

## **We Need Non-Factive Metaphysical Explanation**

**Michael Bertrand**

*Forthcoming in Erkenntnis (DOI: 10.1007/s10670-020-00227-2). Please cite the published version*

Abstract: Suppose that A explains B. Do A and B need to be true? Provided that we have metaphysical explanation in mind, orthodoxy answers “yes:” metaphysical explanation is factive. This article introduces and defends a non-factive notion of metaphysical explanation. I argue that we need a non-factive notion of explanation in order to make sense of explanationist arguments where we motivate a view by claiming that it offers better explanations than its competitors. After presenting and rejecting some initially plausible rivals, I account for non-factive metaphysical explanation by drawing on existing applications of structural equation models to metaphysical grounding.

### **1 Introduction**

Say that A metaphysically explains B. Do A and B need to be true?

Orthodoxy answers yes. One thing cannot metaphysically explain another unless both are so.<sup>i</sup> I argue that in addition to the familiar factive notion, we must recognize and adequately characterize a non-factive notion of metaphysical explanation.

Why is a ball disposed to roll down an incline plane? One answer is that a downward gravitational force acts on the ball. This is a causal explanation. But another acceptable answer is that the ball is disposed to roll in virtue of being spherical. This is a paradigmatic metaphysical explanation.<sup>ii</sup> It identifies what grounds the ball’s disposition to roll supposing that dispositional properties are grounded in categorical ones. Metaphysical explanations are

widely thought to be identical to or underwritten by distinctively metaphysical relations of determination.<sup>iii</sup>

Metaphysical explanation is the subject of a fast growing literature and is quickly assuming a central role in metaphysics. It underwrites the Argument from Explanation for grounding and so backs the “Grounding Revolution” (Audi 2012a; Kovacs 2017). The formulation of theses like physicalism implicate it (Dasgupta 2014; Ney 2016). And since metaphysical explanation is sometimes claimed to be the tightest explanatory connection, giving them is a central aim of metaphysics (Fine 2012; Glazier 2017).

Given this, we might expect explanation to be factive and so to relate only obtaining worldly facts. Metaphysical determination is factive (Schaffer 2009; Audi 2012a; Fine 2012). And work on metaphysical explanation usually assumes Explanatory Realism: the view that all metaphysical explanations identify or provide information about metaphysical determination (Schaffer 2017a). I assume this throughout. I'll also assume that grounding is a primitive ontological determination relation relating worldly facts. It is distinct from metaphysical explanation and underwrites it.<sup>iv</sup>

Still, I argue that metaphysicians need a non-factive explanatory notion. I begin in section 2 with an explanationist argument: one that motivates a view by claiming that it explains better than its rivals. This sort of argument is ubiquitous and often concerns metaphysical explanation. But it is plausible only if understood in terms of non-factive metaphysical explanation. Given this, we need a precise account of non-factive metaphysical explanation (N-explanation). Though there are almost no existing accounts, I'll explore some plausible options in section 3. I offer my own view in section 4 and conclude by sketching a procedure for using N-explanations to help our search for factive metaphysical explanations (F-explanations).

## 2 Explanationist Arguments

Suppose all metaphysical explanatory claims are factive. Now imagine two philosophers arguing about what accounts for shared properties like redness. Rose is a realist about universals. She claims that objects are red when they share the universal redness (Armstrong 1980; Lewis 1983).<sup>v</sup> Neil is a class nominalist. He claims objects are red just when they belong to the class of red objects.

Rose argues that we should prefer realism to class nominalism because realism does better on explanatory grounds. Since universals are sparse, they explain objective similarity among objects. Two objects are similar in virtue of sharing a universal and different in virtue of failing to share one. These paradigmatic metaphysical explanations work by identifying the instantiation pattern of universals that metaphysically determines resemblance. In contrast, properties are abundant according to class nominalism. Any pair of objects share and fail to share infinitely many properties (Lewis 1983).

Rose's argument has a familiar abductive form. If theory R explains better than theory N, we ought to prefer R over N. R *does* explain better than N. So we ought to prefer it.<sup>vi</sup> Realism explains better in virtue of explaining more – it offers the only adequate explanation for resemblance. So we ought to prefer realism to class nominalism.

This isn't the end of the story: Neil may contest Rose's argument, perhaps by offering his own explanation of resemblance (as Lewis 1983 does). But we have enough already: explanationist arguments like Rose's beg the question.

Just as Penelope can't run faster than Genevieve unless Penelope runs, a theory R can't explain better than a theory N unless R explains. By supposition, metaphysical

explanation is factive: if R explains, then it is true. So, if R explains better than N, then R is true. Substituting realism for R and nominalism for N, it follows that if realism explains better than nominalism then realism is true.

Because of this, Rose begs the question when claiming that we should prefer realism over nominalism on explanatory grounds. If realism explains better than nominalism, then realism is true. But it is plausible that true theories are *always* preferable to false ones. So preferring realism involves judging that it is more likely than nominalism to be true. Rose argues that realism is more likely than nominalism to be true in part by assuming that realism is true and nominalism is false.

This generalizes to any explanationist argument featuring mutually exclusive theories. When we claim a theory R explains better than a theory N, we presuppose that R explains and so is true. N must therefore be false. True theories are preferable to false ones and preferring theory R involves judging that it is more likely true than theory N. So we have begged the question against proponents of theory N.

We might have expected explanationist arguments to turn on the merits of the explanatory packages competing theories offer. We identify the explanations each provides and determine which competitor explains best. But provided that a theory R explains, it is the only theory among its incompatible rivals to do so. It automatically explains best since an explanation is better than none at all.<sup>vii</sup> Neil therefore has no option in the case at hand but to deny that realism ever explains. Conceding that it explains means conceding that realism is true and so nominalism is false.

Considerations like the forgoing threaten the role metaphysical explanation plays in theory choice. Weighing the metaphysical explanations offered by competing theories is either too easy or too hard. It is too easy when we know which among a set of incompatible

rivals is true: explanations are better than their rivals when they are the only ones available. It is too hard when we don't know which theory is true. Then we don't know which theories explain at all and so *a fortiori* don't know which explains best. Either way explanationist arguments aren't fruitful.

Something has gone wrong here. And I think our factive notion of metaphysical explanation is to blame (see also Lipton 2004, 58). Perhaps, though, we can avoid a non-factive notion by employing a preface instead. "According to realism, shared universals explain resemblance" doesn't entail that there are universals or that they really do explain. I'll argue, in section 3.3, that this Preface Strategy offers no obvious advantages over accounts of N-explanation (which I'll take up first) but introduces at least one significant cost.

We can also avoid N-explanation by giving up explanationist arguments altogether. Since these are ubiquitous (examples include Daly 1994; Rosen and Dorr 2002), giving them up would be deeply destructive for metaphysics. It would also be very surprising were explanationist arguments to uniformly beg the question.

It's evident that we need a non-factive notion of metaphysical explanation. Now we need to account for it; my project in what remains.

### **3 Generating a Notion of N-Explanation**

Three possible strategies come to mind: one merely subtracts from F-explanation the requirement that it have true relata. Another offers a conditional account, and a third attempts to avoid N-explanation using a preface. I'll argue that each fails in revealing ways. I take it any adequate account of N-explanation will do justice to explanationist arguments like the one in section 2.

#### **3.1 The Subtractive Strategy**

On a natural picture of F-explanation, a set of facts  $\Gamma$  must satisfy a bundle of conditions in order to metaphysically explain another fact, [B]. Perhaps  $\Gamma$  and [B] must be related by metaphysical determination, connected by some metaphysical law, or bring about some sort of unification. Filling this in is a job for accounts of F-explanation. Whatever conditions must be satisfied for  $\Gamma$  to F-explain [B] – call these conditions C – we can be assured that they include or entail that F-explanations are factive.

The *Subtractive Strategy* proposes that the other conditions required for explanation are separable from those ensuring factivity. We can generate conditions for N-explanation just by subtracting factivity-guaranteeing conditions from the other conditions in C. In this way, the Subtractive Strategy defines N-explanation in terms of F-explanation, while preserving as much similarity as possible between them. Though we may want to reserve the term ‘explanation’ for the factive sort, N-explanations seem to be potential (or purported) F-explanations of a kind long recognized in philosophy of science. Potential F-explanations are possibly true N-explanations (Hempel 1965, p. 38; Lipton 2004, p. 58). In light of this, we might think Rose’s explanationist argument can be accommodated with a simple paraphrase: Understand Rose to claim that realism offers a better potential F-explanation than nominalism does. Potential F-explanation is non-factive. So Rose need not beg the question against Neil. I’ll argue shortly that cashing out N-explanation in this way is a non-starter.

A closely related additive strategy, suggested in work by Fine 2012 and Litland 2017, reverses the order of priority between N and F-explanation.<sup>viii</sup> Rather than defining non-factive explanation and grounding in terms of factive versions of these notions, non-factive grounding is primitive and factive grounding is defined by adding a factivity condition. Supplemented with unionism, the claim that grounding just is a distinctive sort of metaphysical explanation, this view adopts N-explanation as its primitive notion.

Both strategies are problematic. The additive strategy takes N-explanation as primitive and so declines to analyze it. Yet, our project here is to analyze N-explanation if we can. The Subtractive Strategy attempts this. But it can't account for cases of N-explanation consistent with the behavior expected of metaphysical explanation. For one thing, I'm skeptical that plausible conditions on explanation are generally separable from truth. For example, do false explanatory claims unify?

Setting this aside, the Subtractive Strategy seems not to adequately define the notion of N-explanation. My worry is similar to Fine's worry concerning the definition of non-factive grounding in terms of factive grounding (Fine 2012, pp. 49–50). The Subtractive Strategy can naturally characterize N-explanation as follows: a set of facts  $\Gamma$  N-explains  $\{B\}$  iff  $\Diamond(\Gamma$  F-explains  $[B])$ . Situations where  $\Gamma$  N-explains  $[B]$ ,  $\Gamma$ , and  $[B]$  are the case verify this metaphysical possibility. However, this is too restrictive because it rules out cases where the propositions in the N-explanation are jointly impossible. Consider that conjunctions are explained by their conjuncts. This slogan corresponds to the conjunction introduction rule for grounding and so F-explanation ( $\wedge$ I) (Fine 2012, 58). And given our characterization of N-explanation above,  $\wedge$ I applies equally to N-explanation. In accordance with this rule, any conjunction of the form  $\Phi \wedge \sim\Phi$  should be N-explained by  $\Phi$  and  $\sim\Phi$ : its conjuncts. However, this is not possible according to the definition of N-explanation since  $\Phi$  and  $\sim\Phi$  are not jointly possible. So,  $\wedge$ I fails to hold generally for N-explanation given our definition. To avoid this, we may loosen the definition by allowing cases like ours to count as well:  $\Gamma$  N-explains  $[B]$  whenever the propositions in  $\Gamma$  are jointly incompatible. But this would allow irreflexivity violations since  $[A]$  would N-explain itself whenever  $[A]$  is impossible (Fine 2012, pp. 49–50). Understood as we have, the Subtractive Strategy struggles

with extensional adequacy: it is either too restrictive or too permissive. Perhaps with work we could eliminate problems like these. However, I see no obvious way to do this.

### 3.2 The Conditional Strategy

The Subtractive Strategy fails to implement the attractive idea that we can generate our account of N-explanation by modifying our antecedent view of F-explanation. Perhaps a strategy built on subjunctive conditionals yields better results.

*Conditional Strategy:*  $\langle A \rangle$  N-explains  $\langle B \rangle$  iff, were [A] and [B] the case then [A] would F-explain [B].

N-explanations are very much like F-explanations given certain assumptions. For example, it is natural to understand Rose's claim as follows: if it were the case that objects shared universals, then objects would be similar in virtue of sharing universals and different in virtue of failing to share them. It is because realism explains resemblance in this way that realism has more explanatory power and explains better than nominalism.<sup>ix</sup>

The Conditional Strategy is friendly to orthodox assumptions about metaphysical explanations provided that these are also understood in terms of subjunctive conditionals. For example, the Conditional Strategy is compatible with a claim in the spirit of Explanatory Realism: for propositions  $\langle A \rangle$  and  $\langle B \rangle$  and their corresponding facts, if  $\langle A \rangle$  N-explains  $\langle B \rangle$  then, if [A] and [B] were the case then [A] would metaphysically determine [B] (or stand in the ancestral of determination). We can evaluate N-explanations according to how accurately they represent the possible (or as we'll see impossible) determinative structures they describe.

Unfortunately, the Conditional Strategy faces its own difficulties. It cannot do justice to the cases that motivated us to posit N-explanation: explanationist arguments. We've assumed this as an adequacy condition. Specifically, to be any good against Neil's



nominalism, Rose but not Neil must provide Rose's N-explanation of resemblance. Recall that Rose's argument against Neil turns on the claim that realism N-explains resemblance and nominalism does not. This is not the case if it follows from the definition of N-explanation that Nominalism can offer every N-explanation realism can.

By agreeing that shared universals N-explain resemblance, Rose and Neil both endorse the following conditional:

*Shared Universals Conditional:* Were it the case that similar objects shared universals and different objects failed to share them, then the sharing or failure to share universals would F-explain similarity and difference.

By Rose's lights, this conditional has a true antecedent and its truth-value is determined wholly by its consequent. When evaluating the conditional, we need only look to whether shared universals F-explain resemblance in the actual world.

Things are very different by Neil's lights. According to him, the Shared Universals Conditional is a counterpossible: a subjunctive conditional with a metaphysically impossible antecedent. After all, Neil claims that nominalism is true and necessarily so.<sup>x</sup> Maybe all counterpossibles are trivially true (Lewis 1973, p. 24; Williamson 2008, pp. 171–5). In that case, nominalism entails the Shared Universals Conditional. Or, we might extend the familiar possible worlds semantics to range also over impossible worlds (c.f. Nolan 1997).

Accordingly, evaluating the Shared Universals Conditional requires us to examine the nearest worlds, possible or impossible, where objects instantiate and share universals. Since realism is necessarily false (says Neil), this means examining those impossible worlds that most resemble the actual world consistent with the truth of the antecedent.<sup>xi</sup> Again, the Shared Universals Conditional is true given nominalism. Though impossible, the realism-worlds described by Rose are similar to the actual world in many respects. They are plausibly among

the nearest impossible worlds in which similar objects share universals. The Shared Universals Conditional is evaluated with respect to these worlds and in them, shared universals F-explain resemblance.

The Shared Universals Conditional is true given nominalism and that means nominalism is consistent with the claim that shared universals N-explain resemblance. Understood as the Conditional Strategy requires, it's no longer the case that realism N-explains resemblance and nominalism can't. Neil can offer exactly the same N-explanation Rose does and so nominalism N-explains resemblance exactly as well as realism. This generalizes and is destructive to explanationist arguments; what motivated N-explanation in the first place. In every case where opposing views are jointly metaphysically inconsistent, *any* N-explanation offered by one party can be co-opted by the other. No party can enjoy an explanatory advantage on these grounds and explanationist arguments are therefore dialectically useless.<sup>xii</sup>

Perhaps there are other ways of pressing Rose's distinctive claim on the relevant N-explanation. Rose's N-explanation of resemblance is only an F-explanation if realism is true. But it can't be anything more than an N-explanation given nominalism. Let's modify the Conditional Strategy to capture the distinctive relationship between N-explanations and the views that, were they the case, would verify their corresponding F-explanations.

*Revised Conditional Strategy:* <A> N-explains <B> according to view R iff, if view R were the case then [A] would F-explain [B].

The problem with the Revised Conditional Strategy is that Neil shouldn't agree that the right side of the biconditional is true in our case.<sup>xiii</sup> Applied to Rose's claim, the revised strategy yields the following:

*Realism Biconditional:* Shared universals N-explain resemblance according to realism iff, if realism were the case then shared universals would F-explain resemblance.

According to Neil, the right hand side of the Realism Biconditional is a counterpossible. It is evaluated in the nearest impossible worlds in which its antecedent is true. Given nominalism, worlds in which universals abound and explain resemblance do not meet that description. A wide variety of realism-worlds are closer by. These worlds preserve nominalist claims about properties and resemblance in nearly every case, deviating only with respect to a single small, local miracle. In them nearly every resemblance fact remains primitive. Perhaps these worlds are exactly like the actual world except that they contain just one uniquely instantiated universal in some isolated corner of the universe. Since the nearest realism-worlds are not ones in which shared universals F-explain resemblance (says Neil), Neil ought to reject the right hand side of the revised strategy's biconditional as applied to Rose's explanation. Supposing the Revised Conditional Strategy, this means rejecting Rose's N-explanation and so denying the alleged explanatory advantage on which her explanationist argument was predicated.

In order to prevent this move, we might modify further. Consider:

*Twice Revised Conditional Strategy:*  $\langle A \rangle$  N-explains  $\langle B \rangle$  according to view R iff, if view R, [A], and [B] were the case then [A] would F-explain [B].

By packing [A] and [B] into the antecedent of the embedded conditional, we prevent Neil's response to the Revised Conditional Strategy. He should agree that if realism were the case, universals were shared, and objects resembled one another, then shared universals would F-explain resemblance. Nonetheless, the concern re-emerges that Neil can co-opt Rose's explanation

Neil now claims shared universals N-explain resemblance according to nominalism too. Nominalism verifies the following: if nominalism were the case, universals were shared, and objects resembled one another, then shared universals would F-explain resemblance. Since Rose and Neil ought to agree that the antecedent of this conditional is logically inconsistent, the nearest impossible worlds in which it is the case are quite distant from the actual world. In them, the Law of Non-Contradiction fails to hold. Nonetheless, this is a world in which similar objects share universals and one in which no obvious disruption to the world's explanatory structure has occurred. In such a world, the relevant counterpossible is arguably true. Supposing the Twice Revised Conditional Strategy, it is open for Neil to argue that shared universals N-explain resemblance according to nominalism. Of course, shared universals don't F-explain resemblance according to nominalism. But on pain of begging the question, F-explanation is not what is at issue in Rose's explanationist argument. We've assumed an extended possible worlds semantics in this discussion. However, if Neil's counterpossible is trivially true then it remains the case that universals N-explain resemblance according to nominalism.

Perhaps we haven't exhausted its resources. But the Conditional Strategy faces a serious challenge. It must find a way of accounting for N-explanation such that relevant N-explanatory claims are consistent with the commitments held by both parties to a well-formed explanationist argument. But it must also ensure that both parties cannot generally claim these N-explanations at once.

### 3.3 The Preface Strategy

Maybe we've overcomplicated things. Let's paraphrase alleged explanatory claims with "according to view R," thereby canceling their factive implications. On this Preface Strategy, Rose means to claim that *according to realism*, shared universals explain resemblance.

Opponents like Rose and Neil can agree about prefaced explanatory claims even when they dispute un-prefaced counterparts. And the Preface Strategy avoids a notion of N-explanation altogether.

Unfortunately, the Preface Strategy struggles to account for claims of relative explanatory goodness. What (if any) parsimony advantage it gains by replacing the relational predicate “N-explains” with a sentential operator is worth little.

Operators and predicates often correspond, as is the case here. “According to realism, universals explain resemblance” seems equivalent to “universals explain resemblance according to realism.” Elsewhere, we choose between operators and predicates largely on pragmatic grounds. For example, some prefer the view on which grounding is a sentential operator because it makes the logic of ground easier and is neutral about the existence of worldly facts (Litland 2017 fn. 8).

However we formulate them, we must say why non-factive explanatory claims are true when they are. On its own, the preface *according to view* R does not do this and so doesn't compete with the views I've discussed. Supplemented with plausible accounts of correctness, the Preface Strategy very closely resembles the accounts we've already considered and rejected.

Several straightforward accounts of correctness are non-starters. Consider: “According to view R, A explains B’ is true iff ‘A explains B’ is among the sentences that constitute R.” This is implausibly restrictive since theories need not include explanatory claims in order to support them. Likewise, an entailment view is a non-starter. Consider: “According to view R, A explains B’ is true iff R entails that A explains B (perhaps together with auxiliary claims).” The entailment view counts false Rose’s explanation since the truth of realism does not entail that shared universals explain resemblance. The realism-worlds

from section 3.2 illustrate this. The Preface Strategy is more attractively understood in terms of subjunctive conditionals: “‘According to view R, A explains B’ is true iff were R the case, A would explain B.” But this is just the Revised Conditional Strategy rejected in section 3.2. We might also treat theories as models. However, the result is very much like the Model Strategy I’ll propose in the next section. Plausible versions of the Preface Strategy are not distinct from the accounts we’ve already considered; they differ only in the semantic machinery they employ.

It’s because of this semantic machinery that the Preface Strategy does worse in at least one respect. Its sentential operator offers less expressive power than comparable relational views. Explanationist arguments often involve competing proposed explanations. Suppose in response to Rose, Neil offers a Lewisian competitor: co-membership in a natural class explains resemblance. To preserve her claim that realism explains better than nominalism, Rose’s explanation must be superior to Neil’s (according to some agreed upon standard of explanatory goodness). The Preface Strategy struggles with the resulting comparative claim: “Rose’s explanation of resemblance is better than Neil’s.”

The explanatory notion appealed to here is factive (the Preface Strategy avoids commitment to N-explanation). But Rose and Neil can’t both have offered factive explanations, since their claims are incompatible. Prefaces are no help since their claims require different ones. And a preface like “according to view R” can’t be applied separately to each part of the comparative claim. Prefaces operate on sentences, not definite descriptions like ‘Rose’s explanation.’

Prefacing the whole comparative claim won’t help either. Consider that “According to our best theory of explanation, Rose’s explanation of resemblance is better than Neil’s” seems to entail “According to our best theory of explanation, Rose explains resemblance.”

After all, if Rose fails to explain, then she fails to explain better. Since the preface strategy only recognizes the factive notion of explanation, it follows that "According to our best theory of explanation, Rose's explanans is true."

The prefaced comparative claim also seems equivalent to "According to our best theory of explanation, Neil's explanation of resemblance is worse than Rose's." This entails that "According to our best theory of explanation, Neil explains resemblance." Via factivity, it follows that "According to our best theory of explanation, Neil's explanans is true." The preface strategy thus predicts that our best theory of (factive) explanation denies the Law of Non-Contradiction: it requires that inconsistent theories be true together. This prediction is false. A better theory is available and avoids inconsistency by supplementing F-explanation with N-explanation.

We may resist this argument by allowing the prefaced comparative claim to come out true when "Neil's explanation of resemblance" fails to refer. This gambit makes our prefaced comparative claim dialectically useless in explanationist arguments. Rose can only make use of the gambit by presupposing that nominalism is false and so Neil's offering fails even to count as an explanation. This begs the question against Neil. In sum, Rose can't offer the prefaced comparative claim in an argument against Neil's nominalism without either making implausible commitments about our best theory of (factive) explanation – namely it entails the Law of Non-Contradiction is false – or by begging the question against her opponent.

We gain little by pursuing the Preface Strategy. Whatever benefit we enjoy by avoiding N-explanation is more than offset by our inability to account for the comparative claims that drive explanationist arguments.

#### **4 The Model Strategy**

I've argued that metaphysics needs a non-factive notion of explanation in order to do justice to explanationist arguments. Though other strategies face serious challenges, I think we can do better by developing the following picture:  $\langle A \rangle$  N-explains  $\langle B \rangle$  iff, in a model of the right sort, [A] is represented as F-explaining [B]. On this view, N-explanatory claims depend on the availability of appropriately understood models. This account is adequate for explanationist arguments like Rose's.

#### 4.1 Structural Equation Models

Though many things would make adequate models, I will use structural equation models (hereafter SEMs) of the sort developed by Pearl 2000. Roughly speaking, SEMs are sets of equations that represent networks of relationships among variables. In addition to their familiarity in the literatures on grounding and causation, SEMs have at least three advantages over alternatives. First, they are more precise than the subjunctive conditionals appealed to by the Conditional Strategy. This makes clear exactly what is and isn't part of our model and makes explicit how parts of the model are connected (Schaffer 2016, 60). Second, SEMs are effective at modeling a wide range of different causal dependencies (Schaffer 2016, 60). Their expressive power offers a reason to be optimistic that SEMs can model explanatory dependence too. Third, this optimism is compounded by the fact that structural equation models are already successfully employed to model the metaphysical determinative relations that are supposed to underlie metaphysical explanation (c.f. Schaffer 2016; Wilson 2018).<sup>xiv</sup>

Formally, SEMs consist of three ingredients: a set of variables, a set of structural equations linking the values of these variables, and an assignment function that specifies what values variables actually take. It's helpful to think of variables as questions answered in different ways, each corresponding to a different value (see Briggs 2012). When modeling



explanations, these ingredients receive a specific interpretation. Understand variables to represent questions about specific states of the world, structural equations to represent relations of explanatory dependence, and assignments as answers to the questions posed by particular variables.

For a simple example, consider a model, SOCRATES, representing the explanation of {Socrates} in terms of Socrates. A natural way to represent this bit of the world makes use of two variables: **Socrates** for whether Socrates exists and **Singleton** for whether {Socrates} does. Each of these variables comes with a range of possible values corresponding to distinct, incompatible ways of answering the questions they pose. Both variables ask whether something exists, so only two possible answers corresponding to two values are needed: 1 for yes and 0 for no. Its variables and space of possible assignments make up SOCRATES *signature*.

Next we need to define a structural equation for the model, representing relationships of direct explanatory dependence between different states of the world represented by variables. These equations have the form  $\mathbf{A} = f(\mathbf{B}_1 \dots \mathbf{B}_n)$ , where  $\mathbf{A}$  is an "endogenous variable" representing dependent conditions and the  $\mathbf{B}_n$ 's are those variables on which  $\mathbf{A}$  depends. This equation says that  $\mathbf{A}$ 's values explanatorily depend on the values of the  $\mathbf{B}_n$ 's in the way stated by the function  $f(x)$  (Woodward 2003; Wilson 2018). In SOCRATES, we need only one structural equation representing the explanatory dependence of {Socrates} on Socrates: **Singleton** = **Socrates**. In words, {Socrates} exists because Socrates does. Together with the signature, this structural equation constitutes the linkage of our model.

Importantly **Singleton** = **Socrates**, and more generally  $\mathbf{A} = f(\mathbf{B}_1 \dots \mathbf{B}_n)$ , do not involve identity or indicate symmetric relations. Rather, they are relations of asymmetric

dependence in which the left-hand variable depends for its values on the right-hand variables.

Lastly, we need to answer the questions posed by the variables in our model. We do this by setting a function that maps each "exogenous variable," representing independent conditions, to exactly one value. This is the assignment of the model. With respect to SOCRATES, we need to say whether Socrates exists. By experimenting with different mapping functions, we can chart the behavior of every endogenous variable based on possible assignments to the exogenous ones. The model encodes various counterfactuals of the form, if **Socrates** were assigned value  $n$ , then **Singleton** would have had value  $n$ . When Socrates exists, it follows that {Socrates} exists because Socrates does. When he doesn't exist, {Socrates} fails to exist because of it. In short, SOCRATES represents the existence of {Socrates} as explanatorily depending on the existence of Socrates.<sup>xv</sup>

Fitting SEMs into the intuitive statement of the Model Strategy produces the claim that  $\langle A \rangle$  N-explains  $\langle B \rangle$  iff the right sort of SEM is available and it represents B as explanatorily depending on A. Importantly, this intuitive sketch is compatible with a variety of existing approaches to causal and metaphysical explanation. It has so far paralleled interventionist treatments of causation and grounding but nowhere presupposes them since its sole concern is explanation (and I have assumed separatism). Nonetheless, it isn't surprising that the network of potential metaphysical explanations is well represented by SEMs supporting interventionist counterfactuals.

#### 4.2 Models of the Right Kind

It's time to say what it takes for a model to be of the right kind. Insofar as we're interested in what a statement would explain supposing that it is true, we should choose our models for their ability to represent the explanatory connections posited or presupposed by

these statements. Intuitively, the model should represent the world as it is supposed to be by the lights of the explanans and its background, so we can see whether that explanans really N-explains as advertised. I encode this idea in the following conditions.

*Apt*: A model, M is of the right kind iff

1. M's signature represents the system under study as it is supposed to be according to the explanans and its background.
2. M's linkage is consistent with those general principles that we take to actually govern factive explanatory dependence.<sup>xvi</sup>
3. M's assignment represents the system as being in a state where the explanans is the case.

Clause 1 of *Apt* ensures that the system being represented is carved up in a way that reflects the ontological commitments supposed or assumed by the explanans of the N-explanation under investigation, where the system being represented is determined by the contents of the explanation. This doesn't require that the explanans make any explicit claims about the system's ontological joints or determinative structure.

With respect to the {Socrates} explanation, the system being described is a subset of the world's set theoretic structure. We need not be realists about that structure in order to represent it as it is supposed to be by those who are. While we might have carved up the system in many ways, doing so in terms of Socrates and {Socrates} is most perspicuous given the explanation under evaluation.

Consistent with the aim of representing the commitments supposed by the relevant explanans, clause 1 ought to prevent the signature of an apt model from being interpreted to represent more features of the system than are included or supposed by the explanans. Since

it is possible to significantly alter or destroy an explanation by adding more information, allowing supplementation undermines the usefulness of our models.

Clause 2 of *Apt* ensures that the structural equations representing alleged explanatory dependence are consistent with antecedent commitments concerning the nature of F-explanation. This ensures that, when its variables and assignments veridically represent the world, the structural equations of our model represent genuine F-explanatory dependence. I will not try to say to which general principles a linkage ought to conform. That job is for theories of factive metaphysical explanation. Nonetheless, our assumed Explanatory Realism furnishes some examples. Explanatory connections only obtain between dependent outcomes and that which constitutively generates them (e.g Schaffer 2017b, p. 304). As a result, explanatory dependence is asymmetric, irreflexive, and transitive.<sup>xvii</sup> We may disagree about which principles in fact govern explanatory dependence and so disagree about what clause 2 entails. In that case, I suggest that we develop rival models in parallel and compare their fruits.

Clause 2 and the assumption of Explanatory Realism jointly entail that apt models must encode specific commitments concerning metaphysical determination. We can make these commitments explicit merely by reinterpreting the linkage of the model so that  $\mathbf{A} = f(\mathbf{B}_1 \dots \mathbf{B}_n)$  says that A's values are *metaphysically determined* by the values of the  $\mathbf{B}_n$ 's in the way stated by function  $f(x)$ . The linkage of an apt model must represent a direction of explanatory priority whose entailed determinative commitments are consistent with those generalizations concerning metaphysical determinative relations that are supposed by the explanans and its background: e.g. sets are grounded in their members (Schaffer 2017b).<sup>xviii</sup> This feature of clause 2 is especially important for ruling out spurious models. And as a

matter of consistency, the relevant theory's other theoretical commitments should constraint entailed determinative commitments.

Finally, clause 3 ensures that the exogenous variables representing the system under study are set to values that represent a state in which the explanans is true.

### 4.3 The Model Strategy Stated

Armed with the notion of an apt model we're in position to officially state the Model Strategy.

*Model Strategy.*  $\langle A \rangle$  N-explains  $\langle B \rangle$  iff there is an Apt structural equation model  $M$ , such that assignments of  $M$ 's variables representing  $\langle B \rangle$ 's being true are based on assignments of  $M$ 's variables representing  $\langle A \rangle$ 's being true.

We can further define notions of complete and partial N-explanation:  $\langle A \rangle$  completely (partly) N-explains  $\langle B \rangle$  iff the assignments of  $M$ 's variables representing  $\langle B \rangle$ 's being true are based wholly (partly) on assignments of  $M$ 's variables representing  $\langle A \rangle$ 's being true.

Call  $M$  an apt witnessing model for the relevant N-explanatory claim.  $M$  satisfies *Apt* and witnesses the correctness of its corresponding claim in the sense that what it is for the claim to be correct is for there to be an apt model representing the N-explanation it asserts. Models that satisfy *Apt* are intentionally designed to represent the world's constituents and explanatory connections as supposed by the explanation under scrutiny. This reflects the intuition that N-explanations ought to be just like F-explanations provided that we've supposed the truth of the explanans. The Model Strategy adds a precise way of understanding this intuition.

The Model Strategy also does justice to the intuition that N-explanatory claims are claims about the world. This is because discovering that one thing N-explains another consists in discovering certain worldly facts: namely, that the world is such that there is an

apt witnessing model. What exactly must the world be like in order for this to be the case? The Model Strategy is itself neutral and this is a virtue of the view. In particular, we can afford to remain relatively neutral with respect to ontological debates concerning abstracta like models. The Model Strategy makes no claims about what models are, only that the world is such that they exist. Since models may exist without being metaphysically fundamental, irreducible, or may exist in a lightweight sense, the Model Strategy is compatible with a wide range of views about their ontology (see Schaffer 2015).<sup>xix</sup>

In order to determine whether an N-explanatory statement is true, we need only determine whether there is an apt witnessing model. It bears repeating that whether this is the case is independent of the outcomes of those metaphysical debates that might appeal to N-explanation. In particular, the Model Strategy is totally silent about whether apt witnessing models are veridical representations of the actual world. It claims only that apt models represent the explanation as correct: there is a model that represents the explanans as true, whose variable basing functions are consistent with the general principles governing factive explanatory dependence, and which bases the values of the explanandum on the values of the explanans.

By the lights of the Model Strategy, symmetric pairs of N-explanations are possible. The priority monist claims that the cosmos ("the ultimate concrete whole") explains its many parts (Schaffer 2010, p. 31). The priority pluralist says it goes the other way round. Each theory offers a successful N-explanation differing only in its direction of explanatory priority. All this is as it should be: the monist and pluralist differ only about what determines what. Yet we can certainly ask which theory explains best (Schaffer 2010, p. 59). There remains an important sense in which N-explanation is asymmetric. There can be no single model, and so no single N-explanation, featuring explanatory circles. Such a model would

violate clause 2 of *Apt* by requiring circles of metaphysical determination, something forbidden given explanatory realism and standard assumptions about metaphysical determination.

The Model Strategy bears some resemblance to the Conditional Strategy considered and rejected in section 3.2. It's natural to think of both strategies as species of a common genus unified by the thought that N-explanations are like F-explanations given that we have supposed or stipulated the truth of their explanantia. On the Conditional Strategy, we attempted to do this by examining the nearest worlds where the explanans is true. The Model Strategy is distinguished by a different approach: we build the truth of the explanans into the model and then examine how this affects the truth of the explanandum. Because the Conditional Strategy directs us to whatever worlds meet the description "the nearest worlds in which the explanans is true," we've seen that it invites disputes about which worlds are nearest. In contrast, the Model Strategy identifies the relevant situation directly, thereby avoiding such disputes.

The Model and Conditional strategies might be brought closer together by understanding apt models as ersatz worlds selected by an appropriate selector function. Though not my view, such a model based conditional strategy better resists the objections offered in section 3.2. On it, N-explanations are subjunctive conditionals verified by models interpreted as ersatz worlds. Nonetheless, model based conditional strategies are importantly different from the Model Strategy I've proposed. Unlike model based conditional strategies, The Model Strategy does not require commitments to ersatz worlds. It requires only that apt SEMs are capable of serving as abstract models of systems as they are supposed to be by the explanantia of some relevant N-explanatory statement.

#### 4.1 The Model Theory Applied

To see in more concrete terms how the Model Strategy works, revisit the disagreement between Rose and Neil. Recall that Rose's explanationist argument against Neil goes as follows: realism explains better than nominalism because it, unlike nominalism, can explain what makes for genuine similarity and difference. The explanations for resemblance that Rose has in mind plausibly share some common forms – X is similar to Y (in some respect) because X and Y both instantiate universal U; X is a duplicate of Y because every universal instantiated by one is instantiated by the other; and X is different from Y (in some respect) because X instantiates U and Y does not. Let's take up the explanation of a particular similarity as a representative instance. Realism explains that the ball and the stop sign have the same color because they both instantiate the universal scarlet.

The Model Strategy invites us to cash out Rose's N-explanation in terms of the existence of apt models whose variables witness the explanation by standing in the right sort of variable-basing relationships: those that represent their targets as explanatorily depending on alleged explanantia.

This requires that there be an apt structural equation model in which the variable representing color resemblance is based on variables representing the instantiation of color universals in the ball and the sign. The Model Strategy proceeds in two steps. First, we need to construct a model witnessing the N-explanation Rose has offered. Second, we need to determine whether the resulting model is apt. If so, then Rose has offered a correct N-explanation's. We can weigh its explanatory virtues and vices just as we do for F-explanations.

A very simple model witnessing Rose's N-explanation is easy to design and can be modified in various ways to make more sophisticated similarity judgments, to capture other



respects of sameness, or to describe comparisons between more than two objects. Call our model RESEMBLANCE. Its signature includes variables representing the instantiation of color universals in the ball (**Ball**) and sign (**Sign**) and a binary variable representing whether or not they are exactly the same color (**Same**). **Ball** and **Sign** are exogenous variables while **Same** is endogenous. Since their values correspond to answers to the question ‘what colors are the ball and sign,’ let **Ball** and **Sign** take as possible values any of the rational numbers. Interpret these numbers to represent maximally determinate color universes (perhaps including the absence of any color universal) assigned via any adequate method we’d like.<sup>xx</sup> Since **Same** is a binary variable, it can be answered either yes (1) or no (0).

According to the explanation that Rose has offered, the instantiation of color universals in the ball and sign collectively fully explain their sameness. Given this, we can assign a linkage to represent the relationship of explanatory dependence that Rose takes to hold. Its structural equations will run from variables representing the instantiation of universals to variables representing sameness.

In this case, just one structural equation is needed:

$$\mathbf{Same} = \begin{cases} 1 & \text{if } |\mathbf{Ball} - \mathbf{Sign}| = 0 \\ 0 & \text{if } |\mathbf{Ball} - \mathbf{Sign}| > 0 \end{cases}$$

Finally, we state what color universals are instantiated by the ball and sign. This means assigning **Ball** and **Sign** values drawn from their space of possible values. The explanans states that both objects are scarlet, so we assign **Ball** and **Sign** to whatever value corresponds to scarlet in our chosen enumeration. Supposing for the sake of simplicity that scarlet is a maximally determinate color represented by the number 1, then **Ball** = **Sign** = 1.

By basing **Same** on **Ball** and **Sign**, RESEMBLANCE represents color sameness as explanatorily depending on the instantiation of color universals in the ball and the sign: these

objects perfectly resemble one another color-wise iff they instantiate the same maximally determinate color universal. RESEMBLANCE witnesses Rose's alleged N-explanatory claim. So, we have completed the first step required by the Model Strategy. What remains is to determine whether RESEMBLANCE is an apt witnessing model. Rose's N-explanation is correct if and only if this is the case.

*Apt* places constraints on all three components of RESEMBLANCE, beginning with its signature. Specifically, RESEMBLANCE counts as apt iff it represents the relevant objects and universals as they are supposed to be according to the explanans. This is plausibly the case in RESEMBLANCE. Its variables carve the system up naturally and at the right level of description. It represents the explanans as a pair of exogenous variables, the explanandum as an endogenous variable, and the contrast space of values is appropriately rich.

RESEMBLANCE also clearly satisfies *Apt*'s constraint on its assignment. An assignment is apt when it represents the explanans as being the case. Since both **Ball** and **Sign** are assigned the value representing the universal scarlet, the relevant bit of the world is represented as it is claimed to be by the explanans.

Finally, *Apt* requires that RESEMBLANCE's linkage be consistent with those general principles that we in fact take to govern F-explanatory dependence. In absence of a canonical theory, it is not clear precisely what this requires. Nonetheless, our assumption of Explanatory Realism furnishes several substantive constraints. Factive metaphysical explanations chart the world's relations of constitutive generation and, in virtue of this, are plausibly irreflexive, asymmetric, and transitive.

The structural equation in RESEMBLANCE is apt when it charts the world's relations of constitutive generation as they are supposed to be by the explanans and its

background. This equation describes an irreflexive relationship. Its interpretation ensures that it is asymmetric, and it isn't complex enough to exhibit intransitivity. Finally, the directed explanatory dependence that it imposes entails specific metaphysical determinative connections that are consistent with those generalizations concerning constitutive generation or grounding that Rose in fact takes to govern the relationship between resemblance and the instantiation of universals. Rose is plausibly committed to the claim that the instantiation of universals in objects metaphysically determines the resemblance between them. The linkage of our model bases resemblance on shared universals consistent with this commitment. So, RESEMBLANCE is plausibly an apt witnessing model and as a result, Rose's N-explanatory claim succeeds.

As before, none of the ingredients in Rose's apt witnessing model are factive: neither representation in the model nor basing relations among variables require that the thing represented actually obtains. So, Rose's model should be amenable to Neil as well: it avoids begging any questions against his nominalism. In order to evaluate the N-explanatory claim, Rose and Neil develop structural equation models representing it, determine whether these models appropriately witness the explanation in question, and test whether they are apt in light of Rose's commitments concerning grounding.<sup>xxi</sup> Whether N-explanation occurs is determined solely by the features that the model actually has. Rose and Neil can agree about them.

In light of this agreement, we might worry that the Model Strategy is subject to the same problems that plagued the Conditional Strategy. After all, the Model Strategy aims to capture what's plausible about the Conditional Strategy and bears some resemblance to it. The problem for the Conditional Strategy (whether in its original or revised form) was that it left Rose unable to press an explanatory advantage over Neil's nominalism. Depending on

how the strategy was understood, the nominalist could either co-opt Rose's N-explanation or reject it. Yet her claim that realism offers an unopposed explanation for resemblance was required for the success of Rose's explanationist argument.

As with the Conditional Strategy, Rose and Neil ought in principle to agree that the sharing (or not) of universals N-explains sameness and difference. On the Model Strategy, this amounts to agreeing that RESEMBLANCE is an apt witnessing model. This much is required in order for Rose's N-explanatory claim to function as neutral evidence useful for adjudicating the debate between realism and nominalism.

However, the Model Strategy affords Rose a distinctive way of claiming her N-explanation that is not available to the nominalist. The N-explanation of similarity in terms of shared universals is appropriately claimed by realism and cannot be appropriated by nominalism simply because RESEMBLANCE represents a system in which realism is true and nominalism is false. In short, any apt model that witnesses Rose's N-explanation also witnesses realism about universals and is incompatible with nominalism. This is in the same spirit as the Realism Biconditional from section 3.2. However, my claim here does not concern what is the case in the nearest worlds in which realism is true. It is instead a claim about the system that the model in fact represents. Indeed, RESEMBLANCE fails clause 2 of *Apt* when offered by Neil: it entails determinative commitments that are inconsistent with other theoretical commitments made by nominalism – namely the non-existence of universals.

This entails something similar to the right hand side of the Realism Biconditional: If RESEMBLANCE were a veridical model, then universals would F-explain similarity and difference. But unlike the right hand side of the Realism Biconditional, Neil shouldn't reject this conditional claim. The nearest worlds in which RESEMBLANCE is veridical are ones in

which shared universals really do F-explain sameness. Neil ought to agree that Rose's N-explanation is genuine. And since RESEMBLANCE does not verify nominalism, Neil can't plausibly co-opt Rose's N-explanation either. The Model Strategy successfully accounts for Rose's explanationist argument against Neil.<sup>xxiii</sup> In doing so, it makes available the tools required to underwrite dialectically useful explanationist arguments.

### **5 Conclusion: N-explanatory diagnostics for F-explanation**

In addition to the factive notion of metaphysical explanation already familiar in the literature, we require a non-factive notion of metaphysical explanation. This non-factive notion is what we appeal to when claiming that, between two incompatible theories, one offers a better metaphysical explanation for some phenomenon than the other. The Model Strategy offers a precise way of evaluating N-explanatory claims that nonetheless remains appropriately neutral: we can agree about what N-explains what in the midst of deep disagreement about what there is and what grounds what. The Model Strategy demonstrates that we can understand N-explanation in a way that is non-factive and at the same time is of a kind with F-explanation and similarly connected to metaphysical determination.

What's the relationship between F-explanation and N-explanation? Unfortunately, I lack the space to answer fully. Nonetheless in the spirit of Fine 2012, we can define the notion of F-explanation in terms of N-explanation as follows:

*F-explanation:* [A] F-explains [B] iff  $\langle A \rangle$  N-explains  $\langle B \rangle$ , [A] and [B] are the case, and stand in the determinative relationship entailed by " $\langle A \rangle$  N-explains  $\langle B \rangle$ ."

This is akin to the additive strategy from section 2: F-explanations are N-explanations that are appropriately backed and have true relata. Unlike that strategy, we've characterized N-explanation. The Model Strategy also suggests a way of defining N-explanation in terms of

F-explanation, since we look to the nature of F-explanation in order to define what linkages count as *apt* in our model.

These notions are inter-definable, so there is no easy demonstration of the claim that F-explanation is more fundamental than N-explanation. Nonetheless, there are strong indicators that this is true. Intuitively, the Model Strategy consists of constructing a model representing a system in which F-explanation occurs, in order to determine what N-explains what. This is parasitic on F-explanation and so is our intuitive grip on N-explanation as potential F-explanation.

Because of the close relationship between N and F-explanation, N-explanation can be useful for those of us who are interested in understanding the nature of metaphysical explanation more generally. We seem to owe our grip on metaphysical explanation to the examination of paradigmatic examples. And it is often claimed that theorizing about explanation should begin with examples and do justice to them. But clear and uncontroversial examples of metaphysical explanation are difficult to find: metaphysics is subject to a high degree of uncertainty and disagreement is the rule.

Under these circumstances, it often proves useful to assume some interesting claims for the sake of argument in order to ask how they would be explanation-wise related if they were true. In other words, it's useful to limn the explanatory connections among commitments in a way that is detached from considerations about their truth. We can do this by appealing to N-explanation.

One way N-explanation might prove useful for general accounts of metaphysical explanation is as a sort of diagnostic for failures of F-explanation. If  $\langle A \rangle$  N-explains  $\langle B \rangle$ , we don't yet know whether the facts they represent are also related by F-explanation. N-explanatory claims do not secure F-explanatory ones. But if  $\langle A \rangle$  fails to N-explain  $\langle B \rangle$ ,

perhaps by requiring explanatory connections that are inconsistent with our other commitments concerning F-explanation and grounding, we learn something about the F-explanatory structure of the world as well as its N-explanations: nowhere in the actual world does [A] F-explain [B] (unless, of course, our background assumptions are wrong). Even when we can afford only agnosticism about whether [A] and [B] are the case, N-explanation offers up a tool for ruling out explanatory connections between them. What remains are those potential F-explanations that are worthy of study and development. By developing theories to account for these cases of N-explanation, we will with luck produce successful theories of F-explanation as well.

Focusing on N-explanation doesn't require anti-realism or a less ambitious view of metaphysics. Metaphysics aims at F-explanation. I've suggested N-explanations feature in a useful strategy for meeting this goal. Though compatible with many metaphysical views, N-explanation pairs naturally with the view of metaphysics as modeling. Metaphysicians, like scientists develop models (often abstract or idealized) as indirect tools for understanding the world and its structure (Paul 2012; Williamson 2017). Integrating N-explanation into the metaphysics-as-modeling view must be left to future work.

---

<sup>i</sup> Though occasionally explicit (Fine 2012; Schaffer 2016), this commitment is more often implicit. For example, grounding is widely claimed to relate facts or to operate on sentences, where the result is true only if its component sentences are. Grounding is then claimed to underwrite or be identical with metaphysical explanations sharing their relata (c.f. Rosen 2010).

<sup>ii</sup> Audi 2012b offers a longer list of canonical examples. Some metaphysical explanations may not be determinative or underwritten by grounding. See Bertrand 2018; Taylor 2018.

---

<sup>iii</sup> This disjunction divides unionism from separatism (Raven 2015). Both disjuncts are denied by opponents of Explanatory Realism (see Thompson 2016; Kovacs 2017; Taylor 2018).

<sup>iv</sup> Unionists hold that determinative relations like composition or property realization back explanation. Jessica Wilson 2014 bases her skeptical challenge to grounding on this.

Determinative relations are factive. And that's all we need.

<sup>v</sup> This recalls Lewis's 1983 argument against David Armstrong but isn't intended as a faithful representation. Lewis' argument is a more complex example of my target.

<sup>vi</sup> I leave open why we should prefer the best explanation. Perhaps explanatoriness is the mark of truth. But Rose may instead endorse the weaker claim that explanatoriness is a theoretical virtue. Other things equal, we should accept the most virtuous theory.

<sup>vii</sup> This case shares its relevant features with more complex cases where explanations compete but doesn't require judgments of explanatory goodness.

<sup>viii</sup> Though Fine's considered view is that the factive notion is more fundamental. See Fine 2012, p. 50.

<sup>ix</sup> Subjunctive conditionals may have true antecedents (Lewis 1973, p. 3).

<sup>x</sup> This is tacitly assumed in the metaphysics of properties, though Miller 2010 disputes it. Theories of properties are accounts of their nature and things have their natures necessarily.

<sup>xi</sup> Impossibility weighs heavily against similarity (Nolan 1997, p. 550).

<sup>xii</sup> Similarly, a strategy built on the material conditional fails:  $\langle A \rangle$  N-explains  $\langle B \rangle$  iff, If  $\langle A \rangle$  and  $\langle B \rangle$  then  $[A]$  F-explains  $[B]$ . Since by the lights of one party the embedded conditional will have a (necessarily) false antecedent, the conditional is trivially true. The explanation can be co-opted by any view on which  $\langle A \rangle$  is necessarily false.

<sup>xiii</sup> Unless counterpossibles are trivially true. If so, the Revised Conditional Strategy entails that any necessarily false theory offers every possible N-explanation.



---

<sup>xiv</sup> Some metaphysical determinative relations are same-making rather than difference-making and aren't well represented by structural equation models. See Koslicki 2016, pp. 106–109; J. Wilson 2016, pp. 189–191.

<sup>xv</sup> This amounts to defining a model and intervening on it. See Woodward 2003. For an explicit procedure concerning grounding see Schaffer 2016.

<sup>xvi</sup> Without broad agreement concerning how explanation works, it's difficult to see how explanationist arguments have purchase.

<sup>xvii</sup> For this reason, we need not allow A to explain itself in the case that troubled the Subtractive Strategy. However, none of these claims are wholly uncontroversial. See Jenkins 2011 on irreflexivity, Schaffer 2012 on transitivity, and Barnes Forthcoming on asymmetry.

<sup>xviii</sup> Why doesn't the problem for non-factive explanatory claims also emerge for non-factive grounding claims? It may. But I don't face it here. I require consistency with the world's determinative structure *as it is taken to be by the explanans and its background*. This concerns the determinative commitments of a theory and not the world's determinative structure.

<sup>xix</sup> If we don't have qualms about abstracta then there's an apt witnessing model because the world has among its constituents the abstract object that is the model. Those with qualms need a suitable nominalism.

<sup>xx</sup> I assume there's some enumeration scheme mapping determinate shades to unique real numbers.

<sup>xxi</sup> The debate needn't unfold this way. My claim is that the existence of an apt witnessing model makes Rose's explanatory claim true.

<sup>xxii</sup> Rose and Neil might both have had genuine N-explanations. The debate would then turn on whose N-explanation is better. One N-explanation is better than another if the explanation represented by its apt witnessing model is more virtuous than its competitor.

---

Determining how explanatory virtue works is beyond the scope of this paper. But it's on this point that we should decide explanationist arguments.

### Acknowledgements

### References

- Armstrong DM (1980) *Universals and Scientific Realism*. Vol. 1 Nominalism and Realism. Cambridge University Press
- Audi P (2012a) Grounding: Toward a Theory of the In-Virtue-of Relation. *J Philos* 109:685–711
- Audi P (2012b) A Clarification and Defense of the Notion of Grounding. In: Correia F, Schnieder B (eds) *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press, pp 101–121
- Barnes E (Forthcoming) Symmetric Dependence. In: Bliss R, Priest G (eds) *Reality and Its Structure*. Oxford University Press
- Bertrand M (2018) Metaphysical Explanation by Constraint. *Erkenntnis* 1–16
- Briggs R (2012) Interventionist Counterfactuals. *Philos Stud* 160:139–166
- Daly C (1994) Tropes. *Proc Aristot Soc* 94:253–261
- Dasgupta S (2014) The Possibility of Physicalism. *J Philos* 111:557–592
- Fine K (2012) Guide to Ground. In: Correia F, Schnieder B (eds) *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press Cambridge, pp 37–80
- Glazier M (2017) Essentialist Explanation. *Philos Stud* 174:2871–2889
- Hempel CG (1965) *Aspects of Scientific Explanation*. Free Press
- Jenkins CS (2011) Is Metaphysical Dependence Irreflexive? *The Monist* 94:267–276
- Koslicki K (2016) Where Grounding and Causation Part Ways: Comments on Schaffer. *Philos Stud* 173:101–112
- Kovacs DM (2017) Grounding and the Argument From Explanatoriness. *Philos Stud* 174:2927–2952
- Lewis D (1983) New Work for a Theory of Universals. *Australas J Philos* 61:343–377
- Lewis DK (1973) *Counterfactuals*. Blackwell

- 
- Lipton P (2004) *Inference to the Best Explanation*. Routledge
- Litland J (2017) Grounding ground. *Oxf Stud Metaphys* 10:279–316
- Miller K (2010) Three routes to contingentism in metaphysics. *Philos Compass* 5:965–977
- Ney A (2016) Grounding in the Philosophy of Mind: a Defense. In: Aizawa K, Gillett C (eds) *Scientific Composition and Metaphysical Ground*. Springer, pp 271–300
- Nolan D (1997) Impossible worlds: A modest approach. *Notre Dame J Form Log* 38:535–572
- Paul L (2012) Metaphysics as modeling: the handmaiden’s tale. *Philos Stud* 160:1–29
- Pearl J (2000) *Causality: Models, Reasoning, and Inference*. Cambridge University Press
- Raven MJ (2015) Ground. *Philos Compass* 10:322–333
- Rosen G (2010) Metaphysical dependence: Grounding and reduction. In: Hale B, Hoffmann A (eds) *Modality: Metaphysics, logic, and epistemology*. Oxford University Press Oxford, pp 109–36
- Rosen G, Dorr C (2002) Composition as fiction. *Blackwell Guide Metaphys* 151–174
- Schaffer J (2016) Grounding in the Image of Causation. *Philos Stud* 173:49–100
- Schaffer J (2009) On What Grounds What. In: Manley D, Chalmers DJ, Wasserman R (eds) *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press, pp 347–383
- Schaffer J (2017a) The Ground Between the Gaps. *Philos Impr* 17:1–26
- Schaffer J (2017b) Laws for Metaphysical Explanation. *Philos Issues* 27:302–321
- Schaffer J (2012) Grounding, Transitivity, and Contrastivity. In: Correia F, Schnieder B (eds) *Metaphysical Grounding: Understanding the Structure of Reality*. Cambridge University Press, pp 122–138
- Schaffer J (2015) What not to multiply without necessity. *Australas J Philos* 93:644–664
- Schaffer J (2010) Monism: The Priority of the Whole. *Philos Rev* 119:31–76
- Taylor E (2018) Against Explanatory Realism. *Philos Stud* 175:197–219
- Thompson N (2016) Grounding and Metaphysical Explanation. *Proc Aristot Soc* 116:395–402
- Williamson T (2008) *The Philosophy of Philosophy*. John Wiley & Sons

Williamson T (2017) Model-building in philosophy. *Philos Future Probl Philos Prog* 159–172

Wilson A (2018) Metaphysical Causation. *Noûs* 723–751

Wilson J (2014) No Work for a Theory of Grounding. *Inq Interdiscip J Philos* 57:535–579

Wilson J (2016) The Unity and Priority Arguments for Grounding. In: Aizawa K, Gillett C (eds) *Scientific composition and metaphysical ground*. Springer, pp 171–204

Woodward J (2003) *Making Things Happen: A Theory of Causal Explanation*. Oxford University Press