

Humean laws, explanatory circularity, and the aim of scientific explanation

Chris Dorst¹ 

Published online: 18 July 2018
© Springer Nature B.V. 2018

Abstract One of the main challenges confronting Humean accounts of natural law is that Humean laws appear to be unable to play the explanatory role of laws in scientific practice. The worry is roughly that if the laws are just regularities in the particular matters of fact (as the Humean would have it), then they cannot also explain the particular matters of fact, on pain of circularity. Loewer (Philoso Stud 160:115–137, 2012) has defended Humeanism, arguing that this worry only arises if we fail to distinguish between scientific and metaphysical explanations. However, Lange (Philoso Stud 165:255–261, 2013, Synthese 195:1337–1353, 2018) has argued that scientific and metaphysical explanations are linked by a transitivity principle, which would undercut Loewer’s defense and re-ignite the circularity worry for the Humean. I argue here that the Humean has antecedent reasons to doubt that there are any systematic connections between scientific and metaphysical explanations. The reason is that the Humean should think that scientific and metaphysical explanation have disparate aims, and therefore that neither form of explanation is beholden to the other in its pronouncements about what explains what. Consequently, the Humean has every reason to doubt that Lange’s transitivity principle obtains.

Keywords Laws of nature · Humeanism · Lewis · Scientific explanation · Metaphysical explanation · Transitivity

✉ Chris Dorst
dorst@live.unc.edu

¹ Washington University in St. Louis, 17 Lincoln Road, Salem, MA 01970, USA

1 Introduction

Humean accounts of laws of nature maintain that the laws are certain sorts of regularities in the particular matters of fact. For example, according to Lewis's Best System Account ("BSA"), the laws are the regularities of that systematization of the particular matters of fact which best balances the desiderata of simplicity and strength.¹ A common objection to Humean views like the BSA is that they render the laws unable to play the explanatory role that is traditionally attributed to them in scientific practice, wherein the laws are routinely used to explain the occurrences of natural phenomena. The worry, roughly, is that if the laws are just *regularities* in the particular matters of fact, then they cannot also *explain* the particular matters of fact, for that would seem to be dangerously close to a circular explanation of the particular matters of fact by themselves. As Maudlin puts it:

If the laws are nothing but generic features of the Humean Mosaic, then there is a sense in which one cannot appeal to those very laws to explain the particular features of the Mosaic itself: the laws are what they are in virtue of the Mosaic rather than vice versa. (Maudlin 2007, p. 172)²

Recently there has been a renewed interest in this sort of argument against Humeanism. In his (2012), Loewer defends Humeanism from the charge of explanatory circularity, suggesting that the worry only arises if we fail to distinguish between scientific and metaphysical explanations: the laws *scientifically* explain the particular matters of fact, but the particular matters of fact *metaphysically* explain the laws (according to the Humean). Loewer thinks that the fact that there are two different kinds of explanation here precludes any worrying kind of circularity.

Loewer's proposal has received some intense scrutiny from the likes of Lange (2013, 2018), Emery (forthcoming), and Shumener (2017), but ultimately I think that Loewer is right: the differences between metaphysical and scientific explanations are what saves the Humean. However, in order to see this, we have to step back from the particulars of these debates and consider why scientists *care* about explanatory considerations in the first place. Ultimately, my suggestion will be that the aims of explanation—our *motivations* for seeking explanatory theories—are very different in science than they are in metaphysics. Or at the very least, Humeans should think so. And these different aims are going to undercut any sort of systematic connection between scientific and metaphysical explanations, thus rendering the charges of explanatory circularity unsuccessful. In brief, the explanatory circularity objection against Humeanism succeeds only if we assume that the aims of scientific and metaphysical explanation align in a way that the Humean has good reasons to reject.

This paper proceeds as follows. In Sect. 2, I explicate the circularity objection against Humeanism in more detail. Then in Sect. 3, I review a recent Humean theory that, I will argue later, is particularly well positioned to avoid the circularity

¹ See Lewis (1973, 1986, 1994).

² Armstrong expresses a similar worry in his (1983, p. 102).

objection. In Sect. 4, I consider the question of why scientists care about finding theories that possess explanatory virtues. I will suggest that Humeans should understand the explanatory virtues not, in the first instance, as a guide to *truth*, but rather as ways of making our scientific theories more *predictively useful*. On the basis of this discussion, in Sect. 5 I will argue that the Humean should conceive of the aims of scientific and metaphysical explanation very differently, and I will show how this undercuts the circularity worry. I conclude in Sect. 6.

2 The circularity problem for Humeanism

In its most basic form, the circularity problem for Humeanism can be stated quite simply. Dating back to at least Hempel and Oppenheim's (1948) Deductive Nomological (DN) model of scientific explanation, philosophers have recognized that the laws of nature can be used to explain their instances. If it is a law that all *F*s are *G*s, then we can explain why some particular *F* is *G* by appealing to that law. More realistically, consider the Lorentz force law, which states that a charged particle traversing a magnetic field will experience a "Lorentz force" perpendicular to its direction of motion. Now suppose that we observe an electron traversing a magnetic field, and as it does so, it curves off of its original trajectory, indicating that it is experiencing a force. Call this event "*e*". We can explain why *e* occurred by appealing to the Lorentz force law: the electron is a charged particle, and as it traverses the magnetic field, the Lorentz force law states that it will experience a Lorentz force. The event *e* is thus both an *instance of*, and *explained by*, the Lorentz force law.³

The problem for the Humean is that she conceives of the Lorentz force law itself as essentially just the *regularity* that every time a charged particle enters a magnetic field, it experiences a Lorentz force perpendicular to its direction of motion.⁴ And if that is what the law amounts to, then part of what *makes it a law* is the fact that *e* occurs. Thus it looks as though, on the Humean view, *e* helps to explain why the Lorentz force law is a law, while at the same time the Lorentz force law can be used to explain why *e* obtains. And this seems problematic, as if *e* is helping to explain why *e* itself occurred.⁵

³ Two clarifications. First, the claim is not that *any* fact "covered" by a law is explained by it. For example, it would be incorrect to say that the Lorentz force law, together with the strength of the magnetic field and the magnitude of the Lorentz force experienced by the electron, explains why the electron has the charge that it does. (One of the main problems with the DN model was its inability to account for this sort of explanatory asymmetry.) Second, we might just as well have considered the law explaining the regularity *as a whole*, but for simplicity I have focused on the explanation of the particular event *e* that is a part of that regularity.

⁴ Of course, this regularity will typically have to exhibit additional characteristics in order to count as a law. For example, according to the BSA, it must be part of a systematization of *all* the particular matters of fact that best balances simplicity and strength.

⁵ That is not to say that *all* cases of self-explanation are unacceptable. Perhaps in some outré time travel scenarios, for example, we would have to admit that self-explanation can occur. But surely we should not allow that self-explanation is *ubiquitous*, as it would appear to be on the Humean view. After all, our

As I mentioned before, Loewer (2012) has suggested that this circularity worry can be circumvented by noting that there are two different kinds of explanation occurring here. On the one hand, the Lorentz force law *scientifically* explains the occurrence of *e*, but on the other hand, *e* helps to *metaphysically* explain the lawhood of the Lorentz force law. Loewer does not provide an in-depth analysis of the differences between these kinds of explanation, but he does think that they are significant enough to preclude the circularity worry. Thus, Loewer would say that in our example, *e* is not helping to explain itself in any univocal sense, because there are two very different kinds of explanation at work.

However, Loewer's suggestion has been challenged by Lange (2013), who suggests that scientific and metaphysical explanations are linked by the following transitivity principle:

TP: If *X* scientifically explains [or helps to scientifically explain] *Y* and *Z* metaphysically explains [or helps to metaphysically explain] *X*, then *Z* scientifically explains [or helps to scientifically explain] *Y*.⁶

Lange motivates TP with a number of compelling examples, many of which have the same character. One such example (*ibid.*, p. 257) is the expansion of a balloon, which is scientifically explained by certain laws as well as the fact that the balloon's internal pressure is greater than the external atmospheric pressure. Then, Lange suggests, since the fact that the internal pressure is greater than the external pressure is grounded in—i.e. metaphysically explained by—the particular values of both the internal and external pressures, it follows by the transitivity principle that the particular values of those pressures help to explain why the balloon expands. And this seems correct.

So the basic thought is that if *Z* metaphysically explains *X*, then *Z* provides the metaphysical basis for *X*, so *Z* should have all of the explanatory power that *X* does. Thus, *Z* should be able to scientifically explain anything that *X* can scientifically explain. Now if TP is true, then Loewer's gambit is no longer going to work. To see this, consider our example with the Lorentz force law explaining event *e*, which was supposed to be a generic type of scientific explanation. In the above explication of TP, plug in the Lorentz force law for *X*, event *e* for *Y*, and event *e* again for *Z*. Thus we have the Lorentz force law scientifically explaining event *e*, and event *e* helping to metaphysically explain the Lorentz force law. So it follows by TP that event *e* helps to scientifically explain event *e*. Again we are left with a worrisome case of self-explanation.

Of course, it is open to the Humean to just reject TP, but doing so without independent motivation would be objectionably *ad hoc*. In that vein, several

Footnote 5 continued

choice of the Lorentz force law, and the event *e*, was relatively arbitrary. The same considerations would hold no matter which law we chose, and no matter which instance we picked. [Lange (2018, p. 1338) also allows that the prohibition on self-explanation may be violated in some exotic cases.]

⁶ Lange actually frames his principle in terms of grounding rather than metaphysical explanation, but it makes the connection with Loewer's argument most explicit if we frame it in terms of metaphysical explanation.

philosophers, such as Hicks and van Elswyk (2015), Miller (2015), and Marshall (2015), have recently argued against TP on various grounds. For example, it has been suggested that TP has unacceptable consequences in cases where macrolevel phenomena are multiply realizable, and also that it fails to account for the fact that a variety of different types of relations can underwrite scientific explanations. Unfortunately for the Humean, Lange (2018) offers compelling responses to all of these objections. Part of his response involves modifying TP slightly to allow for contrastive effects in the explanans and explanandum, but otherwise the principle survives largely intact.

In the rest of this paper, I am going to argue that the Humean has compelling independent motivations for rejecting TP or any of its potential modifications. Indeed, I will argue that the Humean has good reason to think that there are *no systematic connections whatsoever* between scientific and metaphysical explanation. To see this, we first have to review some recent developments in the Humean tradition.

3 LOPP-style Humeanism and the Best Predictive System Account of laws

Some of the most promising Humean theories on the market right now are developments of Lewis's BSA. The particular theories I have in mind are Hicks's (2017) "Epistemic Role Account," Dorst's (2018) "Best Predictive System Account," and Jaag and Loew's (forthcoming) "Cognitive Usefulness Account."⁷ All three of these theories can be viewed as starting from the concept of a "Limited Oracular Perfect Physicist," an idea that traces back to Hall (ms). There, Hall suggests that we think of the BSA as a sort of ideal observer view of laws:

[L]et us imagine someone whom I will call a Limited Oracular Perfect Physicist. What makes our LOPP a perfect physicist is that, given as evidence any information about the world, she is perfectly able to judge what hypotheses about the fundamental physical laws are most strongly supported by that evidence. What makes her oracular is that she has, as evidence, quite a lot of information about the world. But not, of course, all information: else her job would be too easy. (For example, she doesn't directly receive as evidence information about what the fundamental laws are.) Specifically, we will suppose that what she has available to her as evidence are all the [particular matters of fact] [...] The [...] idea, then, is roughly that the laws are whatever she says they are. (Hall [ms], p. 15)

So the suggestion is that the laws are the result of applying a certain *operation* to the totality of the particular matters of fact, namely, that operation (whichever one it is)

⁷ Another noteworthy development of the BSA is the "Better Best System Account" of Cohen and Callender (2009). The developments I am focusing on here are, I think, largely compatible with the Better Best System view, though they focus on different aspects of the Humean metaphysics.

that physicists *themselves* apply when they are investigating the laws. Theories in this tradition are variants of what I call “LOPP-style Humeanism”.⁸

Another way of thinking about LOPP-style Humeanism is that it elevates the epistemic standards used by physicists in discovering the laws to the status of *constitutive* standards that are part of what it is *to be* a law. And one reason that this is a particularly attractive maneuver is that it immediately presents us with an explanation of why those standards are reliable guides to the laws. In other words, LOPP-style Humeanism *reverse engineers* the laws from physicists’ epistemic standards for lawhood, thereby obviating the need to provide an independent justification for the reliability of those standards.

By doing this, we replace one justificatory task with another. Whereas before we wanted to know why physicists’ epistemic standards were a reliable guide to the laws, *now* what we should want to know is why creatures like us would *care* about implementing those standards in the first place. To see what I mean, it will help to have a more concrete proposal for these epistemic standards in mind. Thus consider Lewis’s BSA. If Lewis was right, those standards would consist of a balance of simplicity and strength. Now one question that inevitably arises in these sorts of context—i.e. whenever simplicity is treated as a guide to theory choice—is why simplicity should track the truth. The uncomfortable answer here is that, while philosophers have long searched for epistemic justifications for an appeal to simplicity, success has been hard to come by. It is just not particularly clear why simplicity should be truth-conducive. But if simplicity is rightly understood, not just as an epistemic guide to lawhood, but as a constitutive standard of lawhood, then the question of why simplicity is a reliable guide to truths about the laws lapses: if simplicity is part of what makes something a law, then it becomes much clearer why it is a good idea to use simplicity as part of our epistemic guide to the laws. The question then becomes why creatures like us should be interested in a set of principles that is partly distinguished based on its simplicity. Our justificatory task has thus switched from an epistemic one—making sense of why these standards are a reliable guide to the laws—to a pragmatic one: making sense of why these standards produce a set of principles that is *useful* to creatures like us.

The answer to this latter question is going to depend on which standards scientists actually use to investigate the laws, and particularly on what sorts of principles the *combination* of those standards tends to produce. If Lewis was right, those standards are meant to produce a set of principles that constitutes an *efficient summary* of the totality of the particular matters of fact.⁹ But as Hicks (2017), Dorst (2018), and Jaag and Loew (forthcoming) all argue, an efficient summary of the particular matters of fact is not likely to be particularly useful to creatures like us. For example, an efficient summary may end up giving us a lot of *statistical* facts, such as facts about the average lifespan of stars in this universe, or facts about the standard deviation of galaxy diameters. Facts like this, interesting though they are,

⁸ Versions of LOPP-style Humeanism are also developed in Miller (2014) and Bhogal and Perry (2017).

⁹ For explication of this point, see, e.g., Lange (2009, pp. 101–102), Albert (2015, pp. 23–24), and Beebe (2000, p. 574).

are not particularly useful for predictive purposes, nor do they resemble the dynamical form of actual putative laws in scientific practice.¹⁰

With these points in mind, the theories advanced by Hicks, Dorst, and Jaag and Loew can all be seen as proposing standards that differ from Lewis's balance of simplicity and strength, with the goal of producing principles that (i) are more useful to creatures like us, and (ii) exhibit the general form of putative laws that we find in scientific practice. It will prove useful for our later discussion to examine these views in a bit more detail. Here I will focus on my Best Predictive System Account ("BPSA") for ease of exposition.

The BPSA maintains that the laws are the regularities of the systematization of the totality of the particular matters of fact which is maximally predictively useful to creatures like us. Thus, the standards that figure into the account are designed to produce a set of predictively useful principles. The goal then becomes to characterize the features that would be possessed by such principles. And the rough idea is that predictively useful principles would give us a great deal of information about the temporal evolution of quasi-enclosed subsystems of the universe, without relying on facts that creatures like us are usually not in a position to ascertain. Consequently, the standards that figure into the BPSA include straightforward desiderata like extensive dynamical implications and wide applicability, and the predictive utility of these desiderata is straightforward. But the BPSA also includes desiderata such as spatial and temporal locality, because creatures like us tend not to be able to ascertain facts that are radically far-removed in space and time from the systems we are trying to make predictions about. Similarly, the BPSA also includes desiderata regarding spatial, temporal, and rotational symmetries. The thought here is that since creatures like us are not innately able to locate and orient ourselves in space and time, principles that are predictively useful to us would not require that sort of information in order to be applied. The fact that these various features are reflected in the principles that physicists (past and present) have regarded as laws suggests both that these desiderata are indeed part of physicists' epistemic criteria for lawhood, and also more generally that we are on the right track in thinking about the laws in terms of predictive utility.

In sum, when physicists theorize about the laws, they implement a variety of different standards. The BPSA makes sense of this fact by suggesting that these standards are constitutive of lawhood and that they all serve the common purpose of maximizing the predictive utility of the resulting set of principles. Thus, while the standards are truth-conducive, it can be misleading to think that their *fundamental* purpose is to serve as a guide to truths about the laws. Rather, the fundamental purpose of implementing these standards is to produce a system of laws that is maximally predictively useful, and whose truth is, in a sense, *derivative* upon that predictive utility. The best predictive systematization of the particular matters of fact constitutes the truth about the laws of nature, and it does so *because* it is the best such systematization.

¹⁰ C.f. Section 3 of Dorst (2018) for a more comprehensive discussion of this point.

In this author's opinion, theories like the BPSA constitute some of the most promising recent developments in the Humean tradition. In the next section, I am going to consider how a Humean sympathetic with the BPSA should conceive of the aim of explanation in science. In particular, I am going to consider the following question: Why do scientists care whether their theories possess explanatory virtues?

4 Why care about explanatory virtues?

Why do we care about explanatory considerations when we do science? What is it about explanations, and explanatory theories, that makes them scientifically valuable? One answer is just that explanations are *intrinsically* valuable: they are intellectual achievements that have value much like any other intellectual achievement. Strevens (2008) suggests just this sort of explanation of the value of explanation:

If science provides anything of intrinsic value, it is explanation. Prediction and control are useful, and success in any endeavor is gratifying, but when science is pursued as an end rather than as a means, it is for the sake of understanding—the moment when a small, temporary being reaches out to touch the universe and makes contact. (2008, p. 3)

Strevens thus views explanations as valuable in their own right. By contrast, other philosophers have argued that the value of explanation lies in its connection to practical concerns such as prediction. Hempel and Oppenheim's (1948), for example, famously suggested as much:

It is this potential predictive force which gives scientific explanation its importance: only to the extent that we are able to explain empirical facts can we attain the major objective of scientific research...to anticipate new occurrences and to control, at least to some extent, the changes in our environment. (1948, p. 138)

Similarly, Quine and Ullian maintain that “the hypotheses we seek in explanation of past observations serve again in the prediction of future ones. Curiosity thus has survival value, despite having killed a cat” (1970).

Part of the trouble that arises in trying to answer the question of why scientists care about explanations is that the question itself is ambiguous. On one reading, the question is asking why scientists *in fact* find explanations valuable, whereas on another reading, the question is asking why is it *prudent* for scientists to care about explanatory considerations. On the first reading, an answer along the lines of Strevens's is undoubtedly often right: individual scientists often find intrinsic value in explanatory considerations simply because they have been *trained* to find them valuable, elegant, etc. But on the second reading, it is not so clear that Strevens's answer is particularly helpful. If I want to know why it is prudent for us to care about explanatory considerations, replying that explanations are inherently valuable is not especially elucidating.

It is this second reading of the question that I am concerned with here. I am going to argue that the Humean should answer this question by following Hempel and Oppenheim's original suggestion, and thinking of explanatory considerations as derivative upon predictive considerations. Of course, Hempel and Oppenheim's symmetry thesis—that every explanation is a potential prediction and vice versa—no longer appears tenable. But that doesn't preclude the possibility that there is some less direct connection between explanation and prediction. To see this connection, I propose to come at the issue from the side, as it were, by asking why scientists care about their theories possessing *explanatory virtues*. If we can figure out what is valuable about the explanatory virtues, then we will have the beginnings of an answer to why it is prudent for scientists to value explanatory considerations in general.

Of course, any exploration of the value of the explanatory virtues has to begin with a discussion about what sorts of theoretical virtues *count* as explanatory virtues. Unfortunately there is no widespread agreement on either the terminology used to name these virtues, nor on the exact list of theoretical virtues that are rightly counted as explanatory. Nevertheless, there are a few virtues that appear most commonly (under some guise or other) in discussions of explanatory virtues. Here I will consider four of the most prominent ones.

First, simplicity. Many theorists have suggested that theories that possess a greater degree of simplicity have greater explanatory power.¹¹ However, there are multiple virtues that go by the same name here. One, which I'll call "conceptual simplicity," concerns the nature of the theoretical apparatus appealed to as an explanation. Theories that make use of a more straightforward conceptual framework than others possess the virtue of conceptual simplicity to a greater degree. For example, the picture of ideal gases as a collection of non-interacting point particles possesses a considerable degree of conceptual simplicity, and would be preferable on these grounds, other things being equal, to a theory with more abstruse theoretical postulates that nevertheless has the same observable consequences.

Another type of simplicity is what I'll call "extrapolative simplicity." Extrapolative simplicity concerns the fitting of curves (i.e. theories) to data points. Simpler curves are expressible by equations of lower degree, and such curves possess more extrapolative simplicity than curves that must be expressed using higher degree equations. Here again, other things being equal, theories with more extrapolative simplicity are preferable to theories with less.

Another virtue that is commonly characterized as explanatory is *unificatory power*. A theory possesses more unificatory power if it is applicable, in a uniform manner, to phenomena in very different domains.¹² A classic example here is Newton's theory of gravity, which was able to treat the movements of earthly bodies and the planets in a unified manner, thus constituting a significant improvement over

¹¹ Simplicity is discussed as an explanatory virtue by philosophers such as Kuhn (1977), Quine and Ullian (1970), Thagard (1978), Psillos (2002), and Lipton (2004).

¹² Virtues corresponding to unificatory power go by a variety of names, including "generality" (Quine and Ullian 1970), "consilience" (Thagard 1978), and "unification" (Psillos 2002).

the heterogeneous conjunction of Galileo's law of falling bodies and Kepler's laws of planetary motion.

Finally, another theoretical virtue often counted as explanatory is *precision*. The virtue of precision amounts to the fact that if a theory T_1 gives more details about a given phenomenon than its rivals T_2 , T_3 , etc., then, other things being equal, T_1 is preferable, as an explanation of that phenomenon, to T_2 , T_3 , etc.¹³ These details can take a variety of forms, but the rough thought is that the more information a theory provides about either the *production* of the phenomenon of interest, or about various aspects of the phenomenon itself, the more that theory possesses the virtue of precision.

There are, of course, additional explanatory virtues often cited in the literature, but the four I have mentioned here—conceptual simplicity, extrapolative simplicity, unificatory power, and precision—are some of the most common.

What, then, is the value of these virtues? One place that this question gets addressed head-on is in discussions about the status of inference to the best explanation (“IBE”) in science, particularly in the debates about whether IBE is compatible with Bayesianism. It will be worth briefly reviewing some of these debates here. Supporters of IBE maintain that we should believe the theory that scores the best in terms of its explanatory virtues.¹⁴ Thus a theory T_1 receives a higher credence than theories T_2 , T_3 , etc. just in case T_1 exhibits the explanatory virtues to a higher degree than T_2 , T_3 , etc. The suggestion, then, seems to be that the explanatory virtues are *truth-conducive*: theories that exhibit these virtues to greater degrees are more likely to be true.

A number of theorists have attempted to argue that explanatory considerations, such as the virtues we have just considered, are truth-conducive. Thus, Henderson (2014) develops a view that she calls “emergent compatibilism,” arguing that IBE essentially “emerges” from the framework of Bayesianism. As Henderson puts it, “a Bayesian who adopts constraints on her probabilities, which are reasonable on her own terms, would end up favoring more explanatory theories” (2014, p. 12). And since Bayesianism is concerned solely with the confirmation of hypotheses (and not other other sorts of pragmatic or instrumental goals), Henderson's suggestion that explanatory virtues emerge from Bayesian confirmational considerations implies that the explanatory virtues are confirmationally relevant, i.e. truth-conducive.

Similarly, both Lipton (2004) and Dellsén (2017) argue that IBE should be conceived as a heuristic that *approximates* probabilistic Bayesian reasoning for cognitively limited creatures. The rough thought is that it is easier for creatures like us to make rough-and-ready judgments about the degree to which hypotheses possess explanatory virtues than it is for us to go through detailed probabilistic reasoning in accord with Bayes' theorem, but ultimately these are just two different methods for helping us track the epistemic probabilities of the relevant

¹³ Precision is discussed as an explanatory virtue by Psillos (2002), Lipton (2004), and Cabrera (2017), among others.

¹⁴ Philosophical discussions of IBE trace back at least to Harman (1965). Lipton (2004) gives probably the most extensive contemporary treatment.

hypotheses.¹⁵ Lipton and Dellsén thus also conceive of the explanatory virtues as fundamentally truth-conducive.

By contrast, a number of other theorists have argued that IBE is incompatible with Bayesianism because at least some of the explanatory virtues do *not* track the truth, and therefore are not confirmationally relevant. The most famous such criticism is mounted by Fraassen (1989), who argues that IBE would require agents to assign higher probabilities to more explanatory hypotheses than is required by the Bayesian framework. This sort of criticism is most common regarding the explanatory virtue of simplicity. For example, consider the following remark by Fumerton, which is characteristic of the general skepticism of the truth-conduciveness of simplicity:

[A simpler theory] is certainly more desirable than its competitors in the sense that it would be *nice* if it turned out to be true. But this not being the best of all possible worlds (some theologians aside) what would be nice is not always so. (1980, p. 596)

These sorts of worries about the truth-conduciveness of simplicity would seem to apply to both the conceptual and extrapolative variants, although many theorists are not always explicit about the distinction. In the case of conceptual simplicity, the worry is that it is unclear why we are justified in expecting nature to be describable in conceptually simple terms. After all, why couldn't nature just be massively complex? And in the case of extrapolative simplicity, the worry is the familiar one of curve fitting: given that there are an infinity of curves that fit the observed data, on what basis do we justify a preference for simpler curves?

Similar sorts of worries sometimes arise about the other explanatory virtues. For example, Cabrera (2017) argues that the virtue of precision cannot be appropriately conceived as truth-conducive, for the simple fact that a more precise hypothesis provides more information, and thus is *less* likely to be true, on probabilistic grounds, than a less precise hypothesis. In the most extreme case, we could compare the probabilities of two hypotheses H_1 and H_2 , where H_2 is strictly a precisification of H_1 . Thus we would have $H_2 \models H_1$ but $H_1 \not\models H_2$, and therefore $P(H_1) \geq P(H_2)$. For example, suppose I am trying to find an explanation for the fact that the grass on my lawn is wet. I entertain two hypotheses:

H_1 : It rained last night.

H_2 : It rained for exactly 3 h, 27 min, and 14 s last night.

While both hypotheses would explain the fact that the grass is wet, H_2 contains strictly more information than H_1 , so even though H_2 possesses a greater degree of precision, it is also less likely to be true.¹⁶

¹⁵ It is worth clarifying that Dellsén thinks IBE only functions as a guide to *comparative*, not absolute, probability values.

¹⁶ One might worry that a comparison of H_1 and H_2 on explanatory grounds is illegitimate because one entails the other. In this vein, Dellsén (2016) argues that the hypotheses used in any particular application of IBE ought to be incompatible. But even if that is correct, precision can hardly be regarded as truth-conducive. Imagine comparing H_1 with H_3 : My neighbor's sprinkler system sprayed my yard last night

There are, then, significant debates about how to understand the value of many of the explanatory virtues in science, and currently there is no clear consensus on the right approach. Now the pertinent question for our purposes is this: How should the *Humean* seek to understand them?

The basic suggestion here is that the Humean should pull the same maneuver that she does in the case of laws, and attempt to *reverse engineer* the truth of our scientific theories from the epistemic standards used to decide among them. Thus, the Humean should conceive of the explanatory virtues not merely as epistemic guides to the correct theory, but as partly *constitutive* of the truth of the correct theory. In other words, the Humean should conceive of the explanatory virtues using the framework of the Best Systems approach: the true theory is the one that best satisfies the various explanatory virtues, and that theory is true *because* it best satisfies those virtues.

Conceiving of the explanatory virtues in this way comes with an immediate payoff, but it also gives rise to a novel challenge. The payoff is that we no longer require an account of why the explanatory virtues are truth-conducive, for if their satisfaction is just part of what makes the true theory true, then their epistemic justification is trivial. We thereby avoid all the worries about whether, and *why*, virtues like simplicity, precision, and unification should be expected to track the truth. Instead we are faced with a different challenge: accounting for why creatures like us would *care* about finding a theory that best satisfies these virtues. Just as with LOPP-style Humeanism, then, the challenge shifts from providing an *epistemic* justification to a *pragmatic* one.

The most natural way for the BPSA-friendly Humean to provide such a pragmatic justification is to seek to understand the explanatory virtues as contributing to the *predictive utility* of the resulting theory, much like she conceives of the standards of lawhood as being designed to produce a collection of predictively useful principles. The Humean thus follows the likes of Hempel and Quine in suggesting that explanatory considerations have predictive value. Of course, where the rubber meets the road is in the attempt to explain how each of the explanatory virtues itself contributes to the predictive utility of the resulting theory. To be sure, this is no small task, but I think there are reasons for optimism. For example, here is how the Humean could try to argue that the four explanatory virtues mentioned above are helpful for predictive purposes:

1. *Conceptual simplicity* Consider two theories, T_1 and T_2 , which are empirically equivalent (as far as we know), yet T_1 is conceptually much simpler than T_2 . Why, in that case, is T_1 more useful for predictive purposes?

Footnote 16 continued

for exactly 3 h, 27 min, and 14 s. As explanations of the wetness of the lawn, H_1 and H_3 are incompatible, but the increased precision of H_3 drags down its probability.

On the other hand, maybe the suggestion is that rival hypotheses in an application of IBE must be, in some sense, “complete” explanations, in that they have to be both incompatible and equally precise. (Indeed, this seems closer in spirit to Dellsén’s own suggestion.) In that case, it is hard to see how precision could still be an explanatory virtue, since all hypotheses compared in any application of IBE will be equally precise.

There are at least two reasons. First, theories that are conceptually simpler are easier to use to make *rough-and-ready* predictions. For example, consider someone unfamiliar with the macroscopic behaviors of gases, who only knows the simple postulates of the molecular theory of gases. Even to someone in such a position, it should not be surprising that heating a container of gas will increase the pressure inside the container. For it is not too difficult to reason that heating the gas will make the constituent molecules move faster, therefore causing more forceful and frequent collisions with the walls of the container. Of course, none of this is knowable a priori, but it does accord with our intuitive expectations. On the other hand, a theory with more abstruse theoretical postulates (that nevertheless makes the same predictions as the molecular theory of gases) is unlikely to lend itself to such easy predictions. In other words, it is not as easy to use a more complicated theory to get a rough idea of how the system of interest is going to behave.

Second, conceptually simpler theories are more predictively useful because they lend themselves to more straightforward modifications when confronted with recalcitrant evidence. Suppose a given theory makes some incorrect predictions, and we therefore decide that we are going to have to modify the theory. This task can be immensely easier if it is a simpler theory that is easy for us to grasp and manipulate. Since we are more easily able to understand how the constituent elements of a simpler theory combine to produce the various predictions, we are in a better position to know which elements might have to be changed when the predictions are not borne out by the evidence.

2. *Extrapolative simplicity* How could predictive concerns underwrite a preference for theories with a greater degree of extrapolative simplicity? One of the more prominent justifications of simplicity considerations in the context of curve fitting is found in the literature on the Akaike Information Criterion (AIC). The goal of this criterion is to select a curve that minimizes the chances of fitting “noise” in the observational data. There are some technical discussions of the AIC in the literature [see, e.g. Forster and Sober (1994), Kieseppää (1997)], but we need only the general idea here. The thought is that whenever we gather observational data (say, of the relationship between two variables x and y), our results are the product of two factors: the true underlying relationship between those variables, and observational “errors” or “noise.” Thus if we select a curve that fits the data too exactly, it will be weighing the noise more than the true relationship, and therefore it is likely to lead to incorrect predictions. So the rationale for selecting a simpler curve is that it is likely to be more predictively accurate. As Woodward (2014) puts it:

The simplicity term does not enter into the formula for AIC because simplicity is being accorded some intrinsic value...but rather because the goal of maximizing expected predictive accuracy in itself requires inclusion of this term. In other words, the justification for the simplicity term is entirely in terms of the fact that it is a means to the goal of predictive accuracy. (p. 116)

This, to me, sounds like exactly the kind of pragmatic justification of extrapolative simplicity that the Humean is looking for.

3. *Unificatory power* Why are theories with more unificatory power more predictively useful? Why, for example, was the Newtonian theory of gravity preferable on predictive grounds to the heterogeneous conjunction of Galileo's law of falling bodies and Kepler's laws of planetary motion? Remember that theories with more unificatory power are able to account for disparate phenomena using the *same* theoretical framework; Newton's theory of gravity uses the law of universal gravitation to account for the behaviors of both earthly and heavenly bodies. By contrast, the theoretical frameworks of Galileo's law of falling bodies and Kepler's laws of planetary motion are quite different from each other. The key, then, is to consider how these theories would account for the behaviors of objects that fall on the *border* of both subdomains. For example, how would the conjunction of Galileo's and Kepler's laws treat a projectile that starts in outer space and then gradually falls to earth, such as a comet or meteor? There is no clear direction from this heterogeneous theory about how to treat such phenomena: do we use Kepler's laws, or Galileo's laws, or some mixture of them? But Newton's theory is unambiguous in its instructions: apply the law of universal gravitation, just like you would for any other object. The thought, then, is that theories with more unificatory power are more useful for predictive purposes because they give unambiguous instructions for generating predictions about objects lying on the borders of multiple subdomains of the theory. Radically disjunctive theories, on the other hand, will be correspondingly more equivocal about how one should treat objects at the intersections of their various subdomains.¹⁷
4. *Precision* Why are theories that provide more information about the phenomenon of interest—either about the factors leading to its production or about the phenomenon itself—more predictively useful? This question almost answers itself. If I tasked you with making predictions about some novel physical system that you had never seen before, clearly you would prefer a theory that gives you more information about the system in question. This is roughly like the difference between being told (a) “The system will do something in the future,” and (b) “The system will evolve in exactly this way...”. If your concern is with making predictions, a theory with more precision is better, other things being equal.

Conceiving of the explanatory virtues as methods of designing a theory with higher predictive utility thus looks to be a promising strategy. Of course, there are a

¹⁷ van Fraassen (1980, p. 86) gives a similar explanation of the value of unification, suggesting that the aim of empirical adequacy requires our theories to be able to cover phenomena on the borders of their subdomains.

Additionally, the current state of fundamental physics may provide support for this predictive conception of the value of unification. Specifically, there are radically different physical theories governing Planck-scale and high energy regimes, and these theories seem to generate incompatible predictions about events in the early universe. The search for a grand unified theory may then be viewed as arising, not out of a mere preference for the aesthetic value of unification, but out of the desire to have a theory that accurately predicts the behavior of events in domains where our current theories generate incompatible predictions. Thanks to an anonymous referee for suggesting this point.

number of other explanatory virtues that I have not mentioned here. Cabrera (2017) alone mentions the additional virtues of scope, mechanism, fit with background data, and empirical adequacy, none of which I have discussed. To be sure, then, the Humean would have additional work to do in trying to explain how all of these explanatory virtues could be construed as conducive to predictive utility. But the key point is that *everyone* has their work cut out for them here. After all, we have already seen that the project of conceiving of the explanatory virtues as *truth*-conducive encounters a number of significant obstacles. And it is by no means obvious that the project of conceiving of these virtues as prediction-conducive is any worse off. Indeed, the plausibility of the above suggestions about the predictive value of simplicity, unificatory power, etc. indicates that we may actually have an *easier* time thinking of the virtues in this manner.

In sum, our original question was this: Why do scientists care that their theories exhibit explanatory virtues? The answer I am suggesting on behalf of the Humean is that explanatory virtues are valuable because they contribute to the predictive utility of the resulting theory. This is consonant with some promising recent developments in the Humean tradition, which develop an understanding of laws primarily in terms of their predictive utility. It is also important to note that this suggestion is perfectly compatible with the thought that the explanatory virtues are truth-conducive, though it does require that they be conceived as not *primarily* truth-conducive. Rather, they are primarily *prediction*-conducive, and they track the truth, on the Humean picture, only because the true theory *just is* the one that exhibits the greatest predictive utility.¹⁸

5 Explanation in science and metaphysics

The Humean picture of explanation that I have just been sketching is in broad agreement with a conception of the aims of science that has recently fallen into disrepute. In the early twentieth century, positivistically-minded philosophers of science tended to regard explanation with suspicion. As Salmon (1999) puts it, they were inclined to view explanation as “beyond the scope of science, in such realms as metaphysics and theology” (1999, 338). Karl Pearson, for example, put it bluntly: “Nobody now believes that science *explains* anything” (1911, xi, emphasis in

¹⁸ It is worth noting that, in principle, the Humean has another option open to her: she can view the explanatory virtues as *only* conducive to predictive utility, and *not* thereby also truth-conducive. On this picture, IBE is inherently concerned with fashioning theories that are predictively useful to us, without any regard to whether or not they are true.

While it is open to the Humean to adopt this conception of the explanatory virtues, I do not think it comports well with the overall Humean picture. For example, according to the BPSA, the Humean views the standards of lawhood as truth-conducive because they are *constitutive* of lawhood. She does *not* say that these standards help produce principles that are predictively useful to us, though they are not a guide to the truths about the laws. (Indeed, this latter view begins to sound rather anti-Humean.) So it seems to me that if the Humean wants to be consistent in the way she treats both (i) our epistemic standards for lawhood, and (ii) the explanatory virtues, she should view them *both* as fundamentally prediction-conducive, and as derivatively truth-conducive. (Thanks to Marc Lange and an anonymous referee for raising this point.)

original). In lieu of explanation, these philosophers tended to emphasize the predictive aims of science. This can be observed even in the writings of Hempel and Oppenheim, who, as we have already seen, claimed that the importance of scientific explanation lay in its connection to prediction in their seminal 1948 paper “Studies in the Logic of Scientific Explanation.” Oddly enough, it was that very paper which seems to have engendered a shift in perspective, encouraging many philosophers not only to believe that there *are* such things as scientific explanations, but also to view them as an independent goal in their own right.¹⁹ For example, fifteen years after that paper, Nicholas Rescher wrote:

It is a mistake to think of the definitive task of science as being the specific prediction of the future states of natural systems. It would be more appropriate to construe the root of science in terms rather of explanation than of prediction. (1963, p. 343)

Nowadays when philosophers discuss the aims of science, they tend to mention both prediction and explanation, but as Douglas (2009, pp. 444–445) laments, the overwhelming majority of the discussion focuses on the latter.

My suggestion, then, is that the Humean should agree with Douglas in thinking that the pendulum has swung too far away from prediction. On the Humean picture I have been sketching, prediction is the fundamental aim of science: the fundamental point of engaging in the scientific enterprise is to allow us to make more accurate, more reliable predictions about the behaviors of systems in our environment. Of course, that is not to say that prediction is the *only* aim of science. It’s just that the other aims are *derivative* upon predictive aims. For example, scientists of course care about explanation—contra Pearson, nobody now believes that science doesn’t explain anything—but the ultimate scientific value of explanatory considerations, on this picture, is instrumental: they are ways of producing theories of greater predictive utility.

There is a clear benefit to conceiving of the aims of science in this way, namely, we thereby provide a coherent rationale for our society’s investment in science. As a species, we devote a great deal of time and money to the scientific enterprise. For example, the discovery of the Higgs boson is estimated to have cost over \$13 billion, and was the culmination of many years of research by thousands of people (Knapp 2012). Why, you might wonder, do we care so much about doing science?

If the fundamental aim of science is explanation, then we immediately run into another question: Why ask “Why?”? (Salmon 1979). This question, which Prescott-Couch (2017) calls the “Why Care? Question,” is not so easy to answer, at least not without appealing to more straightforward practical concerns. Prescott-Couch’s point is that even if we were to come up with necessary and sufficient conditions for something to count as a scientific explanation, we would still face the further question of why creatures like us should care about that sort of thing. Of course, one reason we might care about having scientific explanations is that they are particularly insightful or intellectually satisfying. But mere intellectual curiosity,

¹⁹ For discussion of this shift in perspective, see, e.g., Salmon (1999) and Douglas (2009).

powerful as it is to some of us, cannot account for why our society *as a whole* is willing to invest so much time, money, and energy into scientific endeavors.²⁰ For example, it is simply not plausible that the search for the Higgs boson was motivated primarily by intellectual curiosity.

By contrast, if the fundamental aim of science is prediction, then we have a straightforward answer to why we care so much about doing science. And this is because the analogous “Why Care? Question” for *prediction* is easy to answer. Why should we care about being able to make accurate predictions in the first place? Because predictive capabilities are helpful for planning actions and controlling aspects of our environments. Prediction is straightforwardly useful in a way that explanation is not. Consequently, conceiving of science as fundamentally aimed at prediction allows us to give a satisfying account of why society as a whole—not just the intellectually curious among us—is willing to invest so much in scientific endeavors.

So the Humean should conceive of science as fundamentally aimed at increasing our predictive capacities. Explanation, on this view, is a derivative aim of science, and the purpose of seeking scientific explanations is to increase the predictive utility of our scientific theories. Now let us return to the question of Sect. 2: Are scientific and metaphysical explanations linked by a transitivity principle (TP)? If so, we have seen that it spells trouble for the Humean. The difference now is that the Humean has a specific conception of the aims of scientific explanation in hand, and we have seen how this aim can inform the virtues that scientists regard as explanatory. And so the question of whether, and to what extent, scientific and metaphysical explanations *line up* (so as to possibly make TP true) is manifestly going to depend on whether, and to what extent, their *aims* line up. Thus, to evaluate the plausibility of a connection between scientific and metaphysical explanations, we are also going to have to consider the aim of explanation in *metaphysics*.

Of course, there are a variety of different conceptions of metaphysics and metaphysical explanations. Philosophers of different persuasions have suggested that metaphysical explanation aims at diverse tasks ranging from limning the fundamental structure of reality to investigating the nature of our own concepts. I am not going to propose a specific conception of the fundamental aim of metaphysics here, and I do not think the Humean needs to have such a proposal in order to rebut the circularity worries. The reason is that *everyone*, Humeans and anti-Humeans alike, should agree on the following: the fundamental aim of metaphysical explanation, whatever it may be, is *not* to increase the predictive utility of our metaphysical theories. Metaphysics simply does not aim primarily at designing theories that are maximally predictively useful. That is of course not to say that it is not helpful or beneficial when metaphysical theories make predictions that end up coming true; accurate predictions are sometimes an indication that a metaphysical theory is on the right track. But it is to say that our fundamental aim in metaphysics is not—and never has been—to use our theories for predictive

²⁰ This point should be especially salient to philosophers, of all people. Would that philosophy commanded the massive societal investment that something like physics does.

purposes. Indeed, if this *were* our fundamental aim, then we really ought to be alarmed by our results, since metaphysical theories tend not to be particularly useful for predictive purposes. Imagine trying to use the BSA, for example, to make predictions about how a system of particles is going to behave. This would be a hopeless endeavor.²¹

So according to the Humean, the aims of explanation in science and metaphysics must be quite different. Scientific explanation aims primarily at predictive utility, but this cannot be the aim of metaphysical explanation. Consequently, we should expect to find differences in the explanatory virtues appealed to in science and metaphysics. We should expect, that is, to find different virtues in each field, or to find that similar virtues are implemented differently, or that they are *balanced* differently. For example, no one should be tempted to implement the AIC conception of extrapolative simplicity in designing a metaphysical theory. Of course, this is not to say that simplicity is not an explanatory virtue in metaphysics—it clearly is. But its justification as a virtue, and the way it is used to shape our metaphysical theories, must be different than it is in science.

Similarly, while both metaphysicians and scientists tend to prefer theories with greater unificatory power, if the Humean is right then the justification for this preference in each case must be different. Scientific theories with more unificatory power are preferable because they are better able to predict phenomena falling in the intersection of various subdomains of said theories. Notice, though, that this justification of unificatory power is somewhat defeasible: if there is good reason to think that there never *are* phenomena that fall in the intersection of the theory's subdomains, then the preference for theories with greater unificatory power could in principle be overridden by other explanatory virtues. But it is not clear that this

²¹ It might seem that I am cheating by not proposing a specific conception of the fundamental aim of metaphysical explanation. After all, I am defending a metaphysical view in Humeanism, so shouldn't I have a view about the aim of metaphysical explanation that supports that view?

Whether this is true depends on whether our first-order metaphysical views require second-order metametaphysical views for their justification. If that were the case, then indeed it would be unfair of me to defend a first-order view like Humeanism without also advocating a metametaphysical view about the aims of metaphysical explanation that could justify it. But I prefer not to think of our first-order metaphysical theorizing as wholly dependent on metametaphysical foundations. For this seems to launch an infinite regress: if our first-order views need grounding in second-order views, then don't those second-order views need grounding in third-order views—shouldn't our metametaphysics be justified by a metametametaphysics—and so on ad infinitum?

Rather, I prefer to think of our first-order metaphysical views and our metametaphysical views as jointly determined by a process of reflective equilibrium. Whether we find a particular first-order view attractive will in part depend on whether it accords nicely with our metametaphysical inclinations. But likewise, whether we find a metametaphysical view attractive can also depend on whether it accords nicely with our preferred first-order theories. Thus, one does not require a fully articulated metametaphysical view about the fundamental aim of metaphysics before one can advocate a first-order metaphysical view like Humeanism. (Though this is perfectly compatible with knowing that the fundamental aim of metaphysics is *not* predictive.) Rather, we may let our second-order views about the fundamental aim of metaphysics be determined by what strikes us as good first-order metaphysics. Spoils to the victor.

could likewise happen in metaphysics, where pointing out that a theory is radically disjunctive tends to be viewed as a more damning criticism.²²

If the explanatory virtues are different in metaphysics than in science, then there is no reason to think that there are any systematic connections between the explanations of metaphysics and the explanations of science. In particular, there is no reason to think that Lange's transitivity principle holds. Of course, this does not preclude the possibility that the explanations of science and metaphysics may *sometimes* align so that we can chain them together in the manner of TP. Indeed, this may account for the plausibility of some of the examples that Lange uses to motivate TP (more on this in a minute). But on the picture I am suggesting, we should not expect to find that TP holds *in general*. And in particular, we should not expect that it holds in the case of laws explaining their instances. Yes, according to the Humean, the particular matters of fact metaphysically explain the laws. And yes, the laws scientifically explain the particular matters of fact. But it does not follow that the particular matters of fact have to scientifically explain themselves. Why? Because if scientific and metaphysical explanation have different aims, then the fact that *X* metaphysically explains *Y* should not, by itself, compel us to conclude that *X* *scientifically* explains anything that *Y* does. In short, given their different aims, scientific explanation is not beholden to metaphysical explanation in its pronouncements about what explains what.

Furthermore, this account of scientific explanation as fundamentally aimed at predictive utility actually gives us insight into what is bad about cases of self-explanation. If science admitted the possibility of self-explanation, this would *completely undermine* the ability of explanations to function as a guide to predictive utility. For if every fact could be explained just by appealing to itself, then explanatory considerations would pose no constraints whatsoever on the resultant theory. They would be utterly useless for theory building. And that flies in the face of the Humean's view of scientific explanation as an instrument used to generate theories of greater predictive utility. Therefore, the Humean has a way, not only of avoiding the explanatory circularity worry, but also of accounting for why explanatory circularity is something to be avoided in the first place.

Lastly, there remains the question of why we may have found TP plausible in the first place. Many of the examples used to motivate TP involve a less fundamental fact being grounded in, or metaphysically explained by, a collection of more fundamental facts. In Lange's example of the expanding balloon (2013, p. 257), the internal pressure of the balloon helps to scientifically explain why the balloon expands. And the internal pressure itself is grounded in the forces exerted by various gas molecules colliding with the sides of the balloon. Thus, we are supposed to conclude that those forces help to scientifically explain why the balloon expands. This seems perfectly correct, and in accord with TP. Furthermore, it seems likely

²² Case in point: physicists tend not to be alarmed by the fact that textbook formulations of quantum mechanics treat the measurement process as fundamentally different than the rest of the dynamics. Metaphysicians, on the other hand, view this disjunctivity as unacceptable in a fundamental theory of the world, and have long searched for resolutions to the so-called "measurement problem." Perhaps the difference here is explained by the different explanatory aims of physicists and metaphysicians.

that we could produce chains of explanation like this in many other cases where we have a scientific reduction of one domain to another, such as in the statistical mechanical explanation of thermodynamic phenomena. What, then, is the Humean to say about the existence TP-like chains of explanations in cases of scientific reductions?

I think the best response for the Humean is to maintain that these explanations of physically less-fundamental phenomena by more fundamental phenomena are, in fact, not metaphysical explanations at all, but scientific explanations.²³ Thus, in these examples we have chains of *scientific* explanations, but no direct connection between metaphysical and scientific explanations. In support of this position, the Humean could point out that scientific reductions, such as the reduction of thermodynamics to statistical mechanics, all have their origins in scientific practice, not in metaphysics. And therefore it would seem quite odd to claim, for example, that statistical mechanical explanations of thermodynamic phenomena are *not* scientific explanations. If we have to classify it as one or the other, it would seem much more natural to classify it as a scientific explanation than a metaphysical one.

The strongest argument in favor of viewing these sorts of physically reductive explanations as metaphysical explanations seems to be that they appeal to more fundamental facts (i.e. the lower-level realizers) to explain less fundamental facts, and since metaphysics is commonly regarded as articulating fundamentality hierarchies, they must therefore be metaphysical explanations. But there are two points to note here. First, the conception of metaphysics as articulating fundamentality hierarchies is not necessary—I've already mentioned other possibilities for the aim of metaphysics, such as describing the structure of our concepts. And second, even if the Humean adopts this fundamentality conception of metaphysics, this still doesn't establish that the statistical mechanical explanations of thermodynamic phenomena are themselves metaphysical explanations. For the Humean *already* has a need for a distinction between physical and metaphysical fundamentality.²⁴ To see this, just note that the Humean thinks that the fundamental laws of physics are not themselves *metaphysically* fundamental. So there is already a need for the Humean to appeal to an independent notion of metaphysical fundamentality that is distinct from physical fundamentality. And that distinction undercuts the inference from the relative fundamentality of statistical mechanical phenomena to the conclusion that statistical mechanical explanations of thermodynamic phenomena must be metaphysical explanations. Rather, statistical mechanics could be physically more fundamental without being metaphysically more fundamental.

Of course, this response would require further exploration in a number of respects, especially since it presupposes a distinction between physical and metaphysical fundamentality that needs further articulation. Unfortunately, this

²³ Thanks to an anonymous referee for suggesting this response.

²⁴ Bhogal (forthcoming) likewise argues that the Humean should distinguish between scientific and metaphysical fundamentality. He also suggests a similar response to the explanatory circularity problem to the one proposed here, namely that the Humean should distinguish between the aims of metaphysics and the aims of science. His proposal is interesting, and differs from mine in a number of respects. I leave a comparison of our proposals for future work.

would take us well outside the scope of this paper. Nevertheless, I think that the foregoing arguments show that the Humean has every reason to doubt the existence of a systematic connection between scientific and metaphysical explanations.

6 Conclusion

In her (2012) paper on the methodology of metaphysics, Paul defends the legitimacy of metaphysics as a discipline by arguing that it uses the same tools as science. Specifically, both scientists and metaphysicians use inference to the best explanation to guide their theorizing. Paul argues that since IBE tracks the truth in science, we have every reason to think that it also tracks the truth in metaphysics.

The Humean should be wary of this argument, because it assumes that science and metaphysics have, at bottom, the same goal: discover the truth. On the Humean picture I am suggesting here, science aims fundamentally at prediction, and the truth of our scientific theories is a *function* of their predictive utility. But metaphysics does not aim, fundamentally, at prediction, and the truth of our metaphysical theories is *not* primarily a function of their predictive utility. If the Humean is right, the different aims of science and metaphysics should lead us to be cautious about importing the tools of one discipline for use in the other. Tools like IBE, fine-tuned as they are for scientific purposes, are useful for particular goals, and it is inadvisable to assume that since they are useful for one goal, they also must be useful for another.²⁵

Indeed, there is a concern here that applies, not just to Humeans, but to all metaphysicians. Insofar as the aims of science and metaphysics *do* coincide, we metaphysicians should feel a little uncomfortable. After all, we all ought to be wary of one of the recurring questions of undergraduates: “Shouldn’t you really just be doing science?” If we want to have a compelling response to that question, then it helps to be able to suggest that metaphysicians are aiming at something very different than scientists are. And insofar as our aims are different, one would also expect our tools to be different. The Humean, I am suggesting, has an understanding of science that readily lends itself to this sort of difference in both aims and tools. What is less clear at this point is whether the anti-Humean does also.

References

- Albert, D. (2015). *After physics*. Cambridge, MA: Harvard University Press.
- Armstrong, D. (1983). *What is a law of nature?*. Cambridge: Cambridge University Press.
- Beebe, H. (2000). The non-governing conception of laws of nature. *Philosophy and Phenomenological Research*, 61, 571–94.
- Bhogal, H., & Perry, Z. (2017). What the humean should say about entanglement. *Noûs*, 51, 74–94.

²⁵ Of course, this does not preclude the possibility of “tuning” IBE differently, so that we *can* use it for metaphysics. My point is just that we ought not to use it indiscriminately, without careful consideration of whether it has been adjusted for our purposes.

- Bhagal, H. (Forthcoming). Nomothetic explanation and Humeanism about laws of nature. *Oxford Studies in Metaphysics*.
- Cabrera, F. (2017). Can there be a Bayesian explanationism? On the prospects of a productive partnership. *Synthese*, 194, 1245–72.
- Cohen, J., & Callender, C. (2009). A better best system account of lawhood. *Philosophical Studies*, 145, 1–34.
- Dellsén, F. (2016). Explanatory rivals and the ultimate argument. *Theoria*, 82, 217–37.
- Dellsén, F. (2018). The heuristic conception of inference to the best explanation. *Philosophical Studies*, 175, 1745–1766.
- Dorst, C. (2018). Toward a best predictive system account of laws of nature. *British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axy016>.
- Douglas, H. (2009). Reintroducing prediction to explanation. *Philosophy of Science*, 76, 444–63.
- Emery, N. (forthcoming) Laws and their instances. *Philosophical Studies*.
- Forster, M., & Sober, E. (1994). How to tell when simpler, more unified, or less ad hoc theories will provide more accurate predictions. *British Journal for the Philosophy of Science*, 45, 1–35.
- Fumerton, R. (1980). Induction and reasoning to the best explanation. *Philosophy of Science*, 47, 589–600.
- Hall, N. (ms) Humean reductionism about laws of nature. <https://philpapers.org/archive/HALHRA.pdf>
- Harman, G. (1965). Inference to the best explanation. *Philosophical Review*, 74, 88–95.
- Hempel, C., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15, 135–75.
- Henderson, L. (2014). Bayesianism and inference to the best explanation. *British Journal for the Philosophy of Science*, 65, 687–715.
- Hicks, M. (2017). Dynamic humeanism. *British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axx006>.
- Hicks, M., & van Elswyk, P. (2015). Humean laws and circular explanation. *Philosophical Studies*, 172, 433–43.
- Jaag, S., & Christian, L. (forthcoming). Making best systems best for us. *Synthese*.
- Kieszeppää, I. (1997). Akaike information criterion, curve-fitting, and the philosophical problem of simplicity. *British Journal for the Philosophy of Science*, 48, 21–48.
- Knapp, A. (2012). How much does it cost to find a higgs boson? *Forbes*. <https://www.forbes.com/sites/alexknapp/2012/07/05/how-much-does-it-cost-to-find-a-higgs-boson/#4f5d83453948>
- Kuhn, T. (1977). *The Essential tension*. Chicago: University of Chicago Press.
- Lange, M. (2009). *Laws and Lawmakers*. New York: Oxford University Press.
- Lange, M. (2013). Grounding, scientific explanation, and humean laws. *Philosophical Studies*, 165, 255–61.
- Lange, M. (2018). Transitivity, self-explanation, and the explanatory circularity argument against humean accounts of natural law. *Synthese*, 195, 1337–53.
- Lewis, D. (1973). *Counterfactuals*. Malden: Blackwell.
- Lewis, D. (1986). *Philosophical papers* (Vol. II). New York: Oxford University Press.
- Lewis, D. (1994). Humean supervenience debugged. *Mind*, 103, 474–90.
- Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). New York: Routledge.
- Loewer, B. (2012). Two accounts of laws and time. *Philosophical Studies*, 160, 115–37.
- Marshall, D. (2015). Humean laws and explanation. *Philosophical Studies*, 172, 3145–65.
- Maudlin, T. (2007). *The metaphysics within physics*. Oxford: Oxford University Press.
- Miller, E. (2014). Quantum entanglement, bohmian mechanics, and humean supervenience. *Australasian Journal of Philosophy*, 92, 567–83.
- Miller, E. (2015). Humean scientific explanation. *Philosophical Studies*, 172, 1311–1332.
- Paul, L. A. (2012). Metaphysics as modeling: The handmaiden’s tale. *Philosophical Studies*, 160, 1–29.
- Pearson, K. ([1911] 1957). *The grammar of science* (3rd ed.). New York: Meridian Books.
- Prescott-Couch, A. (2017). Explanation and manipulation. *Noûs*, 51, 484–520.
- Psillos, S. (2002). Simply the best: A case for abduction. In A. C. Kakas & F. Sadri (Eds.), *Computational logic: Logic programming and beyond* (pp. 605–622). Berlin: Springer.
- Quine, W. V. O., & Ullian, J. (1970). *The web of belief*. New York: Random House.
- Rescher, N. (1963). Discrete state systems, markov chains, and problems in the theory of scientific explanation and prediction. *Philosophy of Science*, 30, 325–45.
- Salmon, W. (1979). Why ask, ‘Why?’? An inquiry concerning scientific explanation. In: Salmon, W. (Ed.) *Hans Reichenbach: Logical empiricist*. *Synthese Library*. Dordrecht: Springer.

- Salmon, W. (1999). The spirit of logical empiricism: Carl G. Hempel's role in twentieth-century philosophy of science. *Philosophy of Science*, 66, 333–50.
- Shumener, E. (2017). Laws of nature, explanation, and semantic circularity. *British Journal for the Philosophy of Science*. <https://doi.org/10.1093/bjps/axx020>.
- Strevens, M. (2008). *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.
- Thagard, P. (1978). The best explanation: Criteria for theory choice. *The Journal of Philosophy*, 75, 76–92.
- van Fraassen, B. (1980). *The scientific image*. New York: Oxford University Press.
- van Fraassen, B. (1989). *Laws and symmetry*. New York: Oxford University Press.
- Woodward, J. (2014). Simplicity in the best systems account of laws of nature. *British Journal for the Philosophy of Science*, 65, 91–123.