have adopted the simplifying expedient of restricting myself to finitely many variables, so the formalism needs complication (the case of infinitely many parents must be tolerated). Secondly and more interestingly, there is controversy about whether universally quantified claims, or negative existential claims, require an additional ground in something like a totality fact (Russell 1985; Armstrong 1997, pp. 196–201). This is a matter on which the system itself is neutral, and depends on which extra-formal modal constraints are imposed on apt grounding models. If one imposes a necessitation requirement on the grounding relations encoded in apt models and allows for the possibility of expanded domains, then models without a totality fact will fail to be apt, but otherwise—for instance if one merely requires supervenience, as I prefer—then there is no such requirement (Lewis 2001; Bricker 2006).

With the case of substance and mode one sees a first “forked” structure (analogous to a causal explosion, in which one event produces many), as well as a useful opportunity to consider contrastive patterns. So consider a toy model of an apple as a “thick particular”, with its redness and roundness understood as dependent abstractions from it. I am supposing that the natures of this redness and that roundness are both grounded in how the apple is. So one will want a variable answering to the question of how the apple is (*Apple*) and variables answering to the questions of how its color and its shape modes are (*Red*, *Round*). And—just to display a bit more complexity—I will allot *Red* and *Round* each three values {1, 2, 3} to represent whether there is perfect redness/roundness, partial redness/roundness, or none at all. Correspondingly I will allot *Apple* nine values {1–9} to represent the relevantly complete ways it can be, which are going to wind up (once the structural equations are put in) coding points in a  $3 \times 3$  color/shape statespace:

<u><i>Apple=9</i></u> <i>Red=3</i> <i>Round=3</i>	<u><i>Apple=8</i></u> <i>Red=3</i> <i>Round=2</i>	<u><i>Apple=7</i></u> <i>Red=3</i> <i>Round=1</i>
<u><i>Apple=6</i></u> <i>Red=2</i> <i>Round=3</i>	<u><i>Apple=5</i></u> <i>Red=2</i> <i>Round=2</i>	<u><i>Apple=4</i></u> <i>Red=2</i> <i>Round=1</i>
<u><i>Apple=3</i></u> <i>Red=1</i> <i>Round=3</i>	<u><i>Apple=2</i></u> <i>Red=1</i> <i>Round=2</i>	<u><i>Apple=1</i></u> <i>Red=1</i> <i>Round=1</i>

So, supposing that the apple is in fact a 9, making it perfectly red and perfectly round:

$S7 = \langle \{Apple\}, \{Red, Round\}, \mathbf{R7} \rangle$ , where  $\mathbf{R7}$  maps *Apple* to {1–9} and *Red* and *Round* to {1–3}



$L7 = \langle S7, \{Red \leftarrow \text{ceiling}(Apple/3), Round \leftarrow \min(x: x \text{ is a pos. integer congruent to } Apple \text{ mod } 3)\} \rangle$ <sup>33</sup>  
 $M7 = \langle L7, \{ \langle Apple, 9 \rangle \} \rangle$

Graphically:



Contrastive counterfactual dependence comes into play, and one gets plausible results such as that *Apple* = 9 rather than 8 grounds *Round* = 3 rather than 2, but does not ground *Red* = 3 rather than 2 (wiggling *Apple* from 9 to 8 wiggles *Round* from 3 to 2, but leaves *Red* stuck on 3). *Apple* = 9 rather than 6 grounds *Red* = 3 rather than 2, but does not ground *Round* = 3 rather than 2 (wiggling *Apple* from 9 to 6 wiggles *Red* from 3 to 2, but leaves *Round* stuck on 3). And *Apple* = 9 rather than 1 grounds both *Red* = 3 rather than 1 and *Round* = 3 rather than 1.

The monist who thinks of parts as dependent abstractions from concrete wholes essentially deploys the same sort of forked model as this modist. For her the apple and its dependent modes are akin in grounding structure to the cosmos and its abstracted portions. (The pluralist who thinks of wholes as dependent constructions from simpler parts may put forward a conjunction-type model as an alternative. The structural equations framework allows both parties to state their view and hold their debate. By my lights this is an appropriate neutrality.)

I leave the remaining paradigm cases to the reader. I should re-emphasize that (as with causation) the main thing one needs to do is to find the right equations. In some cases—such as the logic cases—the work will essentially reiterate (albeit in different terms) the work of Fine (2012) and Rosen (2010), who have focused on developing grounding principles apt for the logic cases. In other cases—such as the levels cases—the work will need to draw on detailed empirical knowledge. In still other cases—such as the debate between the monist and pluralist over the direction of dependency in parthood cases—the work will turn on subtle metaphysical considerations. The point is not to *replace* substantive metaphysical inquiry into logic or levels or metaphysics or other such matters, but to *unify* these inquiries into a fitting common framework. (By way of analogy, structural equations models for causation do not replace detailed biological or medical or sociological investigations into different sorts of mechanisms etc., but unify these inquiries.)

One sort of complaint I sometimes hear about grounding is that it threatens to replace serious logical, empirical, or metaphysical inquiry into the detailed hows

<sup>33</sup> *L7* deploys a ceiling function, where  $\text{ceiling}(x)$  is the smallest integer  $y$  such that  $y \geq x$ . The function takes  $Apple/3$  and raises the result (if need be) up to the nearest integer value. *L7* also deploys the notion of the minimal positive integer congruent to  $Apple \text{ mod } 3$ .  $\{1, 4, 7\}$  are congruent mod 3, as are  $\{2, 5, 8\}$  and  $\{3, 6, 9\}$ . This function is effectively taking the smallest representative of these three equivalence classes. These functions are just providing a compact encoding of the mappings seen in the statespace diagram in the main text above.

and whys of dependency with the nearly empty formula ‘this grounds that’ (Wilson 2014, pp. 542–548). However, apt that complaint might or might not be on other ways of thinking about grounding, it is not applicable to the structural equations framework, insofar as this framework includes functions codifying the shape of the dependence pattern between conditions, and insofar as the correct shape of these functions turns precisely on the detailed hows and whys of dependency. The structural equations framework does not ignore these matters but integrates them.

### 3.5 What has been wrought

I have presented a treatment of grounding in the image of causation, through the use of structural equation models for grounding. This completes the final task of enfolding the concept of grounding within a formalism (thus detailing the internal structure of the concept and its external connections to other concepts). Since the structural equations framework is our leading approach to understanding the internal structure and external connections of causation (Sect. 2), by enfolding grounding in a parallel formalism one thereby predicts a parallel internal structure and external connections, just as seen in Sects. 1.3–1.4. The grounding-causation analogy thus deepens to include an additional italicized point:

- both are generative relations;
- both are partial orders, admitting a type/token distinction, a component/net distinction, an incomplete/complete/total distinction, and screening-off relations;
- both are backed by non-accidental generalizations, delimit a specific form of necessity, are supportive of and diagnosable by counterfactuals, and can back explanation; and
- *both can be fruitfully formalized via structural equation models.*

I am proposing that the final italicized point connects and systematizes the other points.

As with causation, I do not claim that the structural equations framework is a magical panacea, dissolving all problems. General issues about grounding remain which I cannot attempt to discuss here, including Fine’s (2010) puzzles of ground, and Sider’s (2011, p. 107; see also Bennett 2011b; deRosset 2013) question of what grounds the grounding facts themselves. I am claiming that the structural equations approach to grounding has virtues, including the virtues of allowing one to:

- enfold grounding within a precise formalism (Sects. 3.1–3.3),
- handle the paradigm cases (Sect. 3.4),
- explain the analogy with causation, and
- unify the input to explanation.

The first three bulleted points should need no further commentary at this stage. I linger briefly on the final point. Causation and grounding both do something very special, which few others relations do. Both back explanation. The question looms: *which relations back explanation, and why?*

I offer a speculation: the relations that back explanation are the relations of *directed dependency*.<sup>34</sup> One can explain why this rather than that, by finding a contrast between one thing and the other on which this rather than that depends. In other words, explanations report difference-makers, and structural equation models are an excellent tool for modeling difference-making. Insofar as structural equation models detail relations of directed dependency through contrast spaces, they model contrastive *why*-questions ('Why this rather than that?') and their potential answers ('Because of the one thing rather than the other'), and the dependency connection that makes the answers fit to be answers. But speculations aside, it seems to me that providing parallel formal treatments of grounding and causation can only be a step in the right direction, as far as understanding how both have the special power to back explanation, and even glimpsing the nature of explanation itself.

All in all I think that causation theorists have something to teach grounding theorists. There has been an extensive and ongoing discussion of causation in which all sorts of different ideas have been trotted out and tested, and there is an emerging consensus as to the most useful background conceptual tools for studying causal structure. Grounding theorist need not re-invent the wheel.

I should also flag that my treatment of grounding has not once mentioned a concept often thought central, namely that of *essence* (cf. Fine 2012). While it is possible to think that the structural equations themselves are connected to essences in some way or another, absolutely nothing in the formalism requires this idea, and not many would want to embrace this idea on the causal interpretation. Those of us (including myself) who eye the notion of essence with suspicion may welcome its separation from grounding.

## 4 Comparisons and controversies

So far I have drawn on the structural equations approach to causation to develop a structural equations approach to grounding, and claimed that this allows one to enfold grounding within a precise formalism, handle the paradigm cases, explain the analogy with causation, and unify the input to explanation. I turn to comparisons with alternative treatments of grounding, and the criticisms of grounding-based approaches as either unintelligible or disunified. I claim that a structural equation framework can also:

- reveal shortcomings in current treatments of grounding, and
- show grounding to be an intelligible and unified posit.

### 4.1 Rivals: Fine

Following Fine (2001; see also Schaffer 2009; Rosen 2010), there has been an explosion of interest in grounding and fundamentality. Yet no current treatment of

<sup>34</sup> Ruben (1990, p. 210) suggests something like this: "[T]he basis for explanation is in metaphysics. Objects or events in the world must really stand in some appropriate 'structural' relation before explanation is possible. Explanations work, when they do, only in virtue of underlying determinative or dependency structural relations in the world."

grounding looks even remotely like a plausible treatment of causation. I focus on Fine and on Rosen, for they have produced some of the most beautiful and influential work on the topic, but what I have to say in a critical vein—which is that the wrong forms have been introduced from the start, and that grounding and metaphysical explanation have been conflated—carries over to virtually all other extant frameworks.<sup>35</sup>

Starting on the causal side of the ledger, one wants to distinguish between *causation*—a concrete relation in the world—and *causal explanation*—an abstract pattern over facts or sentences. And one wants to connect these notions by allowing causal relations in the world to *back* causal explanations among facts or sentences. Or so orthodoxy has it, and so I take for granted here (see also Schaffer 2012, p. 124). So one has a concrete relation between entities backing an explanatory pattern over facts or sentences. Bennett (1988, p. 22)—though himself an opponent of certain aspects of this orthodoxy—expresses the view well when he says: “A fact is a true proposition (they say); it is not something *in* the world but is rather something *about* the world, which makes it categorically wrong for the role of a puller and shover and twister and bender.”

My most fundamental criticism of Fine and Rosen—near allies though we three are, and admirable as their work is—is that they have not preserved this image of relation-backing-explanatory-pattern on the grounding side of the ledger, but have instead conflated the image of causation with the image of causal explanation. I start with Fine. Fine (2012, p. 46; see also Correia and Schnieder 2012; Litland 2013), for his primitive notion of grounding, offers a sentential operator for which ‘because’ is offered as the nearest ordinary language equivalent, regimented as:

*Fine:*  $T, U, V, \dots < S$

Fine’s ‘<’ operates on a plurality of sentences ( $T, U, V, \dots$ ) and a single sentence ( $S$ ).<sup>36</sup> As an illustration Fine gives:

The ball is red, the ball is round < The ball is red and round

He glosses this as “The ball is red and round because the ball is red and the ball is round.”

Fine does not argue that this is an appropriate regimentation, save indirectly by showing what he can do with it. That is, he starts off with his preferred regimentation of grounding claims and then tries to put it to use. This is all perfectly appropriate. Only it cannot be assumed that Fine has found the best starting point.

<sup>35</sup> I make an exception for Wilson (manuscript), who has independently come to a view similar to mine. I also note by way of self-criticism that structural equation models for grounding, which I gesture at in Schaffer (2012, pp. 130–132), do provide strictly more structure than the mere partial ordering appealed to in Schaffer (2009): structural equation models map many-one to partial orders (Sect. 1.2). So I think my earlier work on the topic used an impoverished framework.

<sup>36</sup> Fine goes on to distinguish several different variants of this core notion of grounding, ultimately (Fine 2012, p. 55) taking four different primitive grounding operators (his “weak full,” “weak partial,” “strict full,” and “strict partial” notions of ground). Dasgupta (2014) challenges Fine’s specific claim that ‘<’ should be singular on the right. For present purposes the relevant point is that Fine and Dasgupta both regiment grounding as a ‘because’-like sentential operator.

The formalism of *Fine*, as well as Fine's associated glosses, are evidently explanatory locutions. (A 'because' pattern over sentences is a plausible first pass regimentation for explanation<sup>37</sup>.) Moreover, it should be apparent that *Fine* has no plausibility as a regimentation of the causal relations that back these explanatory patterns. No one thinks that a metaphysically perspicuous regimentation of causation looks like:

The ground was dry, the lightning struck  $<_{\text{cause}}$  There was a fire

A metaphysically perspicuous regimentation of causation should involve a relation not a sentential operator, and arguably should also involve contrasts (Schaffer 2005). It is because there is a causal relation between the events involved, under appropriate contrasts, that there even arises a causal explanation at all, one which might be spoken of via: 'There was a fire because the ground was dry and the lightning struck.'

Perhaps most crucially, *Fine* includes none of the core conceptual background of a cause-like relation. Recall (Sect. 2.2) that the core conceptual structure of causation, insofar as the leading formalism is to be trusted, includes situating conditions within contrast spaces, and encoding functions that articulate the form of the dependence pattern. *Fine* covers none of this. The sentences are not outfitted with specific contrast spaces, save in the limiting and generic sense that they can be in the end true or false. As a result, Fine only has the analogue of binary (true/false) variables in his system. It is not possible to look at dependence patterns involving ternary or higher variables. Also, no functions are given codifying the form of dependence from sentence to sentence, save in the limiting and generic sense that grounding can be on or off. As a result, Fine only has the analogue of identity (true  $\rightarrow$  true) functions in his system. It is not possible to distinguish different patterns of dependence, such as the conjunctive pattern (*min*) and the disjunctive pattern (*max*).<sup>38</sup>

Putting this together, I have three criticisms of *Fine* as a regimentation of grounding claims:

- it regiments grounding as a sentential operator, which is at best apt for the image of explanation not the image of causation;
- it regiments grounding in a noncontrastive format; and
- most crucially, it does not include the core conceptual background structure found in structural equation models, including (a) contrast spaces for the

<sup>37</sup> In this vein, Schnieder (2011, p. 445) aptly speaks of providing a logic for "the explanatory connective 'because'."

<sup>38</sup> Instead Fine must try to capture the distinction between conjunction and disjunction at the level of the grounds themselves. For this he needs a primitive notion of *complete (full) ground*, so as to say that  $p$ , if true, is a complete ground of  $p \vee q$  but not of  $p \& q$  (the complete ground of  $p \& q$  is the plurality  $p, q$ ). That said, Fine (2012, pp. 48–50) considers adding the new notion of *nonfactive ground*, which is something like a possible ground. Adding this new notion would enrich his system to the point where it could be used to capture patterns of dependence. For instance, one could then encode the pattern of conjunctive-type dependence via the following four nonfactive-grounding (' $<_{\circ}$ ') claims: (i)  $p, q <_{\circ} p \& q$ ; (ii)  $p, \sim q <_{\circ} \sim(p \& q)$ ; (iii)  $\sim p, q <_{\circ} \sim(p \& q)$ ; (iv)  $\sim p, \sim q <_{\circ} \sim(p \& q)$ .

conditions under study, and (b) functions codifying the shape of the dependence patterns between conditions.

Thus grounding as Fine conceives of it is nothing like metaphysical causation. This means that Fine cannot claim to explain the analogy with causation, and cannot claim to unify the input to explanation. (Perhaps grounding as Fine conceives of it is really more like metaphysical explanation, which might be backed by what I am calling grounding? For a pessimistic assessment of this reading, see Sect. 4.3.)

This is not to say that Fine's system is expressively inadequate. Expressive power can only be assessed for a fixed slate of notions, and Fine is happy to add on further primitive ground-related notions on an as-needed basis. Nor is this to say that *Fine* is exposed to any specific counterexamples. Imagine that someone wanted to study causation, and—without considering the literature—began by writing down a binary many-one sentential operator:

The ground was dry, the lightning struck  $\leftarrow_{\text{cause}}$  There was a fire

The appropriate response to this person would not be to offer a counterexample. (Counterexamples would not even come into view until constraints on their sentential operator had been stated, and then this person could always play the game of revising the constraints or adding further operators.) Rather the appropriate response to this person would be that she is seeking to reinvent the wheel, and moreover that she has started off by sketching a square.

## 4.2 Rivals: Rosen

Much of what I have to say against Fine's regimentation carries over to Rosen's regimentation. Rosen (2010, p. 115; see also Audi 2012b, p. 103; Raven 2012, p. 689) offers a primitive relation between facts, regimented as:

*Rosen*:  $[p] \leftarrow \Gamma$

Rosen's ' $\leftarrow$ ' relates a single fact  $[p]$  (read as: "the fact that  $p$ ") to a plurality of grounding facts  $\Gamma$  (read as: "the facts gamma"). So—applied to Fine's conjunctive property example (Sect. 4.1)—*Rosen* delivers:

[The ball is red and round]  $\leftarrow$  [The ball is red], [The ball is round]

Again I think it should be apparent that *Rosen* is at best a first-pass regimentation of explanatory patterns, and has no plausibility as a regimentation of the causal relations that back these explanatory patterns. No one would think that a metaphysically perspicuous regimentation of causation looks like:

[There was a fire]  $\leftarrow_{\text{cause}}$  [The ground was dry], [The lightning struck]

The move from a sentential operator (as in *Fine*: Sect. 4.1) to a relation is an improvement, but facts are still not the right relata for causation (or so I am assuming here, on orthodox lines)—facts are far more plausibly wheeled in at the level of explanation. Contrasts are still missing, as are the core conceptual backdrop for causation, includes situating conditions within contrast spaces, and encoding

functions that articulate the form of the dependence pattern. So I have three criticisms of *Rosen* as a regimentation of grounding claims:

- it regiments grounding as a relation between facts, which is at best apt for the image of explanation not the image of causation,
- it regiments grounding in a noncontrastive format, and
- most crucially, it does not include the core conceptual background structure found in structural equation models, including (a) contrast spaces for the conditions under study, and (b) functions codifying the shape of the dependence patterns between conditions.

Thus grounding as *Rosen* conceives of it is nothing like metaphysical causation. This means that *Rosen* cannot explain the analogy with causation or unify the input to explanation. This is not to offer any specific counterexample to *Rosen*, but rather to argue that he—like *Fine*—has started off with the wrong form.

### 4.3 Against reconciliation

I think that *Fine* and *Rosen* are at best apt as first-pass regimentations of explanatory patterns, and inappropriate for treating grounding in the image of causation, as the concrete relation that backs the explanatory pattern (Sects. 4.1–4.2). But one possibility is that I have a merely verbal disagreement with proponents of anything like the *Fine* and *Rosen* regimentations. It may be that *Fine* and *Rosen* all along only meant ‘grounding’ to stand for metaphysical explanation.<sup>39</sup> For instance, suppose that one accepts the following schema, on the understanding that ‘grounding’ will feature somewhere in the blanks:

*Analogy schema:* Causation is to causal explanation as \_\_\_\_\_ is to \_\_\_\_\_.

Then one might think:

*Reconciliation:* Causation is to causal explanation as Schaffer-grounding is to Fine-grounding.

(One could equally try to plug ‘*Rosen-grounding*’ into the second blank).

I should not be too upset if *Reconciliation* proves right, insofar as it preserves a crucial role for *Schaffer-grounding* as the analogue of causation, and also gives one a strong basis for understanding metaphysical explanation through the rigorous formalism of *Fine* (2012). From that perspective *Reconciliation* might seem like the best of both worlds.

<sup>39</sup> Indeed *Litland* (2013, p. 19) explicitly says “grounding corresponds to (metaphysical) ‘explanation how’,” and *Dasgupta* (2014, p. 3) says: “As I 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’.” Though *Rodriguez-Pereyra* (2005, p. 28) explicitly rejects this idea: “Explanation is not and does not account for grounding—on the contrary, grounding is what makes possible and ‘grounds’ explanation.” Likewise *Audi* (*Audi* 2012a, pp. 687–688; cf. *Trogon* 2013b) argues for grounding by saying that “explanations require nonexplanatory relations underlying their correctness” and so inferring that there is “a noncausal relation at work in these explanations.”

Two questions arise. First, am I misreading Fine, Rosen, and others? Have they all along been offering a theory of metaphysical explanation, as distinguished from but thought to be backed by some relation theorized as “metaphysical causation”? Secondly, leaving aside what Fine, Rosen, and others have actually thought, is it intellectually possible to reconcile their ideas with mine in this way?

The first question probably holds little interest to outside parties, so I will simply assert that these notions are not distinguished in the current literature, but that one finds causal and explanatory notions used as if equivalent.<sup>40</sup> Also there is no discussion of what relation might underwrite grounding if conceived of as “metaphysical explanation.” This absence is further evidence that grounding has been conceived in a way that blurs the image of causation and the image of explanation.<sup>41</sup>

The second question, as to whether it is at least intellectually possible to hold *Reconciliation*, deserves more interest. Indeed I think that *Reconciliation* is an intellectually possible view. But still I advise against it, for two reasons. The first reason is that I think (following van Fraassen 1980 and Garfinkel 1981) that explanation is itself a contrastive notion. Hence the schema ‘*T, U, V, ... <S*’ is at best a first-pass regimentation of explanatory patterns, and ought to give way to the contrastive:

*S* rather than *S\** because *T* rather than *T\**, *U* rather than *U\**, ...

This point is connected to the previous point that structural equation models specify contrasts for each condition (Sect. 2.2), while neither *Fine* nor *Rosen* does so (Sects. 4.1–4.2). As such I do not actually think that the regimentations given in *Fine* or *Rosen* correspond in the end to anything whatsoever. (At best I would allow that their natural language glosses correspond to *explanatory sentences*, which can be true or false relative to a context, where the context determines the contrasts.)

Virtually everyone agrees that preference is contrastive, simply because one might prefer vanilla to strawberry, but not prefer vanilla to chocolate. (So ‘I prefer

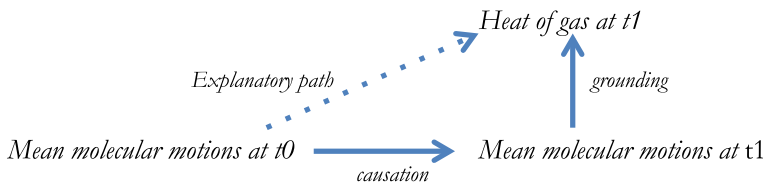
<sup>40</sup> The curious reader may look at Fine (2012, p. 40), Correia (2005, p. 48), Correia and Schnieder (2012, p. 22), Rosen (2010), and Audi (2012b, p. 104 and p. 106) for some examples. I also have in mind all of the many cases where grounding is glossed with both causative terms such as ‘making’ and explanatory connectives such as ‘because.’

<sup>41</sup> Fine (1995; cf. Koslicki 2012) offers a relation of *ontological dependence* between entities understood in terms of *essence*, on which one entity *x* depends upon some others *y1, y2, ...* if and only if *y1, y2, ...* feature in *x*’s constitutive essence (that is, *y1, y2, ...* show up in the “real definition” of *x*). So it might be thought that ontological dependence is the relation backing grounding-as-metaphysical-explanation. But in fact neither Fine nor Rosen go in for anything like this. Rosen (2010, pp. 131–133) instead discusses the idea (“Mediation”) that general grounding principles are to be explained by essences. Rosen is skeptical, and in any case the idea is of the wrong form to provide an analogue of causation: the idea instead concerns whether essences may be playing a role in backing general grounding principles (the laws of metaphysics), not playing a role akin to causation. Fine (2012, pp. 74–80) also discusses the connection between essence and ground, and instead of viewing the former as backing the latter, sees (2012, p. 79) instead “two fundamentally different forms of explanation” and warns (2012, p. 80) of an error “writ large over the whole metaphysical landscape” of “attempting to assimilate or unify the concepts of essence and ground.” Whatever exactly dependence and essence are doing for Rosen and Fine, they are clearly *not* playing the role of “Schaffer-grounding” in backing metaphysical explanation, nor do I see any other concept specified in either of their discussions which might play this role.



vanilla’, which is a grammatical sentence of English, needs some form of contextual supplementation by an intended contrast to get assigned a truth value.) Yet exactly the same argument pattern extends to explanation, twice over. The explanation for why Sutton *robbed* banks may be quite different from the explanation for why Sutton robbed *banks*.<sup>42</sup> But even holding fixed the question as asking why Sutton robbed banks as opposed to following the law, an answer involving his father’s *abuse* of him may be quite different from an answer involving his *father’s* abuse of him. It may or may not have mattered who treated Sutton so abusively. So one might have an explanation of why Sutton robbed banks rather than following the law in terms of his father’s abuse rather than loving treatment. But one might not have an explanation of why Sutton robbed banks rather than gas stations in those terms, and one might not have an explanation of why Sutton robbed banks rather than following the law in terms of his father’s abuse rather than his mother’s abuse (perhaps it made no difference which parent abused him, but only mattered that he was abused).

My second reason for advising against *Reconciliation* is that I also think—in ways that conflict even with *Analogy schema*—that there is no distinctive notion of causal (as opposed to metaphysical) explanation. I think that there is just *explanation (simpliciter)*, and that explanation can be backed by causation, grounding, and whatever other relations belong to the elite handful of relations that can back explanation. My main argument for *the unity of explanation* is that there are *hybrid explanations* involving an element of causation and an element of grounding (and mathematical elements as well). If there were just causal explanation and metaphysical explanation, these hybrid explanations would count as neither fish nor fowl. As an example of a hybrid explanation, consider the “diagonal” explanation of why a given gas has a certain heat at  $t1$ , in terms of the molecular motions of its molecules at  $t0$ :

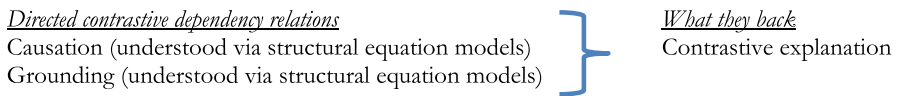


This seems to me to be a perfectly legitimate style of explanation. It is not purely causal because it crosses levels, and it is not purely metaphysical because it crosses times (hence the label “diagonal”). It also presumably involves various mathematical elements as well (drawing on statistical mechanics). It is a legitimate explanation that is backed by a combination of causal, metaphysical, and mathematical factors.

The existence of hybrid explanations shows that there cannot just be causal and metaphysical explanation, but leaves open whether there is just one sort of

<sup>42</sup> *Backstory*: The armed robber Willie Sutton, asked by his prison chaplain why he robbed banks, famously replied ‘That’s where the money is’ (from Garfinkel 1981, p. 22).

explanation (as I think) or instead three sorts of explanation: causal, metaphysical, and hybrid. But saying that there are three sorts of explanation is unappealing for two reasons. First, this sort of move is going to lead to a kind of explosion of types of explanation when one considers mathematical, analytical, and other “sorts of explanation” as well. For now it seems like every plurality of such sorts should hybridize, so instead of having  $n$  sorts of explanation one will wind up with  $2^n - 1$  sorts of explanation. This looks ugly. Secondly and relatedly, these  $n$  or  $2^n - 1$  sorts of explanation are going to be deeply unified. Pending the identification of something like a formal-structural difference between these sorts of explanation, this starts to look like there is a single unified connection of explanation called ‘causal’ when backed by causation, ‘metaphysical’ when backed by grounding, ‘causal-metaphysical hybrid’ when backed by causation and grounding, and so on. So, my recommendation is that the best overall way to conceive of the domain is:



As such I recognize no distinctive notion of metaphysical explanation. Hence my view about the semantic value of ‘because’ (unconsidered in the options mentioned by Correia and Schnieder 2012, pp. 22–24) is that it can univocally denote *explanation*.<sup>43</sup>

Putting this together, I am saying that one should recognize various backing relations, alongside a general and univocal notion of explanation these back. The various backing relations including causation and grounding, treated as contrastive directed dependency relations and understood via structural equation models. The general notion of explanation is likewise contrastive. Accordingly I must advise against *Reconciliation*, and moreover must advise against the regimentations in *Fine* and *Rosen*. These regimentations do not fit the image of causation (that’s Schaffer-grounding) or the image of explanation (explanation on my view is already a general notion, and the image of explanation is just explanation again). Insofar as one of the main points of introducing grounding is to understand explanations in metaphysics (Sect. 1.4), I can only conclude that Fine-grounding and Rosen-grounding—for all their admirable rigor and elegance—are grounding in the image of nothing.

#### 4.4 For intelligibility and unity

With the explosion of interest in grounding and fundamentality has come (predictably enough) a skeptical reaction from certain corners. I think a structural equation based treatment of grounding not only helps reveal what rival treatments of grounding have missed, but also helps speak to the grounding skeptics. Essentially I

<sup>43</sup> A second argument for the unity of explanation: there is a sense of ‘because’ which univocally denotes it.

am saying that extending structural equation models from causation to grounding shows grounding to be just as intelligible and unified as causation.

To begin with I distinguish two very different sorts of grounding skeptics: *nihilists* and *pluralists*. The nihilist denies that there is any such notion as grounding. She finds the very idea unintelligible (Hofweber 2009; Daly 2012).<sup>44</sup> On the other flank, the pluralist affirms that there are many intelligible notions of grounding. She considers the idea disunified (Wilson 2014;<sup>45</sup> Koslicki forthcoming). Accordingly I am defending a form of *grounding monism*, which might—as a first pass—be situated against the nihilist and pluralist alternatives as per:

*Grounding nihilism*: there is no intelligible notion of grounding,

*Grounding monism*: there is exactly one intelligible notion of grounding, and

*Grounding pluralism*: there are many intelligible notions of grounding.

This characterization of *Grounding monism* is only a first pass. It is overly strong, insufficiently detailed, and not clearly evaluable. It is overly strong since virtually every friend of grounding draws further distinctions. Indeed I have already (Sect. 1.3) distinguished between notions of type grounding and token grounding, component and net grounding, as well as incomplete, complete, and total grounding.<sup>46</sup> It is insufficiently detailed because it leaves open Koslicki's (forthcoming) distinction between cases in which there is merely a single species notion, and cases where there is a single genus notion with many species. And it is not clearly evaluable until one considers when in general it is appropriate to consider a notion unified and when disunified. I am not sure that I can do better to convey the distinction between the grounding monist and her nihilist and pluralist foes on either flank, but in the end sorting philosophers into camps is less important than dealing with the underlying issue of the legitimacy of operating with the notion of grounding, and the general questions lurking of when it is appropriate to consider a notion intelligible, and when

<sup>44</sup> Prior to the “grounding revolution”, grounding nihilism was the norm. For instance, Thomson (1983, p. 211) decries both ontological and epistemological priority as “dark notions,” though she does immediately allow that “we have *some* grip on what [these notions] are.” Lewis (1983, p. 358) advertises supervenience relations as providing “a stripped-down form of reductionism, unencumbered by dubious denials of existence, claims of ontological priority, or claims of translatability.” And Oliver (1996, p. 48) declares that “we know that we are in the realm of murky metaphysics by the presence of the weasel words ‘in virtue of,’” going on (1996, p. 69) to proclaim: “‘In virtue of’ really ought to be banned.”

<sup>45</sup> Wilson might better be considered a grounding nihilist, though she does (2014, pp. 561–3) wind up invoking a hyperintensional primitive notion of (absolute) fundamentality, and so in that respect she is very close to grounding theorists, who can be understood as invoking a hyperintensional primitive of relative fundamentality. Her pluralism comes in her thinking that there are many relative notions (many “grounding with a small-‘g’” notions) connected to her one primitive absolute notion (“Fundamentality” with a capital ‘F’). It is puzzling to me why Wilson regards grounding as disunified, but regards her “Fundamentality” and her “small-‘g’” relations including causation and composition as unified. It would be good to have general criteria for unity.

<sup>46</sup> Fine (2012, pp. 48–54) offers a different range of distinctions, including factive and nonfactive grounding, direct and indirect grounding, and distributive and collective grounding. Layered over these distinctions, Fine (2012, p. 77) offers further distinctions between “normative and natural conceptions of ground, which are to be distinguished from the purely metaphysical conception.” So I take it that Fine’s various notions of ground are triplicated, with parallel but distinct conceptual structures to be found in the metaphysical, normative, and natural domains.

it is appropriate to consider a notion unified. So I propose to skip directly to the issues of intelligibility and unity.

I begin with the matter of intelligibility. Hofweber (2009, pp. 266–267) distinguishes “egalitarian metaphysics” which is “expressed in ordinary, everyday terms” from “esoteric metaphysics” which requires “distinctly metaphysical terms” such that “You have to be an insider to get in the door.” He (2009, p. 268) then adds: “The most common way to be an esoteric metaphysician in practice is... [to] rely on a notion of metaphysical priority: some notion that claims that certain facts or things are metaphysically more basic than other facts or things.” Daly (2012, p. 89)—following Goodman—says that, in the absence of a general condition for intelligibility, one can only consult one’s philosophical conscience over cases.<sup>47</sup> He reports that grounding troubles his philosophical conscience. The paradigm cases do not assuage him because (2012, p. 95): “Anyone who does not understand ‘grounding’ will not understand any example that uses it.”<sup>48</sup>

Perhaps this is ultimately a matter on which reasonable people may disagree, but I do not find grounding to be esoteric, nor is my conscience panged. Grounding-style notions appear at the dawn of philosophy, for instance in Plato’s *Euthyphro* (Sect. 1.1). I teach this dialogue to first-year Rutgers undergraduates. They get it. And grounding is a notion that is extremely natural in the sciences, in considering the relation between levels. One need not be versed in an arcane metaphysics to think that the chemical depends on the physical.<sup>49</sup> Philosophers such as Fine, Rosen, and myself have not conspired to invent a new code and secret handshake for the initiated few. Indeed philosophers through millennia, with widely varying ideological scruples, have repeatedly found themselves drawn to ‘in virtue of’ locutions and ‘dependence’ talk. These are natural notions, which metaphysicians should make sense of if they can.

At any rate, I continue to think that, in order to communicate a concept it suffices in general to provide paradigm cases, offer useful analogies, and enfold the concept in a formalism (Sect. 1.1). I have tried to do just that. The denier of intelligibility

<sup>47</sup> This seems to me like poor methodology, and an invitation to let unsystematic hunches and pure prejudices run amok. Certainly there are cautionary tales in the vicinity. Goodman himself invoked his philosophical conscience in refusing to countenance sets (fortunately the mathematicians did not heed his counsel), his colleague Quine refused to countenance modality (fortunately the logicians were not moved), and Russell refused to countenance causation (fortunately the scientists did not blink). I agree with Sider (2011, p. 9), who in defending the intelligibility of his primitive concept of structure as follows: “The perceived magical grasp of more familiar concepts like modality, in-virtue-of, or law of nature, is due solely to the fact that we’ve become accustomed to talking about them.” (I also agree with Sider in placing in-virtue-of on the side of the familiar already grasped notions.)

<sup>48</sup> As Lewis (1986b, p. 205) quipped: “[A]ny competent philosopher who does not understand something will take care not to understand anything else whereby it might be explained.”

<sup>49</sup> It remains open to the grounding nihilist (cf. Hofweber 2009, pp. 269–272) to agree that the chemical depends on the physical but maintain that the dependence at issue is to be understood in some non-metaphysical sense. But I see little pressure to demand some alternative sense, and little prospect for finding one. The priority of physics over chemistry is not merely a conceptual or logical matter. There is counterfactual dependence but that is a sign of metaphysical dependence (Sect. 3.2), not a substitute for it. Indeed when one has asymmetric counterfactual dependence, one wants some underlying asymmetric relation (causation, grounding) by means of which to distinguish right-tracking from wrong-tracking (back-tracking in time, down-tracking in levels).

may take this as an invitation to show where I have failed, or to explain what more is generally needed to render a concept intelligible and soothe a troubled mind.

Turning to the matter of unity, I think that the ultimate test of unity is whether the concept can be successfully enfolded in a formalism in a unified way (Rosen 2010, p. 114). Consider—as a foil—the philosopher who doubts whether there is a unified notion of identity because the notion applies to such diverse entities as tables, numbers, and weddings. I take it that few will be moved by her call to “identity pluralism”. But why? It seems to me that the reason why identity monism is plausible is that the concept of identity can be enfolded in a formalism and equipped with distinctive axioms (e.g. Leibniz’s Law of the Indiscernibility of Identicals) *which do not care whether the terms flanking it are for tables, numbers, weddings, or whatnot*. If identity really were disunified, one would reasonably expect that the formalism would need to draw distinctions between notions like “identity<sub>concrete object</sub>,” “identity<sub>abstract object</sub>,” and “identity<sub>event</sub>,” etc., and outfit these distinct notions with distinctive axioms. That this is unneeded is defeasible evidence for unity.

Or consider the real case of the philosopher, such as Cartwright (2007), who doubts that there is a unified notion of causation. What might be said to decide the matter? It seems to me that causal monism is plausible insofar as the concept of causation can be enfolded in a formalism and equipped with distinctive axioms *which do not care about the type of causation*. If causation really were disunified, one would reasonably expect that the formalism would need to draw distinctions e.g. between baking, raking, and waking. Structural equation models do not need special colored arrows for different types of causal relations.<sup>50</sup>

If I am right that structural equations may be used to model grounding, then *grounding has exactly the same ultimate claim to intelligibility and unity as causation*. It is legitimate to operate with the notion of causation, and the ultimate proof of this is the existence of a fruitful framework for operating with the notion in a unified way. Likewise I say that it is equally legitimate to operate with the notion of grounding. Though at this point I wish to offer another opening to the grounding skeptic: defend causal nihilism or pluralism. (It is a good-making feature of a framework if it clarifies the opportunities for skeptical reaction, and especially if it then reveals fruitful connections to further issues.)

Putting this together, I think that the structural equation models approach can:

- reveal shortcomings in current treatments of grounding (Sects. 4.1–4.3), and
- show grounding to be an intelligible and unified posit (Sect. 4.4).

The way of showing grounding to be intelligible was just the way of communicating a concept, including enfolding it within a formalism (Sect. 1.1), and the way of showing grounding to be unified was just pointing out that it plays a unified role within the formalism. The denier of unity may take this as an invitation to show where I have failed, or to explain what more is generally needed to reveal a concept to be unified.

<sup>50</sup> In this vein Pearl (2010, p. 72) challenges pluralists like Cartwright “to cite a single example” that does not fit his unitary structural equations formalism.

#### 4.5 Is there a real difference between grounding and causation?

By way of conclusion I want to return to the relationship between grounding and causation. It might seem, after all, as if I am now committed to a further—and quite disquieting—unity claim:

*Grounding-causation unity:* Grounding and causation are only nominally distinct relations.

*Grounding-causation unity* is the claim that there is a single unified relation of generation called ‘causation’ when it drives the world through time, and ‘grounding’ when it drives the world up levels (Sect. 1.4). The reason why I might seem committed to this is as follows: I say that grounding and causation can and should be treated via a common formalism (Sects. 2–3), and I say that the ultimate test of unity is through enfolding in a common formalism (Sect. 4.4). So it looks like *Grounding-causation unity* should hold by my own ultimate test.

I agree with Bennett (2011a, pp. 93–94) that *Grounding-causation unity* is a fascinating thesis, not to be dismissed out of hand. But that said, I am not actually committed to the thesis, and indeed I am inclined to think that it is false for at least three reasons.<sup>51</sup> Firstly, there is the following difference: grounding implies an associated (metaphysical) supervenience, causation does not imply an associated (nomological) supervenience. This is because there can be indeterministic causation but not indeterministic grounding. (In indeterministic cases, fixing the causes does not fix the effect.) And this difference shows up formally in indeterministic extensions of the structural equations framework, which a full treatment of causation ultimately needs. So the appearance of a commitment to *Grounding-causation unity* is really just an artifact of my restriction to the deterministic case for causation (Sect. 2.2). Lift that restriction and the formalisms separate. (Though the formalism for grounding is still a special case of the formalism for causation, so this is not a decisive argument against the unity claim either.)

Secondly, there is the difference that causation connects distinct events but grounding connects indistinct entities. Causation is thus an external relation while

<sup>51</sup> One reason for rejecting *Grounding-causation unity* which I do *not* accept is the idea that there is a difference in modal strength between grounding as metaphysical necessity, and causation as connected to mere nomological necessity (*pace* Rosen 2010, p. 118). I do agree that there is a distinction between nomological and metaphysical necessity (*contra* certain kinds of “necessitarians” about laws such as Wilson 2013), but regard the distinction as somewhat arbitrary (these are just some among many equally legitimate restrictions one can place on a modal quantifier), and as cross-cutting (neither is to be regarded as a restriction of the other). Though it is commonly thought that nomological necessity is a restriction of metaphysical necessity, I think that there are nomologically possible but metaphysically impossible scenarios (such that the distinctions cross-cut). For instance, suppose that our world is a trope world, and is governed by the laws of quantum mechanics. The laws of quantum mechanics (e.g. Schrödinger’s equation) invoke properties but do not say anything about the metaphysical basis for those properties (e.g. whether they be universals, resemblance classes of tropes, or mere shadows of predicates, etc.) So a world with universals might also be governed by the very same laws of quantum mechanics. Given that our world is a trope world, and that the actual nature of properties is a matter holding with metaphysical necessity, a universals world governed by the laws of quantum mechanics is then nomologically possible but metaphysically impossible.

grounding is an internal relation. Indeed grounding is what Bennett (2011b, p. 32) calls a *super-internal* relation: fixing the intrinsic nature of the grounding side of the relation alone guarantees that the grounded side exists, has the intrinsic nature that it does, and is grounded in that way. On this point I agree with Rosen (2010) who finds it “more natural to keep causal relations on one side—as external relations among wholly distinct states of affairs—and grounding relations on the other.” This difference shows up in the aptness requirements which must accompany the formalism, particularly with respect to the distinctness/indistinctness of the relata (Sect. 3.3). (Though this is hardly a decisive argument either: one might still be looking at the same relation, just applied to distinct and indistinct portions of reality.)

Thirdly, there is (by my lights at least) a crucial structural difference: grounding needs to be well-founded, causation does not. Grounding must be well-founded because a grounded entity inherits its reality from its grounds, and where there is inheritance there must be a source.<sup>52</sup> One cannot be rich merely by having a limitless sequence of debtors, each borrowing from the one before. There must actually be a *source* of money somewhere. Likewise something cannot be real merely by having a limitless sequence of ancestors, each claiming reality from its parents. There must actually be a source of reality somewhere. Just as wealth endlessly borrowed is never achieved, so reality endlessly dependent is never realized. But a caused entity *qua* caused entity still has intrinsic reality unto itself. Caused entities do not inherit their reality from their causes. Indeed, a caused entity may also be fundamental, and thus ontologically subsistent in its own right.

There is a general question lurking of which structures require foundations. Many (including myself) are *foundationalists* about epistemic justification, thinking that all justification must originate in basic justification, and rejecting the possibility of circular or limitlessly descending chains of justification. Yet many (again including myself) are *non-foundationalists* about causal and temporal structure, allowing for limitlessly backwards causal and temporal sequences.<sup>53</sup> This general question of which structures require foundations is connected to the question of when a regress counts as *vicious* and when *benign*. Following Maurin (2013, p. 432; cf. Johnsson 2009), I think that a regress counts as vicious if and only if there is an endless chain of dependency with *transference* of the relevant status. As Bliss (2013, pp. 405–406)—though herself an opponent of the view—aptly puts the point: “[V]icious infinite regresses involve an underlying dependence relation. What is significant about this dependence relation is that it involves the transfer of a certain significant property...” It is the “transfer model” that leads to the need for a *source*. With grounding one is

<sup>52</sup> As Leibniz wrote to de Volder: “[W]here there is no reality that is not borrowed, there will never be any reality, since it must belong ultimately to some subject” (quoted in Adams 1994, p. 335; cf. Aristotle 1984a, p. 5; Fine 1991, p. 267).

<sup>53</sup> A *curious disanalogy*: With respect to causal and temporal structure, most allow that the structure may have a starting point (e.g. the Big Bang) but also may run limitlessly backwards. But with respect to epistemic structure, virtually no one permits both sorts of structure. Virtually all of the views on the epistemic side of the ledger *require* justificatory structure to be well-founded, or to be circular, or to be infinitary.



looking at a transference of reality: the grounded entity exists *in virtue of* its grounds. This is why a source of reality is needed, in order for there to be anything to transfer. With causation one is not looking at a transference of reality. As before, the effect still has intrinsic reality unto itself, and may indeed be ontologically subsistent in its own right. This is why no first cause is needed.

The difference over well-foundedness must show up as an additional formal restriction on structural equation models when employed for modeling grounding. So the appearance of a commitment to *Grounding-causation unity* is also an artifact of my restriction to models with finitely many variables (Sect. 2.2), since with finitely many variables one cannot form an infinite much less a limitless sequence. Allow infinitely many variables and the prospect of non-well-founded models comes into view, at which point the formalisms must separate. An extra axiom is then needed for grounding.

These final two points of difference concerning distinctness and well-foundedness may be two sides of one coin, or so I speculate. Within each distinct portion of reality, one must find an internal source of its reality (this is the required well-foundedness of grounding). But across distinct portions of reality, one is dealing with metaphysically independent tiles of the cosmic mosaic, and each tile is there from the start without needing a source (this is the permitted non-well-foundedness of causation).

Putting all this together, I take the grounding-causation analogy to encompass the following points:

- both are generative relations;
- both are partial orders, admitting a type/token distinction, a component/net distinction, an incomplete/complete/total distinction, and screening-off relations;
- both are backed by non-accidental generalizations, delimit a specific form of necessity, are supportive of and diagnosable by counterfactuals, and can back explanation; and
- both can be fruitfully formalized via structural equation models.

But I also take there to be at least the following points of disanalogy:

- causation can be non-deterministic, grounding must be deterministic;
- causation can only connect distinct (grounding-disconnected) portions of reality; and
- causation can be non-well-founded, grounding must be well-founded.

So I am inclined to think that causation and grounding are distinct notions. Though note that this is consistent with also accepting the claim that there is a unified general notion of directed determination (connected to explanation-backing), of which causation and grounding are distinct species. The point is just that they would be distinct species.

This concludes my attempt to take the grounding-causation analogy seriously. I offer a treatment of grounding in the image of causation, on the template of structural equation models. I end on a note of guarded optimism: the treatment seems viable and useful. But regardless of the ultimate fate of the project, I hope that the journey itself holds interest, yielding a deeper understanding of the guiding idea of grounding as “metaphysical causation”.



**Acknowledgments** This paper supersedes earlier work appearing in manuscript under the titles of “Grounding as the Primitive Concept of Metaphysical Structure” and “Structural Equation Models of Ground.” My thanks especially to Karen Bennett, David Chalmers, Fabrice Correia, Shamik Dasgupta, Louis DeRosset, Janelle Derstine, Kit Fine, Ned Hall, Christopher Hitchcock, Thomas Hofweber, Thomas Kivatinos, Kathrin Koslicki, Lisa Miracchi, L. A. Paul, Michael Raven, Gideon Rosen, Raul Saucedo, Benjamin Schnieder, Theodore Sider, Alex Skiles, Kelly Trogdon, Tobias Wilsch, Alastair Wilson, Jessica Wilson, and audiences at the Australian National University, Birmingham, Bristol, Fordham, Geneva (Eidos), Manchester, Notre Dame, Princeton, Vermont, Washington University, the Colorado Conference on Dependence, Metaphysical Mayhem at Rutgers, the Epistemology of Philosophy Conference in Cologne, the Central APA, the Metaphysics Reading Group at Rutgers, and the Oberlin Colloquium in Philosophy.

## References

- Adams, R. (1994). *Leibniz: Determinist, theist, idealist*. Oxford: Oxford University Press.
- Anscombe, G. E. M. (1975). Causality and determinism. In E. Sosa (Ed.), *Causation and conditionals* (pp. 63–81). Oxford: Oxford University Press.
- Aristotle. (1984a). Categories. In J. Barnes (Ed.), *The complete works of Aristotle* (Vol. 1, pp. 3–24). Princeton: Princeton University Press.
- Aristotle. (1984b). Metaphysics. In J. Barnes (Ed.), *The complete works of Aristotle* (Vol. 2, pp. 1552–1728). Princeton: Princeton University Press.
- Armstrong, D. M. (1975). Towards a theory of properties: Work in progress on the problem of universals. *Philosophy*, 50, 145–155.
- Armstrong, D. M. (1997). *A world of states of affairs*. Cambridge: Cambridge University Press.
- Audi, P. (2012a). Grounding: Toward a theory of the in-virtue-of relation. *The Journal of Philosophy*, 109, 685–711.
- Audi, P. (2012b). A clarification and defense of the notion of grounding. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 101–121). Cambridge: Cambridge University Press.
- Baumgartner, M. (2009). Uncovering deterministic causal structures: A deterministic approach. *Synthese*, 170, 71–96.
- Bennett, J. (1988). *Events and their names*. Indiana: Hackett.
- Bennett, K. (2011a). Construction area (no hard hat required). *Philosophical Studies*, 154, 79–104.
- Bennett, K. (2011b). By our bootstraps. *Philosophical Perspectives*, 25, 27–41.
- Blanchard, T., & Schaffer, J. (forthcoming). Cause without default. In H. Beebe, C. Hitchcock, & H. Price (Eds.), *Making a difference*. Oxford: Oxford University Press.
- Bliss, R. L. (2013). Viciousness and the structure of reality. *Philosophical Studies*, 166, 399–418.
- Bradley, F. H. (1978). *Appearance and Reality*. Oxford: Clarendon Press.
- Bricker, P. (2006). The relation between general and particular: Entailment vs. supervenience. *Oxford Studies in Metaphysics*, 3, 251–287.
- Briggs, R. (2012). Interventionist counterfactuals. *Philosophical Studies*, 160, 139–166.
- Carroll, J. (1994). *Laws of nature*. Cambridge: Cambridge University Press.
- Cartwright, N. (2007). *Hunting cause and using them: Approaches in philosophy and economics*. Cambridge: Cambridge University Press.
- Correia, F. (2005). *Existential dependence and cognate notions*. Munich: Philosophia Verlag.
- Correia, F., & Schnieder, B. (2012). Grounding: An opinionated introduction. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 1–36). Cambridge: Cambridge University Press.
- Daly, C. (2012). Scepticism about Grounding. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 81–100). Cambridge: Cambridge University Press.
- Dasgupta, S. (2014). On the plurality of grounds. *Philosophers' Imprint*, 20, 1–28.
- deRosset, L. (2013). Grounding explanations. *Philosophers' Imprint*, 13, 1–26.

- Descartes, R. (1985). In J. Cottingham, R. Stoothoff, & D. Murdoch (trans. and Eds.), *The philosophical writings of descartes*. Cambridge: Cambridge University Press.
- Fine, K. (1991). The study of ontology. *Noûs*, 25, 263–294.
- Fine, K. (1995). Ontological dependence. *Proceedings of the Aristotelian Society*, 95, 269–290.
- Fine, K. (2001). The question of realism. *Philosophers' Imprint*, 1, 1–30.
- 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 & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 37–80). Cambridge: Cambridge University Press.
- Garfinkel, A. (1981). *Forms of explanation: Rethinking the questions in social theory*. New Haven: Yale University Press.
- Glymour, C. (2007). Learning the structure of deterministic systems. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 231–240). New York: Oxford University Press.
- Hall, N. (2000). Causation and the price of transitivity. *The Journal of Philosophy*, 97, 198–222.
- Hall, N. (2007). Structural equations and causation. *Philosophical Studies*, 132, 109–136.
- Halpern, J. (2000). Axiomatizing causal reasoning. *Journal of Artificial Intelligence Research*, 12, 317–337.
- Halpern, J. (2008). Defaults and normality in causal structures. In G. Brewka & J. Lang (Eds.), *Principles of knowledge representation and reasoning: Proceedings of the eleventh international congress* (pp. 198–208). Menlo Park, CA: AAAI Press.
- Halpern, J., & Hitchcock, C. (2010). Actual causation and the art of modeling. In R. Dechter, H. Geffner, & J. Halpern (Eds.), *Heuristics, probability, and causality: A tribute to judea pearl* (pp. 383–406). London: College Publications.
- Halpern, J., & Pearl, J. (2005). Causes and explanations: A structural-model approach. Part I: Causes. *British Journal for the Philosophy of Science*, 56, 843–887.
- Handfield, T., Twardy, C., Korb, K., & Oppy, G. (2008). The metaphysics of causal models: Where's the biff? *Erkenntnis*, 68, 149–168.
- Hitchcock, C. (1996). The role of contrast in causal and explanatory claims. *Synthese*, 107, 95–419.
- Hitchcock, C. (2001). The intransitivity of causation revealed in equations and graphs. *The Journal of Philosophy*, 98, 273–299.
- Hitchcock, C. (2007a). What's wrong with neuron diagrams? In J. K. Campbell, M. O'Rourke, & H. Silverstein (Eds.), *Causation and explanation* (pp. 69–92). Cambridge, MA: The MIT Press.
- Hitchcock, C. (2007b). Prevention, preemption, and the principle of sufficient reason. *The Philosophical Review*, 116, 495–532.
- Hofweber, T. (2009). Ambitious, yet modest, metaphysics. In D. Chalmers, D. Manley, & R. Wasserman (Eds.), *Metametaphysics: New essays on the foundations of ontology* (pp. 260–289). Oxford: Oxford University Press.
- Hume, D. (1978). In P. H. Nidditch (Ed.), *A treatise of human nature*. Oxford: Oxford University Press.
- Johnsson, I. (2009). Proof of the existence of universals—and Roman Ingarden's ontology. *Metaphysica*, 10, 65–87.
- Jöreskog, K. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85–112). New York: Seminar Press.
- Koslicki, K. (2012). Varieties of ontological dependence. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 186–213). Cambridge: Cambridge University Press.
- Koslicki, K. (forthcoming). The coarse-grainedness of grounding. *Oxford Studies in Metaphysics*.
- Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy*, 61, 343–377.
- Lewis, D. (1986a). Causation. In D. Lewis (Ed.), *Philosophical papers* (Vol. 2, pp. 159–213). Oxford: Oxford University Press.
- Lewis, D. (1986b). *On the plurality of worlds*. Oxford: Basil Blackwell.
- Lewis, D. (2001). Truthmaking and difference-making. *Noûs*, 35, 602–615.
- Litland, J. (2013). On some counterexamples to the transitivity of grounding. *Essays in Philosophy*, 14, 19–32.
- Loewer, B. (2001). From physics to physicalism. In C. Gillett & B. Loewer (Eds.), *Physicalism and its Discontents* (pp. 37–56). Cambridge: Cambridge University Press.
- Maslen, C. (2004). Causes, contrasts, and the nontransitivity of causation. In J. Collins, N. Hall, & L. A. Paul (Eds.), *Causation and counterfactuals* (pp. 341–357). Cambridge, MA: The MIT Press.

- Maudlin, T. (2007). The whole ball of wax. In T. Maudlin (Ed.), *The metaphysics within physics* (pp. 170–183). Oxford: Oxford University Press.
- Maurin, A. S. (2013). Infinite regress arguments. In C. Svennerlind, J. Almäng, & R. Ingthorsson (Eds.), *Johanssonian investigations* (pp. 421–437). Frankfurt: Ontos Verlag.
- McDermott, M. (1995). Redundant causation. *British Journal for the Philosophy of Science*, 40, 523–544.
- Menzies, P. (2007). In H. Price & R. Corry (Eds.), *Causation in context. Causation, physics and the constitution of reality: Russell's republic revisited* (pp. 191–223). Oxford: Oxford University Press.
- Menzies, P. (2009). Platitudes and counterexamples. In H. Beebe, C. Hitchcock, & P. Menzies (Eds.), *The Oxford handbook of causation* (pp. 341–367). Oxford: Oxford University Press.
- Nolan, D. (1997). Impossible worlds: A modest approach. *Notre Dame Journal of Formal Logic*, 38, 535–572.
- Northcott, R. (2008). Causation and contrast classes. *Philosophical Studies*, 139, 111–123.
- Oliver, A. (1996). The metaphysics of properties. *Mind*, 105, 1–80.
- Paul, L. A., & Hall, N. (2013). *Causation: A user's guide*. Oxford: Oxford University Press.
- Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.
- Pearl, J. (2010). Nancy Cartwright on hunting causes. *Economics and Philosophy*, 26, 69–94.
- Plato. (1961a). Euthyphro. In E. Hamilton, & H. Cairns (Eds.) with L. Cooper (Trans.), *Collected dialogues* (pp. 169–185). Princeton: Princeton University Press.
- Plato. (1961b). Republic. In E. Hamilton, & H. Cairns (Eds.), with P. Shorey (Trans.), *Collected dialogues* (pp. 575–844). Princeton: Princeton University Press.
- Raven, M. (2012). In defence of ground. *Australasian Journal of Philosophy*, 90, 687–701.
- Rodgers, R., & Maranto, C. (1989). Causal models of publishing productivity in psychology. *Journal of Applied Psychology*, 74, 636–649.
- Rodriguez-Pereyra, G. (2005). Why Truthmakers. In H. Beebe & J. Dodd (Eds.), *Truthmakers: The contemporary debate* (pp. 17–31). Oxford: Oxford University Press.
- Rosen, G. (2006). The limits of contingency. In F. MacBride (Ed.), *Identity and modality* (pp. 13–39). Oxford: Oxford University Press.
- Rosen, G. (2010). Metaphysical dependence: Grounding and reduction. In B. Hale & A. Hoffmann (Eds.), *Modality: Metaphysics, logic, and epistemology* (pp. 109–136). Oxford: Oxford University Press.
- Ruben, D. H. (1990). *Explaining explanation*. London: Routledge.
- Russell, B. (1985). In D. Pears (Ed.), *The philosophy of logical atomism*. La Salle, IN: Open Court.
- Russell, B. (2003). Analytic Realism. In S. Mumford (Ed.), *Russell on Metaphysics* (pp. 91–96). London: Routledge.
- Schaffer, J. (2005). Contrastive causation. *Philosophical Review*, 114, 327–358.
- Schaffer, J. (2009). On what grounds what. In D. Chalmers, D. Manley, & R. Wasserman (Eds.), *Metametaphysics: New essays on the foundations of ontology* (pp. 347–383). Oxford: Oxford University Press.
- Schaffer, J. (2010a). Monism: The priority of the whole. *Philosophical Review*, 119, 31–76.
- Schaffer, J. (2010b). Contrastive causation in the law. *Legal Theory*, 16, 259–297.
- Schaffer, J. (2012). Grounding, transitivity, and contrastivity. In F. Correia & B. Schnieder (Eds.), *Metaphysical grounding: Understanding the structure of reality* (pp. 122–138). Cambridge: Cambridge University Press.
- Schnieder, B. (2011). A logic for 'Because'. *The Review of Symbolic Logic*, 4, 445–465.
- Shulz, K. (2011). "If you'd wiggled A, then B would've changed" causality and counterfactual conditionals. *Synthese*, 179, 239–251.
- Sider, T. (2011). *Writing the book of the world*. Oxford: Oxford University Press.
- Spirtes, P., Glymour, C., & Scheines, R. (2000). *Causation, prediction, and search* (2nd ed.). Cambridge, MA: The MIT Press.
- Spirtes, P., & Scheines, R. (2004). Causal inference of ambiguous manipulations. *Philosophy of Science*, 71, 833–845.
- Thomson, J. J. (1983). Parthood and identity across time. *Journal of Philosophy*, 80, 201–220.
- Trogon, K. (2013a). Grounding: Necessary or contingent? *Pacific Philosophical Quarterly*, 94, 465–485.
- Trogon, K. (2013b). An introduction to grounding. In M. Hoeltje, B. Schnieder, & A. Steinberg (Eds.), *Varieties of dependence: Ontological dependence, grounding, supervenience, response-dependence* (pp. 97–122). Munich: Philosophia Verlag.
- Van Fraassen, B. (1980). *The scientific image*. Oxford: Oxford University Press.
- Van Fraassen, B. (1989). *Laws and symmetry*. Oxford: Oxford University Press.

- Weslake, B. (forthcoming). A partial theory of actual causation. *British Journal for the Philosophy of Science*.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Wilson, J. (2010). What is Hume's dictum, and why believe it? *Philosophy and Phenomenological Research*, 80, 595–637.
- Wilson, A. (2013). Schaffer on laws of nature. *Philosophical Studies*, 164, 653–667.
- Wilson, J. (2014). No work for a theory of ground. *Inquiry*, 57, 535–579.
- Wold, H. (1964). Causality and econometrics. *Econometrica*, 22, 162–177.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- Woodward, J. (2008). Mental causation and neural mechanisms. In J. Hohwy & J. Kallestrup (Eds.), *Being reduced: New essays on reduction, explanation, and causation* (pp. 218–262). Oxford: Oxford University Press.
- Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics*, 5, 161–215.
- Wilson, A. (manuscript). Metaphysical causation.