

Methods in the Metaphysics of Science

Kerry McKenzie

1 Introduction

What is the method of the metaphysics of science? One could be forgiven for thinking that there is an easy answer here. For presumably the metaphysics of science, as practiced today, is just a special case contemporary metaphysics more generally; and in that larger context, there is a common refrain regarding how metaphysics gets done. It is that the method of metaphysics consists of (1) the creation of a range of theories that may be thought to describe the metaphysical features of the world, or some domain of it at least; (2) extolling the *virtues* of those theories with respect to simplicity, unification, explanatory power, etc; and finally (3) the eventual acceptance of that theory which scores best with respect to these virtues. And this method is justified by metaphysicians as an appropriate and reliably truth-tracking method via the claim that it is essentially a *scientific* method, perhaps even *the* scientific method ‘once the data is in’. One might then think that – metaphysics of science being a special case of metaphysics – we should say that that is its method too, and that it is a recommendable one for the very same reason. It would certainly be natural for the method of the metaphysics of science to be intimately connected to that in science, given that very words involved suggests there may be an intimate connection between the two.

The problem with this tempting answer, however, is that philosophers of science have in recent years been working overtime to undermine the supposed analogy between the methods of metaphysics and science that underwrites its vindication. This work has not denied that appeals to virtues are part and parcel of theorizing in metaphysics, but rather claimed that the idea that this is essentially just what scientists do is based on a naive understanding of science. Some even go as far as to claim that by casting itself as just like science while presenting a distorted version of the latter, metaphysics has in effect rendered itself a pseudoscience (Ladyman and Ross 2007, Sec. 1.2.3). But if there is one thing that metaphysicians of science pride themselves on, it is the fact their work is based in science, and moreover *real* science, not ‘folk’ or hopelessly outdated science. Relying on false caricatures of the scientific method to make sense of its own is thus no less *verboten*, in this context, than relying on false caricatures of scientific theories to ground metaphysical claims.

Given that this is our predicament, the ‘easy’ answer to the question of what the method of the metaphysics of science consists of is closed off. To help us figure out what we should say instead, I will begin by unpacking what philosophers of science take to be problematic about

the received view on the method of metaphysics more generally. What those objections in sight, I will then ask what can and should be said about the the methods of the metaphysics of science more specifically. I will not end by proposing a full and definitive answer. But I will be arguing that (1) *falsification* is a core part of the method of the metaphysics of science, and that (2) the idea that appeals to ‘virtues’ must be made rests on widely accepted but arguably under-theorized claims regarding *metaphysical underdetermination*. Note also that I will not claim that even this under-characterized statement of the method of scientific metaphysics will fit every project within it, and indeed my focus will mostly be on the metaphysics of physics in particular. But since many metaphysical projects outside of that context have the same general structure, the morals will largely generalize.

2 The Current Orthodoxy

Let’s begin with the problematic current orthodoxy regarding the method of metaphysics in general. The claim here is that the method of metaphysics consists of (i) the creation of metaphysical theories of some domain, followed by (ii) the comparison of those theories with respect to their respective *theoretical virtues* such as simplicity, elegance, and explanatory power, culminating in (iii) the selection of the theory that optimizes these virtues as the best candidate for the true theory. A further but crucially significant claim is that this method is *essentially just that of science*. For example, Ted Sider, John Hawthorne, and Dean Zimmerman write:

Scientists must regularly choose between many theories that are consistent with the observed data. Their choices are governed by criteria like simplicity, comprehensiveness, and elegance. This is especially true in very theoretical parts of science, for instance theoretical physics [. . .] Just like scientists, metaphysicians begin with observations, albeit quite mundane ones: there are objects, these objects have properties, they last over time, and so on. And just like scientists, metaphysicians go on to construct general theories based on these observations, even though the observations do not logically settle which theory is correct. In doing so, metaphysicians use standards for choosing theories that are like the standards used by scientists (simplicity, comprehensiveness, elegance, and so on) (Sider, Hawthorne and Zimmerman 2008, p. 6)

L.A. Paul puts the point more succinctly:

We use theoretical desiderata as guides to truth in metaphysics just as we use such desiderata as guides to truth in science, since the method is fundamentally the same even when the subject matter is different. In both cases, an acceptable theory must maximize theoretical virtues while being empirically adequate. (Paul 2012a, p. 21)

The basic picture here is thus one of deep analogy, or *methodological continuity*, between science and metaphysics: metaphysics is here painted as sort of ‘limiting case’ of theorizing, one in which empirical input tends towards zero (cf (Chakravartty 2017)). It is a view that often connotes the view of W.V.O. Quine, and is often taken to have been underwritten by him. But it is also clear that, if the picture were accurate, then it would be a real boon for metaphysicians. Science after all is taken to set the bar for respectable theoretical enquiry, and presumably part of the reason we respect it so much stems from an appreciation of its method. If metaphysics can be understood as employing essentially that same method, it seems it would be downright hypocritical not to revere metaphysics in much the same way we do science – clearly this is a non-trivial finding, given how metaphysics (unlike science) has been repeatedly singled out for scepticism.

There is therefore undoubtedly a normative dimension to this picture of ‘methodological continuity’, insofar as it can be used to furnish a principled legitimation of metaphysics. While at its heart it is a simple argument by analogy, it will help to display what I will call the ‘continuity argument’ for metaphysics explicitly as follows.

- P1. Scientific theories are underdetermined by empirical evidence, and scientists must therefore appeal to something other than empirical evidence to choose between them.
- P2. Scientists appeal to theoretical virtues to choose between theories.
- P3. Methods used in science are reliable guides to the truth.
- P4. Metaphysicians appeal to theoretical virtues to choose between their theories.
- C₁. (From 2 and 4) The methods used in metaphysical theorizing are also used, at least in part, by scientists.
- C₂. (From 3 and C₁) The methods used in metaphysical theorizing are reliable guides to the truth.

This picture of deep methodological continuity is undoubtedly a neat and elegant representation of the relationship between metaphysics and science. By its own lights, then, perhaps we ought to regard it as true! Unfortunately however almost every aspect of the continuity view has been attacked by philosophers of science, as has the move from assumptions of continuity to the licensing of metaphysics.¹ In particular, philosophers of science have questioned that either *P1* or *P2* accurately describe real science, and whether the conclusions *C₁* and *C₂* would even follow were those premises to be true. In the next section I will outline some these criticisms.² Suitably informed, I will consider what if anything can be resurrected in their place concerning the metaphysics of science specifically.

3 Problems with the picture

3.1 Contesting P1: Scientific theories are underdetermined by empirical evidence.

The first premise of the argument asserts the doctrine of underdetermination: the idea that no body of empirical evidence will ever suffice to determine a scientific theory.³ If this is true, then we can conclude that no body of empirical evidence will ever determine for us which scientific theory to adopt. As such, there is no time at which an appeal to theoretical virtues isn't necessary if we are to settle on a view of the world; such appeals must therefore be part and parcel of science.

Why should we believe this? While the now familiar idea of systematic and unavoidable underdetermination 'entered the modern literature in the conclusion of Quine's *Two Dogmas of Empiricism*', philosophers of science have noted that Quine himself barely presented any actual argument for this claim, as opposed to a few 'suggestive but elusive metaphors' ((Norton 2003, p. 22)).⁴ Moreover, the underdetermination that Quine himself there envisages pertains to 'total' systems, subsuming mathematical entities, logical laws, and Homeric gods in addition to subatomic particles and other more straightforwardly natural or physical entities. Thus whatever the relationship from evidence to theory involved here is, it is not at all clear that it is the inductive relation of confirmation (ibid.).

If we are to accept the crucial premise of underdetermination of theory by evidence as the scientific predicament clearly a better source of justification is needed. But John Norton has argued that there is actually nothing in our best theories of confirmation that supports it as a general principle. Rather, familiar arguments for the thesis 'seem to depend entirely on a flawed account of the nature of induction' (Norton 2003, p. 27): in particular, they rest of the idea that confirmation is simply a logical relation between theories and their observable consequences. In this picture – basically, the hypothetical-deductive model of confirmation – if two theories logically imply the same observations then they can only be equally confirmed by them (for whether something is an implication or not is not a matter of degree). But this simple theory of induction is already known to be wildly over-permissive, and in general insufficiently sophisticated to capture real scientific inference. Whatever arguments can be given for underdetermination via this framework are therefore of dubious relevance.

This idea that there are no convincing *a priori* arguments for systematic underdetermination has also been echoed by P. Kyle Stanford. He claims, for example, that those arguments simply 'exchange underdetermination for familiar philosophical chestnuts', such as the spectre that we are being manipulated by malicious deceivers or that we are mere pixels in a vast computer simulation (Stanford 2001, S1). As such – and analogously to the criticisms of Quine above – these arguments are simply irrelevant to anyone that is concerned with understanding the inductive relationships relationships that exist between scientific theories and empirical data, but that also takes no interest in defeating such scenarios absent any reason to accord them 'prior plausibility' (ibid., p. S3). Surely this describes many within the community of philosophers of science. However, Stanford believes that *historical* support may nevertheless

be found for widespread underdetermination in science. According to his ‘new induction’, the largely uncontested facts of theory change in science demonstrate that there have always been scientific rivals for previously-held theories, even if historical actors were not in any position to conceive of them at the time. This would seem to give us inductive grounds to believe that real scientific theories will always be ‘transiently’ underdetermined by another scientific theory, even if we can’t say without hindsight what that theory is.

Stanford’s argument has been criticized on multiple grounds (see Saatsi et al. 2009). Most relevantly for present purposes, arguably even the pairs of theories that feature in history of science as examples of transiently underdetermined rivals can be challenged as such, once we conceive of confirmation in the less naive terms urged by Norton. To see the point, consider that a classic example of underdetermination in the literature is the Copernican and Ptolemaic systems prior to Galileo’s observations of the phases of Venus in 1610. This prompted some (such as Kuhn) to assert that the Copernican theory can only have been chosen on account of its non-empirical virtues (see e.g. Kuhn 1957, p. 172). However, applying e.g. a Bayesian analysis to these theories in fact reveals the Copernican view to be better confirmed at this time than its rival (see e.g. Myrvold 2003, Sec. 4). To see why, consider the fact that Mars is seen as undergoing retrograde motion only when aligned with the Sun and the Earth – that is, when it is in ‘opposition’. This is *entailed* by the core of the Copernican model but merely *permitted* by the Ptolemaic, and follows from the latter only when fitted with special parameters. This renders the former model *more falsifiable* with respect to the motion of Mars, and thus more strongly confirmed by its observed motion.⁵ Pablo Acuña has given a similar analysis of another stock case of underdetermination in science – the case of Einstein versus Lorentz (Acuña 2014).⁶ By utilizing a more sophisticated theory of confirmation than mere hypothetico-deductivism, and applying it to the ample well-established empirical knowledge that was in the background at the time of the dispute over the status of the ether, Acuña argues that Einstein’s theory was better supported empirically than was its rival.

Of course, none of this entails that there will never be real-life cases of underdetermination in science. But that even these most classic of supposed examples of underdetermination can turn out to be spurious when a more sophisticated approach to confirmation is employed surely undermines the plausibility of the idea that underdetermination is the norm in the history of science. At the very least, then, Premise 1 should be regarded as deeply controversial.

3.2 Premise 2: Scientists appeal to theoretical virtues to choose between theories.

The second premise of the continuity argument is that scientists routinely use virtues to select among theories.

On the face of it, this seems totally uncontroversial. The very first of Newton’s rules of reasoning was after all an explicit principle of parsimony. Since that time, some ground-breaking scientific theories have been justified in part by appeal to parsimony-based reasoning. Sober for example notes that ‘many evolutionists have thought that parsimony is a guide to

reasoning about phylogenetic relationships' (Sober 2009, p. 121). Further, were scientists not routinely appealing to virtues to justify their theories it would be difficult to make sense of van Fraassen's efforts to undermine their evidential significance in his classic defence of scientific anti-realism (Van Fraassen 1980, Chapter 4). And a new, interesting development in this vicinity is recent work in experimental philosophy, which suggests that scientists do view virtues such as simplicity and unification as features that would raise their credence in a theory (Schindler 2022).

But for all that, it remains clear – to some, so obvious as to be 'a little embarrassing' to point out – that the role of virtues in science simply *cannot* be as simple-minded as it presented in the literature on the method of metaphysics (Ritchie 2022, p. 205). We can note first of all that the 19th century was rife with debates about the reality of invisible entities even though all parties (by and large) were in agreement about the explanatory fertility of the atomic hypothesis (Psillos 2011). It is thus impossible to make sense of this period – the so-called 'golden age of science' – were the picture painted sufficient as an account of theory selection in science. Even the case of assent to Darwin's theory, often held up as a clear case of the virtue account in action (see e.g. Paul 2012a, p. 12), is not aptly characterized by the simple idea that providing more 'unified' explanations of phenomena constitute sufficient grounds for belief. While it may still be debated how best to reconstruct what Darwin was trying to do, the first part of the *Origin* consists of a detailed description of the processes of artificial selection that he clearly took to do crucial work in establishing the reality of the mechanism of natural selection. Only later, *once* the existence of the mechanism of natural selection has been rendered plausible by other means, do the sort of explanatory arguments appealed to by metaphysicians enter in (see Novick 2017).

A metaphysician could of course protest at this point that while there may be *more* to theory selection in science than appeal to virtues, it nevertheless is a core component of inference in science (such as for example in that later portions of Darwin's long argument). As such, metaphysicians can still claim for themselves whatever support issues from that limited component of the method. Such a move requires that (a) the same virtues are at play in both cases, as expressed in C_1 , and that (b) the justification for their use in one case can be extrapolated to that in the other, as expressed in C_2 . But both of these can be contested too.

3.3 Contesting C_1 : Science and metaphysics in part use the same method.

Even if it is the case that scientists do appeal to virtues in some course of their theorizing, it can be argued that the virtues used in science are not of the same nature as those used in metaphysics. This is because the virtues that are trusted in science are often entwined with empirical considerations in a way that they are usually not in metaphysics. Consider again the Ptolemaic and Copernican models and how they deal with the fact that the retrograde motion of Mars takes place only 'at opposition'. The Ptolemaic theory deals with retrograde motion through the method of epicycles, and can account for this observed motion by tuning

the velocity of Mars around its epicycle to the period of the Sun around the Earth. But there is nothing whatsoever in the basic framework of the theory that says these things should be connected. For this reason, the Ptolemaic theory can be said to provide a less ‘unified’ account of the phenomena than the Copernican theory, in which the appearance of retrograde motion is inextricably bound up with facts about orbital period (see e.g. Henderson 2014), (Myrvold 2003). But this is the same reason that the latter comes out to be better confirmed than the former. Thus, we may say that ‘on [the] Bayesian account... the power of a theory to unify a body of evidence is not an extra-empirical virtue but contributes directly to the degree to which the evidence supports the theory’ (Myrvold 2003, p. 413). If this is right, then in some cases at least scientists are not appealing to ‘extra-empirical virtues’ even when they do appeal to facts about ‘unity’, and if so they are presumably appealing to something different from those under discussion in metaphysics. Along similar lines, Elliot Sober argues that while in evolutionary biology the ‘principle of parsimony’ is undoubtedly used in hypothesis selection, it is simply a surrogate for the injunction to choosing theories with a higher likelihood – that is, that render more likely our observations. When so used, it isn’t an ‘optional aesthetic frill’ but has ‘demonstrable epistemic relevance’ (Sober 2009, p. 118).

There is much more to say here, but the point for present purposes is that if it is right that ‘theoretical virtues’ as actually used in science are often not ‘extra-empirical’, then metaphysicians and scientists are seemingly not appealing to the same things when they recruit them in theory selection. As such, the inference to C_1 is blocked, on grounds of equivocation.⁷

3.4 Contesting C_2 : The methods used in metaphysical theorizing are epistemically legitimate

Suppose however for the sake of argument that scientists and metaphysicians both appeal to virtues, and that when they do so they are appealing to the same, ‘extra-empirical’ things. Assuming that, would it not be hypocritical to deny that metaphysicians can claim whatever justification scientists claim from this step in their arguments? Some metaphysicians are sure this is the case:

But while it is true that the empirical, conformable features of scientific theories have allowed us to confirm the value of theoretical desiderata for theorizing, if such features are truth conducive in the case of science, they should be truth conducive more generally... This is a central part of my thesis: if we accept inference to the best explanation in ordinary reasoning and in scientific theorizing, we should accept it metaphysical theorizing. (Paul 2012a, pp. 21–22)

It has however been a theme in recent philosophy of science that there is in general no reason to think that if we trust an inference in one context then we are actually entitled to trust it in another. This is not simply van Fraassen’s point that one is not *compelled* to trust

an inference in some domain simply because we trust it in another (see e.g. Van Fraassen 1980, Chapter 2). It is a growing appreciation, largely based on the work of John Norton on the ‘material theory’ of induction, that the reliability of any inductive inference is highly contextual, and that its reliability is controlled by context-specific background assumptions (Norton 2021). To see the basic point here, compare the two inferences :

Some samples of bismuth melt at 271°C. Some samples of wax melt at 91°C.
∴ All samples of bismuth melt at 271°C. ∴ All samples of wax melt at 91°C.

The inferences are formally identical, and yet most of us will trust the conclusion of the first and not the second. We do so because we have detailed background knowledge of the substances involved concerning the intrinsic similarity of their samples from a chemical point of view, their status as natural kinds and so on. But what this simple example shows is that two formally identical inductive inferences can differ wildly in terms of their reliability. This is relevant, given that virtue-based inferences in science are largely presumed to have a non-demonstrative character: that is, they are supported on account of assumptions concerning how they have actually performed in science, rather than taken to be true ‘by the natural lights of reason’ or some purely deductive process.⁸ It follows that even if there is a formal similarity between the arguments used in science and in metaphysics, it can be claimed that this is a superficial feature that in itself tells us little about how trustworthy the inferences are in each case (see also (Saatsi 2017)). Indeed, Norton argues that our judgments of the plausibility of parsimony principles in science is implicitly conditioned on background knowledge in a similar way to our judgements concerning the enumerative inductions above (Norton 2021, Chapter 6). Thus even if we can find such principles being appealed to in science, and even if we can claim that they are not ‘extra-empirical’ in nature, there is nothing to say that their reliability can be exported to contexts of metaphysics.

3.5 Summing up.

It is undeniable that much of the business of metaphysics consists in evaluating and comparing theories with regard to extra-empirical virtues. Leading contemporary metaphysicians liken this method with that used of science, with the implication being that metaphysics should have comparable epistemic standing. But philosophers of science have argued on multiple grounds that this continuity argument can be resisted: it isn’t clearly the case that scientists use the methods that metaphysicians use, and even if they did it isn’t at all clear that metaphysics would thereby be entitled to claim comparable epistemic legitimacy.

So much for metaphysics at large then. But where does all this leave the metaphysician of science? What methods does she use, and what can we say for their legitimacy?

4 The Method of Metaphysics of Science

First on the docket is an outline of what I will take the metaphysician of science to be. She is first and foremost a *metaphysician* – that is, someone whose line of business is constructing and defending a metaphysical view of some sort. She is a metaphysician *of science* insofar as she does not spin her view out of air but rather develops it in close consultation with the best science of her time. While not an exhaustive description of what it is that she does, a big part of her job consists in articulating and analyzing the ‘implicit subject matters of scientific theories and models’ (cf. Chakravartty 2017, p. 204). Anything that features in scientific discourse and in the process of scientific theorizing, without itself being directly scrutinized in the course of that theorizing, is a good candidate for this subject matter. Looming large in her work are thus questions concerning laws of nature and causation; probability and possibility; individuals and kinds; fundamentality and emergence, and properties and powers. The questions themselves will focus largely on the analysis of all these things, thus their fundamentality, and their objective reality. Fittingly for a discipline described as ‘the most general attempt to make sense of things’ (Moore 2011), these are highly *general* questions, directed as they are at whole categories and not (or not usually) any particular member of them, and at features that we often take to somehow pervade reality as a whole. In conducting her investigations and producing her analyses, she will often make free use of concepts and arguments that have been developed by her more armchair counterparts, in accordance with the ‘toolbox’ view of French and McKenzie (French and McKenzie 2012), (French and McKenzie 2016).⁹ But whatever tools she uses and wherever she gets them from, her primary aim is getting the metaphysics of some domain right, on the assumption that that domain is approximately described by a given scientific theory or theories. Since, in her view, her only real access to that domain is via those theories, she cannot so much as begin her investigations without them.

I think it is fair to say that most of what we are apt to call ‘the metaphysics of science’ falls into this genre. Still, in characterizing the metaphysics of science in this way, I make no pretense to be exhaustive. For one thing, under this method metaphysical claims are produced via sustained reflection on theories. As such, this method is not that of the metaphysics of science in practice [**connection here to contribution by Marie Kaiser and Javier Suarez**]. Furthermore, this method takes the relevant theory as authoritative: rather than correcting such a theory its primary aim is to interpret it metaphysically. In this it is distinct from more explicitly normative projects that aim to correct some prevailing assumptions within the certain sciences, as exemplified, for example, by Brian Epstein in *The Ant Trap* (Epstein 2015). However, metaphysics of science as I have characterized it here will most closely approximate mainstream metaphysics, given that little if any of the latter is practice-oriented; as such, it is this brand of metaphysics of science whose method may most meaningfully be compared with the canonical view. And while the examples that I employ below have all been drawn from physics, it should be noted that comparable projects go on in the metaphysics of chemistry and biology.¹⁰ My conclusions may therefore be expected to generalize somewhat.

The metaphysician of science, as I understand her at least, will thus arm herself with a scientific theory and aim to draw metaphysical conclusions on its basis. What methods does she employ in so doing? How does she argue that, say, according to our best physics identity is not primitive and that relations are fundamental (as the ontic structural realist thinks)? Or that reality is fundamentally non-separable (as the holist thinks)? Or that some kind of primitive governing based conception is the correct interpretation of physical laws (as many anti-Humeans believe)? Does the canonical characterization of the method of metaphysics in general, in which theories are appraised on the basis of their possession of virtues, fit well with the work of the metaphysician of science?

The first thing to say is that examples of this methodology in action are plentiful in the metaphysics of science. For example, Mauro Dorato and Matteo Morganti argue against structuralist views on primitive identity in physics, ‘if only on mere grounds of simplicity and minimisation of metaphysical revision’ (Dorato and Morganti 2013, p. 597). Alyssa Ney argues for configuration space realism in quantum mechanics – a view according to which spacetime is not fundamental and only a mesmerizingly high-dimensional space is – on the grounds that a metaphysics with these properties is simpler (p. 128) and in a sense more congenial to intuitions (p. 129). And many have voiced the idea that ‘the best argument’ for Humean metaphysics is simply ‘that it is metaphysically parsimonious and faces no knock-down objection’ (Huggett 2006, fn. 3). Nor should this be surprising. It is not surprising firstly because there is not necessarily a bright line demarcating the ‘metaphysics of science’ from all the rest (Chakravartty 2017). Furthermore, the idea that metaphysical theories in general are underdetermined by the data is at least as familiar (to metaphysicians at least) as the idea that scientific theories are, and might well be taken to hold independently of the thesis of scientific underdetermination. If this general underdetermination holds true, then it seems we have no choice but to appeal to virtues if we want to commit, as metaphysicians generally do, to a particular metaphysical view. So far, then, the picture of the methods of the metaphysics of science closely resembles the standard view.

Insofar as this *is* the method of the metaphysics of science, there are a number of open issues. First, is the thesis of underdetermination any better justified in the metaphysics of science than it is in science itself, and thus some appeal to virtues always necessary? Second, is the resulting appeal to virtues any better justified in the case of metaphysics of science than it was in metaphysics more generally? Let us take the first question first. It is true that metaphysics is traditionally regarded as unfalsifiable by definition: this was Popper’s take on the issue, and one can still find the view in currency today, such as when Uriah Kriegel demotes any claim that can be settled by science to the ‘only *superficially* metaphysical’ (Kriegel 2013, 6, italics in original). If this is the case, then it would seem to follow that metaphysical propositions, being wholly unresponsive to data, must be underdetermined by it. However, the metaphysician of science has likely chosen her approach to metaphysics in part because she is of the conviction that science has *overturned* some longstanding and philosophically consequential metaphysical views (such as regarding locality, essentialism, or determinism). This does not seem compatible with the idea that metaphysical theories are unfalsifiable; nor does it seem right to regard such doctrines ‘superficial’ in any sense. Of

course, one might still hold despite this that the metaphysics *of* a given scientific theory is necessarily underdetermined by the data that supports that theory even if not by all data whatsoever. Steven French, for example, regularly speaks of the ‘underdetermination of metaphysics by physics’ (French 2014, Chapter 2), and it is this situation that he seems to have in mind. But it is unclear what the *a priori* argument for this is supposed to be, any more than was the case for scientific theories. It is true that one could simply make up on the spot an interpretation of a given scientific theory and claim it as a piece of metaphysics. Indeed, the ‘familiar philosophical chestnuts’ that we saw Stanford rejecting as worth taking seriously in the debate over underdetermination in science – e.g. the hypothesis that we are all pixels in a simulation, mentioned above in Section 3.1 – seem closer to theories of metaphysics than to those of science. But it is not clear to me why the metaphysician couldn’t protest that such cooked-up theories, proposed only to make a sceptical point about underdetermination, lack the ‘prior plausibility’ that makes them worth taking seriously as real contenders for theories of metaphysics just as did Stanford in the case of scientific theories. Certainly, it now seems hasty to take as axiomatic that metaphysical underdetermination is fundamentally unavoidable, given that thanks to Quine we have been doing similarly in the case of scientific theories and now regard doing so as unwarranted.

The second open question is whether these virtue-trading methods of metaphysics are in any better shape in the context of metaphysics of science, given that the latter is somehow ‘closer’ to science than is metaphysics at large and assuming that science does, in some form or other, make appeal to virtues. Never say never, but it is hard to see how they would be. For one thing, some virtues – such as coherence with intuition – seem *terrible* guides in the metaphysics of modern physics, a discipline seemingly describing regimes where intuition goes to die ((Maudlin 2022); (Vickers 2022, Chapter 6)). For another, metaphysicians’ concern with more abstract and general questions suggests that their inferences must be governed by different principles than those that operate in science. By Norton’s lights, this will block any easy inference to the trustworthiness of metaphysics from the trustworthiness of science. Recall that, according to him, the status of non-demonstrative inferences are controlled by (often implicit) background assumptions. But what does it even mean to talk of the ‘background assumptions’ of metaphysics, when it itself is supposed to be ‘the most general attempt to make sense of things’?¹¹

Insofar as the method of the metaphysics of science relies essentially on appeals to virtues, then, there seems little more to recommend it as a truth-tracking enterprise than regular metaphysics. One could of course use this to motivate an explicitly anti-realist approach to metaphysics. Perhaps, that is, one could argue on this basis that metaphysical theory-building is essentially imaginative and aimed not at truth so much as understanding (McSweeney 2023), or that metaphysics has a purely heuristic value of some other sort (though see (Stanford 2017) for scepticism here). But doing so would be premature – the reason being that much metaphysics of science does not consist in virtue-trading at all. Indeed, metaphysics of science is arguably at its best when it explicitly and self-consciously does not rely on appeals to virtues to argue for its correctness. To me, the most exciting argumentation that goes on in the metaphysics of science is aimed at showing that certain metaphysical views

are straightforwardly *inconsistent* with extant science – that is, by the very theories that the metaphysician of science assumes as at least approximately true. While sometimes listed as such, consistency both ‘internally and with the facts’ are not mere ‘virtues’ on the same footing as simplicity and explanatory power, but rather are ‘rock-bottom criteria of minimal acceptability’ (Van Fraassen 1980, p. 94). Moreover, since establishing inconsistencies is a deductive, demonstrative endeavour, this form of argumentation is immune from the sorts of concerns Norton raises regarding inductive inference.

As an illustration of what I mean here, consider Tim Maudlin’s *The Metaphysics Within Physics* (Maudlin 2007). This is a work widely taken to represent an exemplary work of naturalistic metaphysics. It consists of a battery of arguments against (then) popular views in metaphysics – primarily, the broad system of Lewisian metaphysics, subsequently developed in the context of philosophy of physics by David Albert and Barry Loewer – all of which are based squarely on considerations of our best current physics. For example, he argues (in Chapter 2) that quantum mechanics is incompatible with the assumption of *separability*, so central to David Lewis’ transcription of the basic Humean program, and that the Humean interpretation of laws is not consistent with the uncontroversial fact that special relativity is a limit case of general relativity.¹² He argues (in Chapter 3) that contemporary particle physics, and specifically the ‘gauge principle’ that lies at the heart of it, refutes the age-old idea that properties can be thought of as universals. And he argues (in Chapter 4) that well-documented phenomena in particle physics, such as K-meson decay, are not compatible with the idea that physics fails to recognize any distinction between the past and the future – an assumption that is replete with consequences for the metaphysics of time. What he is doing throughout this book, then, is arguing that the best physics of our day *refutes* some otherwise live and popular metaphysical views.

To be sure, challenges to each and every one of these claims that physics refutes some-or-other metaphysical position have been raised by critics – in some cases persuasively. But the point for present purposes is simply that the methodology in play here seems completely different from the standard line presented in metaphysics in general regarding how metaphysics gets done. Maudlin’s point throughout is not that certain rival views aren’t intrinsically elegant, simple, or otherwise ‘lovely’, or that they are less lovely than those he favours; it is rather that, as he sees it, they are borderline untenable given the current state of physics. Of course, such rampant negativity in itself does not (or at least not obviously) get us to a positive view of metaphysics. Indeed, as Mauricio Suárez points out in his review of the book (aptly-titled ‘The Many Metaphysics Within Physics’), Maudlin’s arguments for the positive views he defends ‘all turn out to be essentially criticisms of Humeanism’ (Suárez 2009). Since, he argues, there are a variety of metaphysical packages that deny the basic Humean tenets, even if we accept Maudlin’s criticisms we are still faced with choices. It is at this juncture, of course, that recourse to virtues is supposed to be inevitable. But it seems to me that there is nothing preventing Maudlin (or anyone else) attempting to utilize physics against any one of these remaining views, and – as noted above – I am not aware of any convincing *a priori* argument as to why we can never expect the physics to whittle down our choices to one.

5 Conclusions

Given this jumble of approaches and questions, what in the end can we say about the ‘method of the metaphysics of science’? Is there a discernible method here at all? And what relationship, if any, does it stand in to that of science? Here one might be tempted to channel the spirit of Paul Feyerabend and say that there is nothing here that deserves to be dignified with the term ‘method’. Not only is it clear that different practitioners do different things, but taking this line may also resonate with the experience of any metaphysician who has struggled to complete the mandatory ‘Proposed Methods’ section of a big funding application. (What is it, exactly, that we do all day?)

I think we can say more, however. As is reflected in my statement of what I take to be the most exciting and persuasive examples of it to be, I think that if there is any method discernible in metaphysics of science – certainly when at its best – it is not the subjective assessment of ‘theoretical virtues’ but rather the method of *falsification*. This may be a little ironic, but it is worth taking seriously. Of course, the kind of falsification in play here is not exactly that which Popper meant in the case of scientific theories: there, falsification means ascertaining the falsity of a singular statement deduced from a universal law (Popper 2005, p. 10). Rather, it is meant in a more general sense as the obtaining of an *inconsistency* between a metaphysical claim and an accepted piece of physics. Given that the latter is treated as authoritative in metaphysics as I see it, in such cases it is the metaphysics that has to go.

Thinking of theories within the metaphysics of science as open to, and constantly subjected to, attempts at falsification invites some interesting and neglected questions. A central one is how informative metaphysical falsifications are in comparison with scientific theories, given the deep disanalogies that may – contra Quine – be argued to exist between the two. It has frequently been pointed out by critics of Popper that determining that a given theory is false may not tell you much at all about how the world in fact is. But given that metaphysical theories are so much more abstract than their scientific counterparts, could one not argue that it is much more informative to know the falsity of a metaphysical theory than it is in the case of a scientific theory? Debates in metaphysics, after all, often seem to carve the world into grand dualities (the debate over physicalism being over whether everything reduces to the physical *or not*; that over Humeanism being over whether there is primitive modality *or not*, etc). Given that this higher level of abstraction makes less room for alternative theories, if one knows that the world is *not* one way has one not potentially made a great leap forward to understanding how the world positively is, in a way that may be lacking in science itself?

Put that way, it again seems much less clear that metaphysical undeterminedness is an inevitable fact of life, despite that presumption having been so front-and-center in discussions on the methodology of metaphysics. The considerations here also suggest that it is by pushing the disanalogies between science and metaphysics that we get a better sense of what the metaphysics of science is all about and how it works, in stark contrast with the Quinean picture of seamless continuity that governs most contemporary discussions. For me,

then, if we are to progress with this issue first on the order of business will be to retire the problematic Quinean metaphors interwoven throughout the contemporary literature on the grounds that they may be distorting, rather than illuminating, the subject matter.

Notes

¹Both Premise 3 and Premise 4 will be taken for granted here. Premise 4 is simply undeniable (as a flick through the contemporary literature will repeatedly confirm). Premise 3 by contrast is highly contested. But this contestation is essentially the basis of the entire debate over scientific realism, and there is no point in debating this here: we may safely presume that only the scientific realist will maintain an interest in the metaphysics of science.

²(Note that e.g. (Kriegel 2013) also attacks the role of virtues in metaphysics in general, on the grounds that many pairs of theory are equally well supported by competing virtues.

³There are of course different recognized forms of underdetermination. For example, there is one according to which there is, for any theory, another specific theory with which it will always be equivalent (so-called ‘permanent’ underdetermination); there is another according to which for any theory, there is at any given time at least *some* other theory with which it is equivalent (but not necessarily always one and the same theory: this is ‘transient’ underdetermination). However, I don’t think we need to get into the subtleties to make the basic point here.

⁴ P. Kyle Stanford notes similarly that ‘Quine’s classic (1975) paper, so often cited as providing evidence for an important underdetermination predicament, simply blusters, “Surely there are alternative hypothetical substructures that would surface in the same observable ways”’ (Stanford 2001, p. 313).

⁵Note that I am not claiming that any of this answers the long-contested question of why in fact the scientific community at that time adopted one theory over another (although one can find the basic sentiment of the analysis expressed in e.g. the work of Kepler: see (Myrvold 2003, p. 46)). Note also that one could say that the Copernican model was more unified than its rival for these reasons; but if so, unity ceases to be an extra-empirical virtue. See Section 3.3 below.

⁶This case study appears in many places in the metaphysics literature to legitimate appeals to virtues in theory selection. For just one example, see (Goff 2019, p. 50).

⁷It is quite possible that scientists implicitly understand this connection. If so, that might account for the experimental work showing that scientists view virtues – especially that of unification – as ‘not merely pragmatic’. Partly as Schindler’s empirical work on scientist’s attitudes to virtues explicitly does not consider unification understood in terms of empirical accuracy (see Schindler *op cit*, p. 544), I do not think that this is something that his analysis tests for.

⁸To take one example, Alan Musgrave holds that appeals to simplicity are acceptable if we can show that ‘theories constructed under [the] aegis [of simplicity] are empirically successful, while theories which violate it are not... It may not be absurd to think that Nature is simple, if we can point to the empirical success of science in vindication of our belief’ (Musgrave 1985, p. 203).

⁹The ‘toolbox’ approach treats *a priori* metaphysics as a repository of concepts that may be useful to the metaphysician of science when building her own metaphysical theories. By treating them purely instrumentally, French and McKenzie remain agnostic on questions concerning the status of *a priori* metaphysical theories as true or justified. (Whether this is a coherent position could no doubt be debated.)

¹⁰Robin Hendry, for example, defends several positions in the metaphysics of chemistry, and takes ‘the metaphysics of chemistry’ to be that which ‘focuses on how things would be if chemistry’s account of those aspects of the world that it studies is broadly correct.’ (Hendry forthcoming, p. 1).

¹¹To see a concrete example of what I mean here, see (McKenzie 2022, Sec. 4.2).

¹²Maudlin takes ‘separability’ to be the principle that ‘the complete physical state of the world is determined by (supervenies on) the intrinsic physical state of each spacetime point (or each pointlike object) and the spatio-temporal relations between those points.’

References

- Acuña, P. (2014). “On the empirical equivalence between special relativity and Lorentz’s ether theory”. In: *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 46, pp. 283–302.
- Chakravartty, A. (2017). *Scientific Ontology: Integrating Naturalized Metaphysics and Voluntarist Epistemology*. Oxford University Press.
- Dorato, M. and Morganti, M. (2013). “Grades of Individuality. A pluralistic view of identity in quantum mechanics and in the sciences”. In: *Philosophical Studies* 163, pp. 591–610.
- Epstein, B. (2015). *The Ant Trap: Rebuilding the Foundations of the Social Sciences*. Oxford University Press.
- French, S. (2014). *The Structure of the World: Metaphysics and Representation*. Oxford University Press.
- French, S. and McKenzie, K. (2012). “Thinking Outside the Toolbox: Towards a More Productive Engagement Between Metaphysics and Philosophy of Physics”. In: *European Journal of Analytic Philosophy* 8.1, pp. 42–59.
- (2016). “Rethinking Outside the Toolbox: Reflecting Again on the Relationship between Philosophy of Science and Metaphysics.” In: *Metaphysics in Contemporary Physics*. Brill Rodopi, pp. 25–54.
- Goff, P. (2019). *Galileo’s Error: Foundations for a New Science of Consciousness*. Vintage.
- Henderson, L. (2014). “Bayesianism and inference to the best explanation”. In: *The British Journal for the Philosophy of Science*.
- Hendry, R. F. (forthcoming). “Structure, Essence and Existence in Chemistry”. In: *Ratio*. DOI: 10.1111/rati.12387.
- Huggett, N. (2006). “The Regularity Account of Relational Spacetime”. In: *Mind* 115.457, pp. 41–73. DOI: 10.1093/mind/fzl041.
- Kriegel, U. (2013). “The epistemological challenge of revisionary metaphysics”. In: *Philosophers* 13.
- Kuhn, T. S. (1957). *The Copernican revolution: Planetary astronomy in the development of Western thought*. Vol. 16. Harvard University Press.
- Ladyman, J. and Ross, D. (2007). *Every Thing Must Go: Metaphysics Naturalized*. Oxford University Press.
- Maudlin, T. (2007). *The Metaphysics Within Physics*. Oxford University Press.
- (2022). *Speculations in High Dimensions*.
- McKenzie, K. (2022). *Fundamentality and Grounding*. Cambridge University Press.

- McSweeney, M. M. (2023). “Metaphysics as Essentially Imaginative and Aiming at Understanding”. In: *American Philosophical Quarterly* 60.1, pp. 83–97. DOI: 10.5406/21521123.60.1.07.
- Moore, A. W. (2011). *The Evolution of Modern Metaphysics: Making Sense of Things*. Cambridge University Press.
- Musgrave, A. (1985). “Realism versus constructive empiricism”. In: *Images of science*, pp. 197–221.
- Myrvold, W. C. (2003). “A Bayesian Account of the Virtue of Unification”. In: *Philosophy of Science* 70.2, 399–423. DOI: 10.1086/375475.
- Norton, J. D. (2003). “Must Evidence Underdetermine Theory”. In: *The Challenge of the Social and the Pressure of Practice*, pp. 17–44.
- Norton, J. D. (2021). *The material theory of induction*. University of Calgary Press.
- Novick, A. (2017). “Metaphysics and the vera causa ideal: The nun’s priest’s tale”. In: *Erkenntnis* 82.5, pp. 1161–1176.
- Paul, L. A. (2012a). “Metaphysics as Modeling: The Handmaiden’s Tale”. In: *Philosophical Studies* 160.1, pp. 1–29. DOI: 10.1007/s11098-012-9906-7.
- Popper, K. (2005). *The logic of scientific discovery*. Routledge.
- Psillos, S. (2011). “Moving molecules above the scientific horizon: On Perrin’s case for realism”. In: *Journal for General Philosophy of Science* 42, pp. 339–363.
- Ritchie, J. (2022). “On the Continuity of Metaphysics with Science: Some Scepticism and Some Suggestions”. In: *Metaphilosophy* 53.2-3, pp. 202–220. DOI: 10.1111/meta.12547.
- Saatsi, J. (2017). *Explanation and explanationism in science and metaphysics*. Oxford University Press Oxford.
- Saatsi, J. et al. (2009). “Grasping at Realist Straws”. In: *Metascience* 18.3, pp. 355–390.
- Schindler, S. (2022). “Theoretical virtues: do scientists think what philosophers think they ought to think?” In: *Philosophy of Science* 89.3, pp. 542–564.
- Sider, T., Hawthorne, J. and Zimmerman, D. W., eds. (2008). *Contemporary Debates in Metaphysics*. Malden, MA: Blackwell.
- Sober, E. (2009). “Parsimony arguments in science and philosophy—A test case for naturalism p”. In: *Proceedings and Addresses of the American Philosophical Association*. JSTOR, pp. 117–155.
- Stanford, P. K. (2001). “Refusing the Devil’s Bargain: What kind of underdetermination should we take seriously?” In: *Philosophy of Science* 68.S3, S1–S12.
- (2017). “So Long, and Thanks for All the Fish”. In: *Metaphysics and the Philosophy of Science: New Essays* 127.
- Suárez, M. (2009). “The Many Metaphysics within Physics: Essay Review of The Metaphysics within Physics by Tim Maudlin”. In: *Studies in History and Philosophy of Modern Physics* 40, 273–6.
- Van Fraassen, B. C. (1980). *The Scientific Image*. Clarendon Press: Oxford.
- Vickers, P. (2022). *Identifying future-proof science*. Oxford University Press.