Explanatory asymmetry in non-causal explanation

Andrew Wayne

Forthcoming: *Journal of General Philosophy of Science*

Abstract

The problem of explanatory asymmetry remains a serious challenge for non-causal accounts of explanation. This paper proposes a novel solution, and it does so by appealing to the theoretical context in which an explanation is offered. The paper develops the problem of explanatory asymmetry for non-causal dependency accounts of explanation, focusing specifically on Alexander Reutlinger's Counterfactual Theory of Explanation and recent work by Marc Lange and Lina Jansson. It defends the idea that nomological possibility with respect to a global theory is the right constraint on explanation, and it shows how this breaks the apparent symmetry in counterfactual dependence that is the source of the problem. This solution appeals to theoretical context, and the paper develops and defends the version of contextualism in explanation that is required.

1. Introduction

There is an intuitive asymmetry in our explanatory practices such that, for example, the height of a flagpole, plus a law of geometrical optics and fixed angle of the sun, plays a part in explaining the length of its shadow; but the length of the shadow, with the same law and angle, normally plays no part in explaining the height of the pole. Many explanations found in science exhibit exactly this sort of built-in asymmetry. If scientific explanation works by citing the relevant causal information, then causal asymmetry straightforwardly accounts for explanatory asymmetry. It is easy to see that the flagpole's height is a causal factor in producing the shadow, but the shadow is no part of the cause of the pole's height. If some explanations are not causal, however, problems arise. Accounting for this asymmetry is a well-known challenge for non-causal theories of scientific explanation.

On traditional approaches, explanation works by subsuming the phenomenon to be explained under a law (Hempel 1965; Kitcher 1989). Recent work on non-causal explanation replaces this subsumption with a dependency condition, and it emphasizes the centrality of scientific models to explanation. According to these accounts, elements of models, including non-causal dependencies, can explain phenomena in a target system (Bokulich 2011; Batterman and Rice 2014; Saatsi 2016; Reutlinger 2016; Reutlinger 2018; Woodward 2018). As proponents of these accounts recognize, they have as great a difficulty accounting for explanatory asymmetries as did earlier non-causal accounts. For Juha Saatsi and Mark Pexton, the price of relinquishing causal explanation is giving up any solution to the problem of explanatory asymmetry. As they put it, “The causal aspect… plays an indispensable role in responding to a familiar puzzle about explanatory asymmetries in connection with explananda concerning singular states of affairs” (Saatsi and Pexton 2013, 614-5). It is impossible, according to them, for a non-causal account of explanation to solve the problem of explanatory asymmetry when it comes to explaining particular phenomena, such as the height of a flagpole (615). In a review of non-causal dependency accounts, Alexander Reutlinger concludes that how a non-causal counterfactual theory of explanation can account for explanatory asymmetries is “one of the deepest puzzles of the current philosophy of explanation” (Reutlinger 2016, 739). Michael Strevens maintains that there is no solution to the problem of explanatory asymmetry without invoking causes. His account—which otherwise bears many similarities to recent non-causal accounts—invokes causes at precisely this point. Anti-realism about causation at the level of fundamental physics “clearly cannot solve the problem of explanatory asymmetry” (Strevens 2008, 34-5). “The raw material of explanation…. is a fundamental-level structure of causal influence revealed by physics” (Strevens 2018, 102). As James Woodward poses the problem, “One role that the [causal] notion of an intervention plays is that it excludes forms of counterfactual dependence that do not seem explanatory… how we can recognize and exclude these [non-explanatory forms of counterfactual dependence] if we don’t have the notion of an intervention to appeal to?” (Woodward 2018, 123).

A view widespread among both proponents of non-causal explanation and those more skeptical of these accounts is that appealing to causes is the best—or only—way to solve the problem of explanatory asymmetry. Yet we have good reason to believe that some *bona fide* explanations are simply not causal, at least in physics. If this is so, then we urgently need a non-causal solution to the problem of explanatory asymmetry. This paper proposes a solution, and it does so by appealing to the scientific context in which an explanation is offered. I argue that model-based explanations in physics must have a certain kind of connection to the global theories in which they are embedded. Section 2 articulates the problem of explanatory asymmetry for non-causal accounts of explanation, focusing specifically on Reutlinger’s Counterfactual Theory of Explanation (2016, 2018). Section 3 examines recent attempts by Marc Lange and Lina Jansson to solve the problem of explanatory asymmetry by taking into account contextual factors, including the global theoretical context of an explanation. Section 4 develops and defends my approach, that nomological possibility with respect to a global theory is the right constraint on explanations via local models, and it shows how this resolves the problem of explanatory asymmetry for non-causal dependency accounts. My solution appeals to theoretical context, and Section 5 develops and defends the version of contextualism in explanation that is required.

2. The problem of explanatory asymmetry in non-causal explanation

Explanatory asymmetry is a well-known difficulty for the deductive-nomological account of scientific explanation, on which explanations are arguments showing how a phenomenon to be explained follows deductively from laws and empirical conditions (Hempel and Oppenheim 1948; Hempel 1965). It is equally a vexed question for unificationist accounts, in which science explains by reducing the number of phenomena that have to be taken as brute (Friedman 1974; Kitcher 1981; Kitcher 1989). As we shall see, recent non-causal dependency theories of explanation do no better than Hempel's and Kitcher's subsumption theories to debar asymmetrical non-explanatory arguments.

Philosophical accounts of scientific explanation typically focus on what may be called *local conditions* on explanation. These have to do with characteristics of the explanans and the relation between the explanans and the explanandum. Hempel's deductive-nomological account, for instance, requires that the explanandum be entailed by the explanans, the explanans contain one or more laws and empirical conditions, and the explanans be true or approximately true.

Far less attention has been paid to *global conditions* on explanation. These are not characteristics of the explanans, the explanandum, the relation between the two, or their relation to a target system. Rather, global conditions concern the relation between the explanation—or more precisely, the model described by the explanation—and the larger scientific field in which it is embedded ([self-reference omitted]). These contextual features of explanation have been largely neglected in the recent literature, and as we shall see, non-causal dependency accounts of explanation are typically framed entirely in terms of local conditions. By contrast, global conditions are front and center in Kitcher's unificationist account, where the explanatory power of an argument is determined in part by the degree of unification of the global theory in which the argument is embedded (Kitcher 1981; Kitcher 1989). In large part, Kitcher's account aims to explicate how a global theory gets its explanatory power. This issue can be set aside for present purposes, as we are concerned with explanatory asymmetries among arguments within the *same* global theory. Rather, our task is to explore the nature of this embedding, and more specifically the connections between explanatory global theories and local models. It is here the solution to the problem of explanatory asymmetry lies.

A global theory, such as Newton's mechanics with its three laws of motion and gravitational law, provides a scientific context for a local model, such as the simple pendulum. Some connections between global theory and local model enable the global theory to confer explanatory power on the local model. For Kitcher, these are cases in which an empirical generalization in an explanans is identical with a law of the global theory, or deducible from the global theory plus empirical conditions described in the explanans, or perhaps deducible from a straightforward and well-justified approximation of the global theory (Kitcher 1989, 452-454). This approach to global conditions rests on the assumption that what makes a local model explanatory, in the context of a global theory, has to do exclusively with the relation between laws of the global theory and generalizations in the local model. It is here that the problem of explanatory asymmetry arises for Kitcher (as it does for Jansson, who takes the same approach). Once we accept that a deductive or approximately deductive relation is all that is needed for explanatory power to flow from global theory to local model, it becomes very difficult to exclude the asymmetrical derivations. For example, on Kitcher's account a derivation of the period of a pendulum is explanatory because the simple pendulum law is (approximately) a deductive consequence of Newton's laws, and Newtonian mechanics is an explanatory global theory. From the simple pendulum law and empirical conditions one can derive and explain the period of a pendulum. The trouble is that the reverse derivation of the length of a pendulum from its period uses the same pendulum law, appears to satisfy Kitcher's conditions for explanation, but clearly does not explain the pendulum's length (see Kitcher 1981, 525-6).

A frequent criticism levelled at Hempel and Kitcher is that their requirement that the explanans subsume the explanandum under a generalization is inadequate for a successful account of explanation (Woodward 2003, 203-204). Rather, to explain a phenomenon is to show what it depends on. This is surely the right intuition about many scientific explanations, but it does not help resolve the problem of explanatory asymmetry. In a model of a simple pendulum, the pendulum law supports the counterfactual that if the length were different, the period would be different, and thereby explain the period. However, the same generalization appears to support the counterfactual that if the period were different, the length would be different, and thereby putatively explain the length. It seems that non-causal dependency theories of explanation do no better than Hempel's and Kitcher's non-causal subsumption theories to debar non-explanatory arguments. As we have seen, proponents of such accounts acknowledge as much. The remainder of this section diagnoses why this is so.

Non-causal dependency accounts of explanation make use of counterfactuals, which describe non-causal dependency relations in models, to do the explanatory work. They typically begin with the interventionist account of causal explanation developed by Woodward (Woodward 2003) and then modify or relax various requirements to cover non-causal situations (Bokulich 2009; Bokulich 2011; Rice 2015; Saatsi 2016; Woodward 2018). These accounts include that the explanandum depends on the explanans, the explanans contains a difference-maker, and the explanans answers what-if-things-had-been different questions about the explanandum. Causal explanations are those in which all the generalizations in the explanans describe causes, where the distinguishing features of causal dependencies include their directional and temporal asymmetries. Non-causal explanations are those in which at least some of the generalizations in the explanans do not describe causal dependencies. The dependencies may fail to be causal, for example, because they are not temporally asymmetric, or they may be conceptual or logical dependencies, or the dependencies may not allow interventions on the explanans variables.

Let us focus on one well-developed non-causal dependency account, Reutlinger's Counterfactual Theory of Explanation (CTE) (2016; Reutlinger 2017; 2018). On the CTE account, explanations are deductive arguments with true or approximately true premises. Premises must include one or more nomic generalizations (G), statements about empirical conditions such as initial or boundary conditions (IC), and auxiliary assumptions (A). In order to debar cases of explanatory irrelevance, CTE includes a minimality condition such that "no proper subset of the set of explanans statements $\{G_1, \ldots, G_m, IC_1, \ldots, IC_n, A_1, \ldots, A_o\}$" satisfies all the conditions of the CTE (2018, 79). What makes CTE a dependency account is the Dependency Condition:

> (DC) $G_1, \ldots, G_m$ support at least one counterfactual of the form: if the initial conditions $IC_1, \ldots, IC_n$ had been different than they actually are (in at least one specific way deemed possible in the light of the nomic generalizations), then E, or the conditional probability of E, would have been different as well (2018, 79).

According to Reutlinger, these conditions are individually necessary and jointly sufficient for scientific explanation (2018, 79). He emphasizes that CTE is a monist account of explanation, in that its conditions cover both causal and non-causal explanations (2018, 79). As Reutlinger acknowledges, there is quite a bit of similarity between CTE and Woodward's account of causal explanation. The main difference is that for Woodward, DC is limited to interventionist counterfactuals, whereas CTE has no such restriction. The key point for our purposes is that CTE is framed entirely in terms of local conditions on explanation: features of the explanans, the explanandum, and the counterfactual relations between them.

Consider Reutlinger's application of CTE to the case of the simple pendulum (2018, 81-2). The key generalization in the explanans is the pendulum law, which states that the period of a pendulum is proportional to the square root of its length. The explanandum is a specific period, for example, that the pendulum takes 2.1 seconds to complete one swing. The explanans includes the empirical condition that the pendulum is 1 metre long and auxiliary assumptions that the period is small and air resistance, flexibility in the rod and other such factors are neglected. This explanation satisfies DC since it supports the counterfactual: if the length of the pendulum were different from 1 m, the period would be different from 2.1 s. According to CTE, this is a bona fide explanation of the period of a pendulum.

The worry here is that, absent a causal requirement, DC is simply too weak to exclude non-explanatory arguments—precisely the problem of explanatory asymmetry. Consider an argument with the same pendulum law, the same auxiliary assumptions, an empirical condition that the period of the pendulum is 2.1 s, and a conclusion that the length of the pendulum is 1 m. Reutlinger acknowledges that this argument satisfies DC as well, since it supports the counterfactual that if the period of the pendulum were different, the length would be different. CTE yields the incorrect result that the asymmetrical argument is *also* a bona fide explanation, whereas normally the period of a pendulum is no part of the explanation of its length. The point generalizes to a plethora of counterexamples that satisfy CTE's necessary and sufficient conditions for explanation. This is exactly the problem acknowledged above: a counterfactual or dependency theory of explanation lacking a causal requirement cannot be a successful philosophical account of explanation in science.

Reutlinger attempts to address this issue in two ways, neither of which is satisfactory. One strategy is to supplement CTE with an additional constraint in an attempt to exclude the counterexamples (2017, 252-3). He considers the flagpole and shadow case, and he raises the worry that CTE is unable to exclude the putative explanation of the height of the flagpole in terms of the length of the shadow.

> This worry is, however, unjustified, because the CTE is a theory of *causal* and non-causal explanations. Applying the CTE to causal explanations may take the form of supplementing the CTE with Woodward's interventionist account of causation as the underlying theory of causation (2017, 252).

The correct explanation here, according to Reutlinger, is a causal explanation in which the height of the flagpole causes the length of the shadow; the reverse argument can be ruled out simply because it is not causal. Reutlinger is effectively supplementing DC with an additional constraint, which we may call the Dependency Condition Addendum:

> (DCA) If an explanation appeals only to causal dependencies, then all explanations with the same (or perhaps similar) generalizations and auxiliary assumptions, but with different empirical conditions, must also appeal only to causal dependencies.

DCA successfully eliminates the reverse, non-explanatory shadow derivation and saves CTE from this counterexample.

This modification of DC requires a clear distinction between causal and non-causal dependencies, and it treats the two situations differently. As noted above, Reutlinger emphasizes throughout his discussion of CTE that he is taking a monist approach to explanation: "the CTE provides *one* philosophical account of *two* types [causal and non-causal] of explanations" (2018, 74). On a monist approach, it does not much matter whether the counterfactuals describe causal or non-causal dependencies, as CTE conditions apply equally in both cases. Reutlinger does propose a "Russellian strategy" of identifying characteristic features of causal relations (2018, 88-9). While one might or might not find it helpful to use Russellian criteria to categorize explanations as causal or non-causal, this categorization is superfluous to the application of the original CTE. With the addition of DCA, however, applications of CTE to causal and non-causal cases diverge, undermining monism. Even more problematic, the causal/non-causal distinction now becomes a fundamental part of the account of explanation. One must first determine whether a particular generalization in the explanans is causal or not before applying DC and assessing the explanatory merits of the argument.

Giving up monism about explanation and requiring a clear distinction between causal and non-causal dependencies are the costs of supplementing CTE as Reutlinger suggests. What are the benefits? They are limited, simply because this strategy does not help at all with non-causal explanation, where the supplementary DCA does not apply. For example, Reutlinger considers the pendulum law to be a "law of coexistence," following Hempel, and asserts that the explanation of the pendulum's period is a non-causal explanation (2018, 89-90). As he puts it,

> [T]he relevant counterfactual dependencies are not time-asymmetric, since the dependence holds between physical states (length and period) at one and the same time…. Thus, the explanation is non-causal (2018, 90).

As we have seen, CTE admits as explanatory both the derivation of the pendulum's period from its length and the reverse derivation of the pendulum's length from its period. DCA can do no work in non-causal cases to exclude the asymmetrical counterexamples.

Reutlinger suggests a second strategy to resolve explanatory asymmetries: simply accept that, for non-causal explanations, arguments in both directions may be explanatory (2018, 92-3). It seems that in the simple pendulum and other cases of non-causal explanation Reutlinger considers, the explanandum counterfactually depends on the empirical conditions, and at the same time the empirical conditions counterfactually depend on the explanandum (given the same generalizations and auxiliary assumptions). Considering one such example, Reutlinger concludes that "according to the CTE there is no explanatory asymmetry in this explanation" (2018, 92). Here Reutlinger seems to slide from the lack of asymmetry in CTE's *treatment* of the case to their being no asymmetry in the explanation itself. In the pendulum case, CTE yields the result that the putative explanations are symmetrical. In actual scientific contexts, such as Newtonian mechanics, as a rule this is not the case. A satisfactory explanation of the length of a pendulum normally appeals to additional empirical conditions, generalizations and assumptions. For a pendulum constructed by humans, such an explanation might include the intentions of the designer. An explanation of the length of a pendulum that is part of the phenotype of an organism might appeal to Darwinian evolutionary theory. And so on.

The simple pendulum is just one example of the asymmetries present in many non-causal explanations in science. We seem to have a case of an inconsistency between a normative theory of scientific explanation, CTE, and well-accepted features of explanatory practice in science, where it is the normative theory that stands to be corrected. I have focused on Reutlinger's account because it is one of the most well-developed theories of non-causal explanation in the literature, and because it is based on the widespread assumption that explanation is best analyzed in terms of change-relating counterfactuals.[1] I suggest that the problem here is not with the details of Reutlinger's account, but with a methodology shared by other non-causal dependency accounts, namely an exclusive focus on local conditions on explanation.

3. Lange and Jansson on explanatory asymmetry

In different ways, Marc Lange (Lange 2016; Lange 2018; Lange 2019) and Lina Jansson (Jansson 2015; Jansson 2020) both attempt to solve the problem of explanatory asymmetry by widening the focus to take into account contextual factors and global theory. As we shall see in this section and the next, neither account succeeds in capturing an important aspect of the role of global theory in many non-causal explanations in physics.

As Marc Lange frames it, the problem of explanatory asymmetry is resolved by focusing on "explanatory priority" (2016, Ch. 3; 2018, 34-37). Many non-causal explanations contain a symmetrical counterfactual dependence between the law or constraint being explained and laws or constraints doing explanatory work. Lange distinguishes laws or constraints which function as *explanatorily fundamental laws* (EFLs), thus candidates for an explanans from those that are *explanatorily derivative laws* (EDLs) and part of an explanandum. EFLs have explanatory priority over EDLs, and this order of explanatory priority fixes the asymmetry of explanation. However, for Lange there is no general or systematic way to distinguish, in a given non-causal explanation, which are EFLs and which are EDLs. Instead, the order of explanatory priority is grounded differently in different cases (Lange 2019, 14).

One example Lange considers is the standard explanation of the Lorentz transformations, which appeals to the principle of relativity and the invariance of the spacetime

[1] Kareem Khalifa, Gabriel Doble, and Jared Millson have recently challenged this assumption and defended a pluralist approach to the sorts of modal claims underwriting scientific explanations (Khalifa, Doble et al. 2020).

interval to explain why the Lorentz transformations hold (2016, Ch. 3; 2018, 33-36). The Lorentz transformations have the same modal strength, within the structure of special relativity, as does the interval's invariance. He asks:

> [W]hy don't the Lorentz transformations themselves qualify as EFLs and so explain the interval's invariance (which they entail), rather than the reverse? What makes the interval's invariance explanatorily prior to the Lorentz transformations (rather than the reverse, for instance—or the relativity of simultaneity being explanatorily prior to each)? (2018, 34).

Lange's answer appeals to the idea that invariant quantities describe features of the world, while frame-dependent quantities, such as the Lorentz transformations, describe how things appear from a given perspective.

> The behavior of invariant quantities is explanatorily prior to the behavior of frame-dependent quantities because invariant quantities are features of the world, uncontaminated by the reference frame from which the world is being described, whereas frame-dependent quantities reflect not only the world, but also the chosen reference frame. How things *are* explains how they *appear* from a given vantage point… Reality explains mere appearances, and so the law that a certain quantity is invariant takes explanatory priority over the law specifying how a certain frame-dependent quantity transforms (2018, 35).

Lange is clear that very different considerations come into play in other cases: "although reality's explanatory priority over appearances grounds the EFL/EDL distinction in this case, it cannot do so generally" (2018, 35); "what makes one constraint an EFL rather than an EDL may have little to do with what makes another constraint an EFL rather than an EDL" (2018, 36). For Lange, the grounds for determining explanatory priority necessarily appeal to the context in which the explanation is given. However, this context may or may not include what I have called global theory. Lange emphasizes the diversity of successful explanation in science. In some contexts, explanation may succeed in the absence of higher-level laws or global theory at all (2016). In other contexts, there may be a global theory, but it is irrelevant to explanation. To explain the Lorentz transformations, as we just saw, a metaphysical principle independent of global theory—the primacy of reality over appearance—was sufficient to determine the order of explanatory priority and solve the problem of explanatory asymmetry.

Lina Jansson likewise attempts to resolve the problem of explanatory asymmetry by paying closer attention to the contexts in which explanations are given (2015, 2020). Unlike Lange, however, she focuses on one aspect of this context: how higher-level laws of nature licence explanatory inferences. While Lange is primarily focused on explanations of laws and constraints, Jansson targets explanations of particular states of affairs. And contra Lange, Jansson believes that appealing to higher-level laws, laws of the global theory, is sufficient to determine explanatory dependence.

Jansson's starting point is the observation that laws of nature apply to particular cases only under certain conditions. For example, the simple pendulum relation applies when the angle of oscillation is small and there is no air resistance or other forces acting on the pendulum. As we have seen, this supports a symmetrical counterfactual dependency of the period on the length and the length on the period. On Jansson's approach, *conditions of application* of simple pendulum relations in the explanans introduce the required explanatory asymmetry. The period of the pendulum is sensitive to violations of these conditions of application: the presence of large oscillations, air resistance or a driving force, for instance, would change the period. By contrast, the length of the pendulum is not affected by these violations of conditions of application of the pendulum relation. This differential sensitivity to conditions of application gives us good reason to think that the period of the pendulum is explained by the length, but not vice versa. Asymmetrical explanatory dependence is revealed when we look at the sensitivity of these counterfactuals to violations of these conditions.

To clarify the role of laws and their relevance to the relations appearing in a scientific explanation, consider the following explanation sketch:

(1) Explanandum: The period of the pendulum is T.
(2) Particulars: the length of the pendulum is l.
(3) General condition: the length and period of the pendulum satisfy the general condition $T = 2\pi\sqrt{l/g}$.
(4) Laws: there are laws L_1, … L_n such that (a) "they licence the inference from the general condition obtaining to the explanandum obtaining (and a spelling out of how they do so); and (b) licence the inference from a failure of the general condition to obtain to the failure of the explanandum to obtain (and a spelling out of how they do so)."[2]

While (1), (2) and (3) are all elements of the explanation, (4) is not. (4) tells us something about the global theory in which the explanation is embedded and about how this context determines explanatory dependence. We can fill out L_1, … L_n by considering laws that are not part of the explanandum, such as dissipative forces, driving forces, buoyancy, and so on. These laws satisfy

[2] (Janssen 2015, 591-2) provides this sketch for the flagpole and shadow case. I have modified it to apply to the pendulum.

(4)(b), for example, by showing that in the presence of a driving force, the simple pendulum relation fails to apply, and in this way they account for how the pendulum's period will be different from *T*.

Now consider the asymmetrical situation, the putative explanation of the length of a simple pendulum in terms of its period:

> (1´) Explanandum: the length of the pendulum is *l*.
>
> (2´) Particulars: The period of the pendulum is *T*.
>
> (3) General condition: the length and period of the pendulum satisfy the general condition $T = 2\pi\sqrt{l/g}$.
>
> (4) Laws: there are laws L_1, … L_n such that (a) "they licence the inference from the general condition obtaining to the explanandum obtaining (and a spelling out of how they do so); and (b) licence the inference from a failure of the general condition to obtain to the failure of the explanandum to obtain (and a spelling out of how they do so)."

In this case, the L_1, … L_n described above do *not* licence the inference from the failure of the general condition to apply (as when, for example, there is a driving force) to the pendulum's length being different from L. More importantly, there simply are no laws in the global theory that *do* licence this sort of inference. The length of the pendulum is insensitive to the conditions of application of the general condition. (4)(b) is violated, and thus (2´) and (3) do not explain (1´). The conditions of application (4) are not part of the explanandum, but as Jansson rightly emphasizes, "The conditions of application are part of our explanatory practices" (2015, 583). The time-symmetric formulation of the simple pendulum relation supports symmetrical counterfactuals, but this does not imply that we have symmetrical explanations.

Lange and Jansson both appeal to theoretical context to resolve the problem of asymmetry in non-causal explanation. Global theory, however, plays very different roles in their accounts. Lange is clear that higher-level laws of a global theory may play no role whatsoever in fixing explanatory priority (indeed, scientific explanation may be successful in the absence of any such laws or global theory). On Jansson's account, context is relevant to explanation in a unique and specific way: laws of the global theory licence explanatory inferences, and they do so by fixing the conditions of application of the dependency relation in the explanans. Her focus is on explanations of particular facts in physics, and these explanations are only successful when there are higher-level laws tell us precisely when and how the explanandum is sensitive to violations of the conditions of application of dependency relations in the explanans.

My worry is that missing from both of these accounts is one important aspect of role of global theory in many, perhaps most, non-causal explanations in physics. As I shall argue in Section 4, an explanation in terms of a local model is only successful when the model has an appropriate connection with global theory. Unlike Lange, I contend that non-causal explanation in physics requires a robust global theory; without one, local models do not have explanatory power. As for Jansson, she acknowledges that her solution to the problem of explanatory asymmetry does not take into account idealizations and modeling practices (2015, 598). Many explanatory models in physics are highly idealized (Woodward 2003; Strevens 2008; Batterman 2010; Bokulich 2011; Batterman and Rice 2014; [self-reference omitted]). For idealized models, dependency relations doing explanatory work (Jansson's "general conditions") are not deducible from laws of the global theory, and in fact may be inconsistent with those laws ([self-reference omitted]). Moreover, determining when and how the explanandum phenomenon is sensitive to violations of the dependency relations in an idealized model may have nothing to do with inferences licenced by higher-level laws in such cases where these laws do not apply to the model.

4. Global theory and nomological possibility

My proposal is that nomological possibility with respect to a global theory is the right global constraint on explanations via local models, at least in physics. This condition enables us to diagnose and resolve the problem of explanatory asymmetry in CTE and other non-causal dependency accounts, especially in the context of idealized-model explanations.

Our starting point is the idea is that all entities in a local model that are indispensable to an explanation must be nomologically possible according to a relevant global theory. Roughly speaking, an entity is *indispensable* to a model-based explanation when all explanations of this kind of explanandum include claims about the entity. For example, the value of a variable (such as the period of a pendulum) is a kind of explanandum, where different specific values constitute different explananda. A *nomologically possible* entity is one that is consistent with nomologically possible synchronic states and dynamics according to the global theory. A perfectly rigid pendulum rod, for example, is nomologically possible in Newtonian mechanics, even if it is not physically realizable in the laboratory. It is nomologically possible for a massive object subject to an applied force to accelerate. By contrast, the acceleration of a massive object in the absence of an applied force is nomologically impossible in Newtonian mechanics.

The explanation of the period of the simple pendulum in terms of its length clearly satisfies this nomological-possibility constraint. The rigid rod and frictionless pivot are the only entities in the model indispensable to the explanation, both of which have nomologically possible properties and dynamics in Newtonian mechanics. As we have seen, the explanation satisfies the Dependency Condition as well. Consider a pendulum with length *L* and period *T*. If the length of this pendulum were *L**, then the period would be *T**, in accordance with the pendulum law. As I shall argue shortly, this counterfactual is reasonably regarded as true. With both the nomological-possibility constraint and DC satisfied, our philosophical analysis lines up with scientific practice and with our explanatory intuitions: we have explained the period of the pendulum in terms of its length. In the asymmetrical case, the relevant counterfactual is: if the period of this pendulum (of length *L*) were *T**, the pendulum's length would be *L**. As I shall show, this counterfactual is false. The simple pendulum model, in the context of Newtonian mechanics, does not support an explanation of length based on period. A non-causal dependency theory of explanation, supplemented by this nomological-possibility constraint, enables us to exclude the pathological case and solve the problem of explanatory asymmetry.

This approach to excluding non-explanatory arguments crucially relies on claims about the truth or falsity of certain counterfactuals featuring in the Dependency Condition. We can provide some support for these sorts of claims by taking a strategy pioneered by Robert Stalnaker and David Lewis, on which truth conditions for counterfactuals are cashed out in terms of similarity relations among possible worlds (Stalnaker 1968; Lewis 1973). The basic idea is that the counterfactual "if *A* were true, then *C* would be true" is true in world *w* just in case there is a world in which both *A* and *C* obtain that is more similar to *w* than any world in which *A* and not-*C* obtain. According to Lewis, there is no one true similarity metric; rather, measurements of similarity between worlds are context sensitive. However, he proposes two sorts of guideposts to similarity. First, worlds with violations of physical law relative to *w* are less similar to *w* than worlds without such violations. So worlds with the same laws as *w* are more similar to *w* than worlds with small and localized violations of the laws of *w*, which in turn are more similar to *w* than worlds with large and widespread violations. Second, worlds with more extensive matches of particular fact are more similar to *w* than worlds with less. The match must be perfect; according to Lewis, an approximate match or similarity of particular fact is of little relevance (Lewis 1979).

In our case, the object is a universe of models in the context of a global theory. Consider the following Lewis-style definition of Truth Conditions for Model-based Counterfactuals (TCMC):

> (TCMC) The counterfactual "if *A* were true, then *C* would be true" is true of a model *m* in the context of global theory *P* just in case some possible model *m** in which both *A* and *C* are true is more similar to *m* than any possible model *m*** in which *A* and not-*C* are true.

Similarity between models is determined by the components and dependency relations in the models. Two models with exactly-alike components are more similar than models with different components. Following Lewis, the similarity metric weights a perfect match between the components of two models more heavily than any approximative match.

A similarity metric between local models is context-dependent in the sense that it can only be operationalized in the context of a global theory. In Newtonian mechanics, for instance, differences in components are captured in variables describing fundamental physical properties such as length, mass, shape, and rigidity. The goal here is to capture one important way we use counterfactuals in this context, which is to explore what would have happened if the model were to have had slightly different initial or boundary conditions yet the dynamical evolution were to proceed in accordance with Newtonian laws. Consider two Newtonian models with exactly-alike components, one of which contains a small perturbative force slightly altering the dynamics of the model. Scientists would describe these as very similar models, because in practice they could be used to represent an identical target system exhibiting two slightly different behaviours. By contrast, two models with different components—even if the components are of the same kind or approximately alike—are normally used to represent distinct target systems and thus regarded as less similar models.

We are now in a position to defend the above claims about truth values of counterfactuals in the Dependency Condition. Consider a model *m* in which the pendulum's period is *T*, its length is *L*, and there is no driving force. For explanations of a pendulum's period in terms of its length, the relevant counterfactual is:

(C) If the length of this pendulum were *L**, its period would be *T**.

According to (TCMC), we assess the truth of (C) by considering possible models of pendulums with length *L** and assessing their similarity to *m*. The first model, call it *m**, is one in which the pendulum has length *L**, period *T**, and no driving force. In this model, the antecedent and consequent of (C) are both true. The second is a class of models *m*** of pendulums in which the length is *L** but the period is not T*, due to a small driving force on the pendulum (sufficient to alter its period but not disrupt periodic motion). Now, the *m** model differs from *m* only in one component, namely that the length of the pendulum is $L^* \neq L$. The *m*** models differ from *m* in same way (the length of the pendulum is $L^* \neq L$), but they also differ because of the presence of the driving force. For this reason, no *m*** model is more similar to *m* than is *m**. According to (TCMC), counterfactual (C) is true, hence the Dependency Condition is satisfied and the length of the pendulum can explain its period.

The pathological case is a bit more challenging. We would like to determine the truth value of the following claim:

(*D*) If the period of this pendulum were *T**, its length would be *L**.

Again, *m** is the model in which the pendulum has period *T**, length *L**, and no driving force. Consider within the class of models of pendulums with period *T** all those models in which the pendulum has length *L*. These are models with pendulums quite similar to *m*—indeed, all the pendulums in these models have components exactly like *m* with the same length, mass, shape and rigidity. In some of these models, call them *m***, there will be small driving forces which alter the period of the pendulum to $T^* \neq T$. *m* and *m*** contain pendulums that are exactly alike, but differ in that *m*** contains a component (a driving force) that is absent in *m*. By contrast, *m* and *m** contain pendulums of different lengths, but neither contains a driving force. Unlike in the explanatory case, in the pathological case both *m** and *m*** differ from the original model in only one way: *m** with respect to length and *m*** with respect to a driving force. The pendulums in models *m* and *m*** are exactly alike. For a scientist working in Newtonian mechanics, these models can be used to represent the same target system under two scenarios, one with a driving force and the other without. By contrast, models *m* and *m** differ much more significantly, as they contain non-alike pendulums of different lengths. For the working scientist, these two models would be used to represent distinct target systems with different components. I contend that *m*** is more similar to *m* than is *m**, and so by (TCMC) counterfactual (D) is false. The derivation of the pendulum's length in terms of its period is unexplanatory.

My approach sets an admittedly high bar for idealized models to support non-causal explanation in physics, and in two ways. First, explanation requires a robust global theory to govern the universe of nomological possibilities, properties and entities that constrain local model-building. Without this, there is no way to evaluate the truth value of counterfactuals in the explanans. Indeed, in the absence of global theory it is difficult to make sense of the counterfactuals appearing in the Dependency Condition as being true or false at all. In situations in which a well-defined global theory is lacking, my account yields the result that local models cannot support non-causal explanations. Consider the pendulum model in its original scientific context, namely Galileo's work on pendulums and falling objects. Galileo measured the elapsed time for pendulums of various lengths to swing to vertical, from small pendulums on his desktop to a large pendulum he swung over his second-storey office balcony (Drake 1989). He noticed a roughly constant ratio between the length of the pendulum and the square of the time. He constructed what contemporary physicists would call phenomenological or data model, one that included a generalization with a good fit to empirical results. What makes this model phenomenological is that it includes regularities based upon and abstracted from a set of empirical observations. Phenomenological models are widespread in physics. These idealized models may have predictive and heuristic power, but they do not support non-causal explanations. That a model contains a regularity fitting a limited data set may enable predictions, but it does not support scientific explanations of the values of the independent variables in the regularity, absent an established causal relation. In Galileo's context, his simple pendulum model can predict the period of a pendulum based on its length for a limited range of initial conditions, but it cannot explain the period (or the length!). In Galileo's time there was no global physical theory to determine the relevant nomological possibilities, nor was there an account of forces or causal structure. Those only come with Newton.

Secondly, non-causal explanation requires a robust, detailed local model in which features of entities and dynamics are fully specified.[3] For example, a key premise in the above argument is that the rod is rigid in the Newtonian pendulum model. This constraint does not come from general theory, since there are deformable rods in Newtonian mechanics. Nor will it necessarily come from consideration of the target system being modeled. There are Newtonian models that do not specify the elasticity of their components, and in these sorts of models we simply do not know whether or how components may deform in counterfactual situations. On my approach, all components in a local model essential to an explanation must be fully realized, so to speak. In the case of Newtonian mechanics, solid objects must have determinate values for such properties as mass, position, velocity, acceleration, dimension and elasticity in three dimensions. Obviously, causal approaches to explanation can get away with a lot less—an explanation of the period of a pendulum that appeals to the direction of causation need say nothing about the rigidity of the rod.

I recognize that these requirements of an explanatory global theory and fully-specified models are programmatic at best. What makes a global theory well-defined or robust enough to underwrite such explanations? Kitcher's explanatory unification account famously included precise criteria for measuring the degree of unification and thus the explanatory power of a global theory (1981; 1989). For present purposes, it suffices to say that some global theories in physics are of broad scope, well-confirmed and exhibit a high degree of unification. These features are conducive to understanding and explanation, and *prima facie* these global theories have independent explanatory power. Our focus is on the relation between local model and global theory in contexts in which there is a consensus that the global theory is explanatory. It is also beyond the scope of the present paper to develop precise criteria for what counts as a model sufficiently developed to support non-causal explanation in physics. These criteria will likely vary within physics. The criteria for a fully-specified model in Newtonian mechanics

[3] I am grateful to Michael Strevens for bringing this point to my attention.

(roughly, determinate values for most or all Newtonian properties) will differ from the criteria in other branches of physics.

5. Contextualism in explanation

If I am right, we can debar non-explanatory arguments by appealing to theoretical context to evaluate the truth of the relevant counterfactuals. Theoretical context, then, plays an essential role when assessing the explanatory value of an argument. What sort of contextualism does my account entail? Philosophers of explanation generally take one of two positions on contextualism in explanation. A few, such as Bas van Fraassen and Andrea Woody, argue for what we might call *strong contextualism* (van Fraassen 1980; Woody 2015). On this view, there is a diversity of kinds of relations between an explanans and an explanandum—explanatory relevance relations—and context is always relevant to whether a given answer to a given question is explanatory. What counts as an explanation is determined by the aims of the relevant community. Many, however, endorse a *weak contextualism* that sanctions a much more limited role for contextual factors. As Woodward puts it, our knowledge states and interests influence what we want to explain. These contextual factors determine legitimate explanatory questions in a given scientific community at a given time (Woodward 2003, 229). But once we have determined the question of interest, a genuine explanation is one that bears a unique, context-independent relation to this question. The role for context entailed by my account, which I call *moderate contextualism*, falls somewhere between these approaches.

Bas van Fraassen famously developed a pragmatic approach to explanation that highlights some unpalatable consequences of explanatory contextualism (Van Fraassen 1977; van Fraassen 1980; van Fraassen 1985). Consider a question asked in the context of what van Fraassen calls "accepted background theory and factual information" (1980, 145). A question consists of three components: the explanandum, a contrast class of alternatives, and a relevance relation. The explanans—an answer to this question—bears the relevance relation to the explanandum (and contrast class). For example, "Why is this conductor warped?" may be identified with: an explanandum, "this conductor warped"; a class of alternatives, "one of the other conductors warped"; and a relevance relation, the causal history of the warping. In this case, the explanans describes the causal events leading up to the warping of this conductor.

Notoriously, van Fraassen leaves the relevance relation entirely unconstrained, allowing almost any explanans to count as a good explanation of just about any explanandum. In a given scientific context, for almost any pair of propositions, there is a relevance relation such that the first counts as a bona fide explanation of the second (Kitcher and Salmon 1987). Consider again the question, "Why is this conductor warped?" In a scientific context, accepted background theory and factual information include, for instance, both the causal history of the conductor and the positions of the sun, moon and stars. The former will be part of an explanation of the warping that appeals to causal relevance. On van Fraassen's account, the latter may constitute an explanation of the warping that invokes astrological relevance. This is not to say that the astrological explanation includes any claims of the sort that the positions of heavenly bodies *caused* the warping—that is simply false. Rather, the astrological explanation includes only scientific facts about the positions of these bodies, plus a claim that these facts are explanatorily relevant to the warping. Clearly, this strong contextualism provides no way to exclude pathological cases of explanatory asymmetry.

Van Fraassen frames his account primarily in local terms, to do with the relation between explanans and explanandum. The role of global theory ("accepted background theory and factual information") is limited to ensuring that the explanandum is true, members of the contrast class are all false, and there is some explanatory proposition that bears the relevance relation to the explanandum (1980, 144-5). By contrast, my account entails a unique relevance relation for non-causal dependency explanations, highly constrained by the local conditions of CTE as well as by the laws of the global theory in which the model is embedded. The account appeals to nomological possibility in global theory in order to determine the truth values of counterfactuals that appear in non-causal explanatory relevance relations of the sort articulated in CTE and similar non-causal accounts. The Dependency Condition and the Truth Condition for Model-based Counterfactuals are not context-dependent but rather universal and invariant, at least for non-causal dependency explanations in physics.

Is my approach then consistent with the much more limited role for context endorsed by Woodward and other philosophers of explanation? Not entirely. Weak contextualists, for example, allow that one may explain the warping of this conductor *rather than one of the other exactly-alike conductors* by giving an account of how the warping would have been different if the conductor design had been different (e.g., the design of the cooling fins in this prototype is less efficient than in the other prototypes). One may explain the warping of this conductor *rather than similar conductors made of other metals* in terms of its micro-causal history (e.g., the thermal resistance of the steel conductor is much higher than that of the copper conductor). However, once the question plus contrast class are fully specified, there is only one correct explanation, and this explanation picks out the right sort of relevance relation between explanans and explanandum.

On my *moderate contextualism*, theoretical context plays a somewhat larger role. Context is always relevant to whether "*A* explains *B*," at least in the sort of explanations based on idealized models that are widespread in physics. While *A* and *B* describe elements of a local model, whether *A* is explanatorily relevant to *B*—the explanatory relevance relation—relies on facts that go beyond the local model, namely facts about nomological possibility fixed by global theory. Consequently, *A* may explain *B* in one local model, in the context of one global theory, yet *A* may not explain *B* in the same (or very similar) local model in the context of a different global theory. For example, on a standard Woodward-style dependency account of explanation, it is a context-independent fact that Galileo's pendulum model supports an explanation of the period of a pendulum. On my account, exactly the same derivation of the period of a pendulum is not explanatory in the context of Galileo's original development of his model, but it is explained by a very similar pendulum model in the context of Newtonian mechanics. In this way, a correct explanation is determined not just by the question plus the contrast class, but also by the context fixed by global theory.

Admittedly, the moderate contextualism my solution requires does not come without a cost. An oft-cited virtue of weak contextualism is that, because there is a (context-independent) "best explanation" for any question, the explanatory power of an argument remains an objective fact about it. In this way, weak contextualism enables inference-to-the-best-explanation arguments for realism. Moderate contextualism undermines this sort of argument, since the same argument may be explanatory in one context and not explanatory in another. In my view, this is a price we should willingly pay; there may be good arguments for scientific realism, but they do not rely on inference to the best explanation.

6. Conclusion

This paper proposes a novel solution to problem of explanatory asymmetry for non-causal explanation in physics. My solution requires the following: a robust global theory that fixes nomological possibility; a fully specified local model; and a non-causal dependency account of explanation such as CTE. As we have seen, these requirements set an admittedly high threshold for explanation, but I believe it is a bar that is both descriptively and normatively representative of actual explanatory practice in physics. This paper demonstrates how to resolve the problem of explanatory asymmetry, at least for a simple model of a pendulum. Elsewhere, I have shown how this approach can distinguish explanatory from non-explanatory arguments in an extended case study ([self-reference omitted]). Whether or not this proposal is ultimately successful in resolving other cases of explanatory asymmetry in physics, I hope to have shown that a global condition—the one I have defended or something like it—is needed for a satisfactory philosophical account of explanation.

References

Batterman, Robert and Collin Rice. 2014. "Minimal Model Explanations." *Philosophy of Science* **81**(3): 349-376.

Batterman, Robert W. 2010. "On the Explanatory Role of Mathematics in Empirical Science." *British Journal for the Philosophy of Science* **6**(1): 1-25.

Bokulich, A. 2011. "How Scientific Models Can Explain." *Synthese* **180**(1): 33-45.

Bokulich, Alisa. 2009. "Explanatory Fictions." *Fictions in Science : Philosophical Essays on Modeling and Idealization; Routledge Studies in the Philosophy of Science ;*. Mauricio Suárez. New York, NY: Routledge.

Drake, Stillman. 1989. "Introduction." *Two New Sciences, Including Centers of Gravity and Force of Percussion*. Stillman Drake. Toronto: Wall & Thompson**:** i-xxxv.

Friedman, Micheal. 1974. "Explanation and Scientific Understanding." *Journal of Philosophy* **71**(1): 5-19.

Hempel, Carl G. 1965. *Aspects of Scientific Explanation*. New York: The Free Press.

Hempel, Carl G. and Paul Oppenheim. 1948. "Studies in the Logic of Explanation." *Aspects of Scientific Explanation, and Other Essays in the Philosophy of Science*. New York: The Free Press**:** 245-290.

Jansson, Lina. 2015. "Explanatory Asymmetries: Laws of Nature Rehabilitated." *Journal of Philosophy* **112**(11): 577-599.

---. 2020. "Network Explanations and Explanatory Directionality." *Philosophical Transactions of the Royal Society B: Biological Sciences* **375**(1796): 20190318.

Khalifa, Kareem, Gabriel Doble, et al. 2020. "Counterfactuals and Explanatory Pluralism." *The British Journal for the Philosophy of Science* **71**(4): 1439-1460.

Kitcher, Philip. 1981. "Explanatory Unification." *Philosophy of Science* **48**: 507-531.

---. 1989. "Explanatory Unification and the Causal Structure of the World." *Minnesota Studies in the Philosophy of Science, volume XIII*. Philip Kitcher and Wesley C. Salmon. Minneapolis: University of Minnesota Press**:** 410-506.

Kitcher, Philip and Wesley Salmon. 1987. "Van Fraassen on Explanation." *The Journal of Philosophy* **84**(6): 315-330.

Lange, Marc. 2016. *Because without Cause : Non-Causal Explanations in Science and Mathematics*. New York, NY: Oxford University Press.

---. 2018. "Because without Cause: Scientific Explanations by Constraint." *Explanation Beyond Causation*. Oxford: Oxford University Press.

---. 2019. "Asymmetry as a Challenge to Counterfactual Accounts of Non-Causal Explanation." *Synthese*.

Lewis, David K. 1973. *Counterfactuals*. Cambridge, Mass.: Harvard University Press.

---. 1979. "Counterfactual Dependence and Time's Arrow." *Nous* **13**: 455-476.

Reutlinger, Alexander. 2016. "Is There a Monist Theory of Causal and Noncausal Explanations? The Counterfactual Theory of Scientific Explanation." *Philosophy of Science* **83**(5): 733-745.

---. 2017. "Does the Counterfactual Theory of Explanation Apply to Non-Causal Explanations in Metaphysics?" *European Journal for Philosophy of Science* **7**(2): 239-256.

---. 2017. "Explanation Beyond Causation? New Directions in the Philosophy of Scientific Explanation." **12**(2): e12395.
---. 2018. "Extending the Counterfactual Theory of Explanation." *Explanation Beyond Causation : Philosophical Perspectives on Non-Causal Explanations*. Alexander Reutlinger and Juha Saatsi: Oxford University Press**:** 1 online resource (288 pages).
Rice, Collin. 2015. "Moving Beyond Causes: Optimality Models and Scientific Explanation." **49**(3): 589-615.
Saatsi, Juha. 2016. "On Explanations from Geometry of Motion." *The British Journal for the Philosophy of Science*.
Stalnaker, Robert. 1968. "A Theory of Conditionals." *Studies in Logical Theory: Essays; American Philosophical Quarterly Monograph Series*. James W. Cornman. Oxford: Blackwell. **2:** 98-112.
Strevens, Michael. 2008. *Depth : An Account of Scientific Explanation*. Cambridge, Mass.: Harvard University Press.
---. 2018. "The Mathematical Route to Causal Understanding." *Explanation Beyond Causation*. Oxford: Oxford University Press.
van Fraassen, Bas. 1985. "Salmon on Explanation." *The Journal of Philosophy*: 630-651.
Van Fraassen, Bas C. 1977. "The Pragmatics of Explanation." *American Philosophical Quarterly* **14**(2): 143-150.
---. 1980. *The Scientific Image*. Oxford: Clarendon press.
Woodward, James. 2003. *Making Things Happen : A Theory of Causal Explanation*. Oxford: Oxford University Press.
---. 2018. "Some Varieties of Non-Causal Explanation." *Explanation Beyond Causation : Philosophical Perspectives on Non-Causal Explanations*. Alexander Reutlinger and Juha Saatsi: Oxford University Press**:** 1 online resource (288 pages).
Woody, Andrea I. 2015. "Re-Orienting Discussions of Scientific Explanation: A Functional Perspective." *Studies in History and Philosophy of Science Part A* **52**: 79-87.