Erkenn (2017) 82:1161–1176
DOI 10.1007/s10670-016-9863-1

ORIGINAL RESEARCH

Metaphysics and the *Vera Causa* Ideal: The Nun's Priest's Tale

Aaron Novick[1]

Received: 15 August 2016 / Accepted: 10 November 2016 / Published online: 16 December 2016

Abstract L.A. Paul has recently defended the methodology of metaphysics on the grounds that it is continuous with the sciences. She claims that both scientists and metaphysicians use inference to the best explanation (IBE) to choose between competing theories, and that the success of science vindicates the use of IBE in metaphysics. Specifically, the success of science shows that the theoretical virtues are truth-conducive. I challenge Paul's claims on two grounds. First, I argue that, at least in biology, scientists adhere to the *vera causa* ideal, which allows the theoretical virtues to play a much more limited role in scientific reasoning than Paul requires for metaphysical reasoning. The success of biology thus does not vindicate the methodology of metaphysics. Second, I argue that, at least in many cases, the successful reliance on the theoretical virtues in scientific contexts shows only that the theoretical virtues are truth-conducive within those local contexts, and not that they are truth-conducive generally. The upshots are (1) that Paul's defense of the methodology of metaphysics fails, and (2) that any attempt to rescue her defense must pay more careful attention to what precisely is vindicated by successful science.

> The hypothesis put forth in the "Vestiges," though it had the merit of connecting the organic evolution with the cosmical evolution, uniting the hypotheses of Lamarck and Meckel with the nebular hypothesis of Kant and Laplace, laboured under the general disadvantage of reposing on two principles which only a metaphysician could accept as *veræ causæ*. George Henry Lewes (1868, p. 356).

✉ Aaron Novick
amn61@pitt.edu

1 Department of History and Philosophy of Science, University of Pittsburgh, 1017 Cathedral of Learning, 4200 Fifth Avenue, Pittsburgh, PA 15213, USA

1 The Nun's Priest's Tale

In "The Nun's Priest's Tale" from Chaucer's *Canterbury Tales*, a chanticleer is caught by a clever fox, having been tickled by the fox's silver tongue. Happily, the chanticleer, by a clever trick of its own, escapes, and learns not to trust flattery:

> Lo, swich it is for to be recchelees.
> And necligent, and truste on flaterye…

A similar moral applies to contemporary discussions of methodology in metaphysics.

Specifically, some metaphysicians (e.g. Hawley 2006; Paul 2012) have flattered themselves that, in relying on inference to the best explanation (IBE), they share a methodology with the sciences, and that the evident successes of science vindicate the use of IBE in metaphysics. This rests, in my view, on a mistaken picture of scientific reasoning, with the result that metaphysicians have adopted an at worst problematic, at best ill-defended methodology.

Recently, Paul (2012) has offered a thorough and explicit defense of the methodological continuity of science and metaphysics. She claims that IBE is widely used in both the sciences and in metaphysics, and thus that "the success of science indirectly confirms metaphysical theories that maximize the theoretical virtues" (Paul 2012, p. 25). At the heart of her argument is the claim that IBE, in both metaphysics and the sciences, involves preferring theories that maximize certain theoretical virtues (e.g. simplicity)—the same virtues in both cases. Insofar as science is successful, it shows that these virtues are truth-conducive. Thence the vindication of metaphysics.

I contend that Paul mischaracterizes scientific methodology in two ways. First, she overestimates the role these virtues play in scientific inference. By reconsidering a paradigm case of IBE in science, I illustrate the importance of the *vera causa* ideal in biology. This ideal allows IBE to play only a narrowly circumscribed role in scientific inference, narrower by far than the role they play in Paul's methodology. This methodological discrepancy between metaphysics and biology undermines Paul's claim that science and metaphysics are methodologically continuous.

But perhaps Paul might look outside of biology to find successful sciences that rely upon IBE. Here the second way that Paul mischaracterizes scientific methodology becomes relevant. Paul's argument assumes that the truth-conduciveness of particular theoretical virtues can be justified globally. I argue, using parsimony analysis in phylogenetic systematics as an example, that such virtues are often shown to be truth-conducive only locally. Thus, in looking outside biology for successful uses of IBE, Paul must show not only that the theoretical virtues are truth-conducive in some scientific inferences, but must show also that the justification for using these theoretical virtues transfers from science to metaphysics.

The upshot is that Paul's defense of the use of IBE in metaphysics fails as presented, and any attempt to rescue her argument must proceed via a closer examination of scientific reasoning than she offers.

2 Paul on the Methodology of Metaphysics

Paul defends three theses about metaphysical inquiry. First, metaphysicians and scientists study distinct problems. Metaphysicians engage in "*ontological* projects" in which they seek "systematic, general truths" about "the fundamental natures of the world" (Paul 2012, p. 4). These natures include numbers, composition relations, properties, and the like. These natures are "metaphysically prior" to the material entities and processes studied by science, which instantiate these natures. Scientists, for instance, study how two hydrogen atoms and an oxygen atom jointly compose a water molecule, but do not study composition per se.

Second, scientists and metaphysicians share a methodology. Given some explananda, both develop competing explanatory models, which they then assess by considering to what extent each model possesses the theoretical virtues. This form of model assessment is known as inference to the best explanation (Lipton 2004). The basic idea is that multiple models that are empirically equivalent (in the minimal sense that they are consistent with all available data) may nonetheless be distinguished in terms of their explanatory power. The model that maximizes the theoretical virtues (e.g. simplicity, avoidance of ad hoc modifications, fruitfulness, etc.) may be inferred to be true on the grounds that it provides the best explanation of the phenomena in question.[1]

Third, the success of science vindicates the methodology of metaphysics by showing that the theoretical virtues are truth-conducive. Given some form of epistemic scientific realism (the view that our best current scientific theories are at least approximately true), and in particular realism about theories supported via IBE, we may conclude that the theoretical virtues are truth-conducive in science. If truth-conducive in science, they should be truth-conducive everywhere, including metaphysics (Paul 2012, p. 26). Metaphysicians thus reason in truth-conducive ways.

I shall grant the first thesis and criticize the second. Because the second thesis is an essential premise supporting Paul's third thesis, my criticism shall thus undermine the third thesis as well. I grant, further, that Paul accurately characterizes the methodology of much (though not all[2]) contemporary metaphysics. Paul errs, not in her description of the methodology of metaphysics, but in her description of the methodology of science. This methodology is distinct from that used in metaphysics, and the success of science does not vindicate the methodology of metaphysics.

My fundamental criticism is that Paul mischaracterizes the use of IBE in the sciences, or at least in biology. Here it is important to distinguish Paul's defense of metaphysics and my critique thereof from a superficially similar defense of

[1] For an especially clear instance of this methodology in action, see Markosian (2008). Markosian argues that, while unrestricted composition enjoys certain theoretical virtues (most notably simplicity and consistency with our background commitment to denying metaphysical vagueness), it flies in the face of our intuitive judgments about what objects exist and brings along commitments to other controversial or implausible views, and so should be rejected in favor of restricted composition.

[2] Henceforth, I will drop this caveat and speak simply of "metaphysics," but I will be referring only to metaphysics of the sort that Paul describes and not all of metaphysics.

metaphysics. Like Paul, Hawley (2006) defends the methodology of metaphysics by linking it to scientific realism, but she does so in an importantly different way. Speaking of a form of skepticism about metaphysics defended by van Fraassen (2002), Hawley (2006, p. 454) writes:

> Such scepticism is typically based upon rejection of the inferential methods of metaphysicians, and rejection of inference to the best explanation in particular. As such, it is incompatible with standard versions of scientific realism.[3]

Hawley's point is that the standard defense of scientific realism relies on inference to the best explanation (in the form of the "no miracles" argument), so any critique of metaphysics on these grounds is tied to anti-realism. Here the crucial comparison is between the methodology of metaphysicians and the methodology of scientific realist philosophers.

Paul's defense of metaphysics, by contrast, compares the methodology of metaphysicians to the methodology of scientists. This is a virtue of her defense, because the successes of science are vastly more impressive than the successes of the (unpersuasive) standard defense of scientific realism. My criticism that Paul has misdescribed scientific methodology is fully compatible with believing that the actual methodology of science is truth-conducive, and thus my critique is compatible with scientific realism.

3 The *Vera Causa* Ideal

The only way to determine whether the methodology of metaphysics resembles that of the sciences is to compare the two. A full survey of scientific methodology is impossible in this short space, so I shall focus on a single science, biology. The choice is not arbitrary, for Paul (2012, p. 12) herself makes the suggestion:

> Nevertheless, the importance of the role of inference to the best explanation as a means to grasp scientific truths about the nature of the world is well confirmed, at least when we look past the context of fundamental physics to wider scientific contexts, for example, to the context of evolutionary biology.

However, Paul does not examine any particular examples of the use of IBE in evolutionary biology. In the next two sections, I consider two cases of the use of IBE in biology, arguing that these cases illustrate the role played by the *vera causa* ideal. First, however, the basic features of that ideal must be introduced.

The roots of the *vera causa* ideal lie in Newtonian philosophy of science. John Herschel expanded on these ideas in a form that proved of particular importance for biology, due to his influence on Charles Lyell, Charles Darwin, and others. Herschel (2010, §141) wrote:

> Whenever, therefore, any phenomenon presents itself for explanation, we naturally seek, in the first instance, to refer it to some one or other of those real

[3] See also Hawley (2006, §5) for further discussion.

causes which experience has shown to exist, and to be efficacious in producing similar phenomena.

That is, when claiming that a certain causal process is responsible for producing a certain phenomenon, we ought to possess independent reason to believe that said process exists and is competent to produce similar phenomena (cf. Hodge 1992).

Satisfying the *vera causa* ideal involves completing three epistemic tasks. When one wishes to explain some phenomenon in terms of some cause, one must show that the cause *exists*, that it is *competent* to explain phenomena similar to the explanandum, and that it is in fact *responsible* for producing the explanandum. Though the specific language of *veræ causæ* has fallen out of favor, biologists still make these distinctions in their work.[4]

These tasks are distinguishable in principle, though, since we only know entities through their causal powers, the first two stages are not always distinct. Furthermore, since responsibility presupposes competence and competence presupposes existence, it can seem that there is a necessary temporal order for completing these tasks. In practice, however, a case for the responsibility of some cause may be made in advance of having any evidence for the existence of that cause. The *vera causa* ideal requires, however, that the explanation not be accepted (except tentatively) until the case for existence and competence has been made.

The *vera causa* ideal specifies how these three tasks are to be accomplished, and in doing so allows the theoretical virtues to play only a limited epistemic role. In showing existence and competence, the ideal requires one to provide evidence that is independent of the putative cause's hypothetical explanatory power. In other words, that a hypothesis invoking that cause maximizes the theoretical virtues does not suffice to warrant acceptance of the existence and competence of that cause. The existence and competence of a cause are to be shown on non-explanatory grounds.

The theoretical virtues are allowed a role only in accomplishing the third task, making the case for responsibility. That an entity or process exists and is competent to produce a phenomenon does not entail that it actually does produce that phenomenon. Theoretical virtues may help to bridge this gap. When choosing between two (or more) causes known to be competent to produce some effect, determining which furnishes the better explanation may allow one to make the decision. But they must be known causes—i.e. their existence and competence must have been demonstrated on non-explanatory grounds.

When biological reasoning fails to satisfy the *vera causa* ideal, i.e. when it invokes as causes entities or processes not known (on non-explanatory grounds) to exist and to be competent to produce the phenomenon in question, biologists are hesitant to accept its conclusions. Biologists mistrust arguments that rely on theoretical virtues to support, not merely the responsibility, but also the existence and competence of the entities and processes they invoke. While an explanation that fails to satisfy the ideal may be taken as pursuit-worthy, it is not considered belief-worthy.

[4] E.g. Arthur (2001, p. 272): "acceptance of the occurrence of a phenomenon does not necessarily imply acceptance either of its commonness or of its power."

4 The Ideal in Action I: Evolution by Natural Selection

Charles Darwin, who read Herschel's *Preliminary Discourse* with great enthusiasm, attempted to adhere to the *vera causa* ideal in his scientific work (Hodge 1992). By considering the reasoning he offered in support of two hypotheses—evolution by natural selection and pangenesis—we can elucidate the nature of the *vera causa* ideal and its role in biological reasoning. Further, by considering how Darwin's peers responded to his work, we can see that this ideal was generally accepted within the biological community of his time. Although Paul does not specify where in evolutionary biology IBE is used, Darwin's reasoning in *On the Origin of Species* is often taken as a paradigm case of IBE's use in science and is thus a reasonable guess as to what Paul had in mind.[5]

Darwin's (2001) reasoning in the *Origin* self-consciously conformed to the *vera causa* ideal. The structure of the work reflects the distinction between existence, competence, and responsibility. Per Hodge (1992, p. 462), Darwin's argument involved three "evidential cases":

> …first, a case for [natural selection's] *existence* as a causal process going on in the world; second, a case for its *adequacy*, its competence to produce, adapt, and diversify species; and, third, a case for its *responsibility*, for, that is, its having produced the species now living and the extinct species found in fossils.

Darwin's argument for the existence and competence (=adequacy) of natural selection as an evolutionary mechanism occupies the first nine chapters of the *Origin*, while chapters ten through thirteen concern its responsibility. In the first nine chapters, Darwin showed: first, how selection operates under domestication (demonstrating its existence and some of its causal powers); second, that the conditions necessary for this principle to operate obtain in nature; third, that certain factors that limit the power of natural selection under domestication (e.g. limited time) do not apply in nature; and fourth, that arguments purporting to show that natural selection is *in principle* incompetent (e.g. to produce complex organs) are unsound.[6]

The evidence offered in these chapters did not concern the explanatory power of natural selection, but instead used established facts from domestic breeding and natural history to show that natural selection exists and (at least plausibly) is competent to explain the diversity of life. Only after establishing this did Darwin leverage the explanatory power of his theory to argue that natural selection is responsible for the history and current diversity of life.[7]

[5] The following authors treat Darwin as offering an IBE in the *Origin*: Thagard (1978), Lipton (2007), Doppelt (2011), Dawes (2013), Lewens (2015), Brössel (2015), Park (2015). Of these, only Lewens explicitly recognizes the Herschelian influence on Darwin's "IBE."

[6] Lennox (1991) has shown the crucial role that thought experiments played for Darwin in addressing these "in principle impossible" arguments.

[7] Lennox (2009, pp. 122–123) thinks that Hodge overstates the extent to which Darwin took the case for responsibility to concern the responsibility of natural selection (as opposed to merely supporting common descent). Though it is true that natural selection is irrelevant for explaining many of the phenomena

Darwin thus held himself to the *vera causa* ideal. His peers' response was mixed. They widely accepted that evolution occurs, but disputed the importance of natural selection. Both supporters and detractors based their assessments on the extent to which they believed Darwin had satisfied the *vera causa* ideal.

Among supporters, Lewes (1868, p. 356) argued that Darwin's major achievement lay in showing that natural selection is "demonstrably the cause of *much* variation." This marked a fundamental advance over prior transmutationists, who, Lewes argued, had failed to invoke *veræ causæ* (see this paper's epigraph). Despite this advance, Lewes (1868, p. 355) nonetheless cautioned: "Darwinism is undoubtedly a better explanation than any of its forerunners; but it will probably give place to some successor, as the hypotheses of Geoffroy St. Hilaire, Meckel, Lamarck, Bonnet, and Robinet gave place to it."

By contrast, Jenkin (1867), one of Darwin's more strident critics, offered powerful arguments challenging the competence of natural selection to accumulate variation past species boundaries. These challenges were not fully dispelled until the twentieth century, when population geneticists showed how natural selection could work to accumulate variations inherited according to Mendelian principles.[8]

5 The Ideal in Action II: Darwin's "Provisional Hypothesis of Pangenesis"

The case of evolution by natural selection may be contrasted with the case of Darwin's reasoning in support of his "provisional hypothesis of pangenesis" (Darwin 1988, chap XXVII). According to this hypothesis, each part of an organism's body throws off minute particules called "gemmules." These disperse throughout the body and collect to constitute the sexual elements. Under the right conditions, they are capable of multiplying by self-division and becoming fully developed units similar to the parts from which they originated.

Darwin (1988, p. 303) divided the evidence for this theory into two kinds. One line of evidence served to show that the "necessary assumptions" of his theory (that gemmules exist and exhibit the requisite behaviors) are not "improbable," while the other served "to bring under a single point of view the various facts" that, in Darwin's view, any theory of heredity needed to explain. Abstractly described, this scheme answers to the *vera causa* ideal: explanatory considerations play a substantial role in establishing responsibility (the second line of evidence), but only after the existence and competence of the entities invoked have been established on non-explanatory grounds (the first line of evidence).

Unfortunately, the evidence Darwin had to offer in support of his "necessary assumptions" was rather meager. It primarily consisted of "analogous facts."

Footnote 7 continued
discussed in the later chapters of the *Origin*, it does play a role in explaining some of them (e.g. it is essential in the discussion of embryology; Darwin 2001, pp. 439-450).

[8] For a detailed history, see Gayon (1998). Gayon (1998, p. 398) describes the post-*Origin* history of Darwinism as in large part constituting an attempt "to go beyond indirect proof through explanatory power," a description in keeping with the account of methodological standards in biology provided here.

Darwin marshaled cases in which known biological entities exhibit behavior similar to the purported behavior of gemmules. Darwin hoped to show that it would not be impossible for gemmules to exhibit those behaviors. Darwin had no evidence, however, that showed that gemmules actually do exist and actually do exhibit those behaviors. The brunt of the evidence work favoring their existence and competence was thus played by the ability of gemmules to bring various facts "under a single point of view."

The reception of Darwin's theory was primarily negative, and critics emphasized the lack of non-explanatory evidence for the existence and competence of gemmules. Lewes' reaction is again representative.[9] Whereas, in the case of natural selection, "the sole part played by pure inference in the construction of the hypothesis is the inferring that what is proved to be true in many cases is also true in all," in the case of pangenesis "all its elements are inferences; not one of them can be admitted as proven" (Lewes 1868, p. 507). Lewes used 'inference' to mark reliance on explanatory power, and this criticism may thus be rephrased: Darwin's case for natural selection meets the *vera causa* ideal, while his case for pangenesis does not.

Lewes' criticism illustrates the point made above (§3) that hypothetical explanatory power may establish a hypothesis as pursuit-worthy but not as belief-worthy. Lewes (1868, p. 503) granted that pangenesis "surpass[ed] all previous attempts in the same direction" and was worthy of further investigation. Before it could be accepted, however, independent evidence for the existence and competence of gemmules was needed. This evidence never materialized, and today pangenesis is largely forgotten.

The contrast between the reception and fate of Darwin's theory of evolution by natural selection and Darwin's theory of pangenesis illustrates the importance of the *vera causa* ideal in biological reasoning. The moral to draw is not that theoretical virtues are not truth-conducive (but see below, §7), but that biologists ask theoretical virtues to do only limited evidential work. Biology's success at best shows the theoretical virtues to be truth-conducive within these limits.

One might grant that biologists adhere to the *vera causa* ideal while questioning whether they ought to do so, but this line of critique is not available to someone who, like Paul, claims that the actual methodology of science vindicates contemporary metaphysical theories.[10]

6 True Causes and Metaphysics

I have shown how the *vera causa* ideal allows the theoretical virtues to play only a limited role in scientific inference. It remains to compare the use of IBE in biology to its use in metaphysics. If metaphysicians cannot satisfy the *vera causa* ideal, if they rely on theoretical virtues to show the existence and competence of

[9] Cf. Delpino (1869) and Mivart (2009, Chap. X) for similar criticisms.

[10] See Novick and Scholl (manuscript) for explicit justification of adherence to the *vera causa* ideal as good epistemic policy.

metaphysical natures, then Paul's claim of methodological continuity between metaphysics and science is unsupported. As this claim of methodological continuity is an essential premise in Paul's argument that the success of science vindicates the methodology of metaphysics, if my criticisms are well-placed, Paul's main conclusion also lacks support.

It is difficult to see how metaphysicians could satisfy the *vera causa* ideal. Because metaphysical natures do not stand in causal relations to material entities, "the features of the world described by metaphysicians are not manipulable or testable the way the features of the world described by science are" (Paul 2012, p. 17). Furthermore, because scientific instruments work by creating informative causal links between an entity of interest and human perceptual systems, there can be no instruments "we could use to detect the presence of numbers, or the presence of composition, or of necessity, or the category of properties" (Paul 2012, p. 18).

Because the primary source of independent evidence that an invoked entity or process is a *vera causa* comes from the perception or manipulation of the entity or process in question, the non-causal status of metaphysical natures suggests that a metaphysician cannot satisfy the *vera causa* ideal. This is not the superficial objection that, because metaphysical natures are not causes, they ipso facto cannot be true causes. The *vera causa* ideal, applied to metaphysics, requires that there be independent evidence of the existence and competence of the natures invoked in metaphysical explanations. In the case of material entities and processes, such evidence is acquired by exploiting their causal properties, but that is not possible for metaphysical natures. Some other source of independent evidence is necessary. The worry is that there is, in principle, no such source.[11]

Paul (2010) argues that ordinary judgments about metaphysical features of the world may provide empirical, quasi-perceptual evidence for metaphysical theories. Leaving aside worries about using such judgments as evidence in metaphysics,[12] they might appear to provide a source of independent evidence for the natures that appear in metaphysical models. However, Paul herself does not think of them in this way. In Paul's methodology, metaphysical theories may lead to the rejection of ordinary judgments provided they discharge "an especially heavy explanatory burden" (Paul 2010, p. 468). In debates over composition, for instance, we may reject our apparent experience of composition if we are sufficiently motivated by the ontological sparseness of compositional nihilism.

[11] This criticism is superficially similar to the criticism of "non-naturalistic metaphysics" offered by MacLaurin and Dyke (2012). They distinguish between scientific theories, which appeal to observation, and non-naturalistic metaphysical theories, which merely appeal to "alternative theoretical virtues." Their argument, however, presupposes that these virtues are "more aesthetic than epistemic" and that theory choice on such grounds "is carried out completely independently of [the theories'] likely truth-value" (MacLaurin and Dyke 2012, p. 304). Because they presuppose this and do not argue for it, their argument furnishes no response to Paul's claim that these virtues are truth-conducive. See also McLeod and Parsons (2013) for criticism of their attempt to define which metaphysical theories have observational consequences, and see Dyke and MacLaurin (2013) for a response thereto.

[12] Benovsky (2016, Chap. 6) argues that many supposed tensions between perceptual experience and "counterintuitive" metaphysical theories (e.g. eliminativism about macroscopic objects) are not genuine tensions. I agree, but will not pursue the worry here.

The same is not true for *veræ causæ*: they are not defeasible in this way. The judgment that some cause is a *vera causa* is defeasible, but not by explanatory considerations that concern only its responsibility. Thus, for instance, proponents of the neutral theory of evolution deny that the explanation of evolutionary change provided by the theory of natural selection is as good as is often thought, but they conclude only that natural selection is not responsible for as much evolutionary change as often thought, not that selection does not exist at all. Likewise, proponents of epigenetic inheritance do not deny that genetic inheritance exists, but merely that it is the only form of inheritance of evolutionary importance.

Metaphysicians thus cannot satisfy the *vera causa* ideal. At this juncture, Paul might grant the point, but insist that the metaphysician may still use IBE as a second-best method, useful because there is none better.[13] I do not see that I can object to this, but I cannot object precisely because I do not think this is an attractive position for Paul. One of Paul's (2012) aims was to show that, given the acceptance of scientific realism, the success of science vindicates the use of IBE in metaphysics, and thus allows for realism in metaphysics.[14] What I have shown, however, is that biological uses of IBE vindicate only those uses of IBE that are restricted to establishing responsibility for known causes. Uses of IBE that are not so restricted are not considered to warrant belief. In these instances, the theoretical virtues are taken not to be truth-conducive, and the appropriate attitude to take when IBE is all that is on offer is agnosticism. The model may be pursuit-worthy, but should not be treated as belief-worthy. Importantly, this restriction does not come about simply because IBE is "second-best." When Lewes and others criticized Darwin's IBE in support of pangenesis, there was no other way to reason about heredity. In the case of theories of heredity, the requisite tools to meet the *vera causa* ideal did not exist until the twentieth century (Novick and Scholl manuscript). In the late nineteenth century, IBE was the best they had, but it was still recognized as being not good enough.

This line of response thus comes at the cost of giving up Paul's defense of realism in metaphysics. Nor is this the only cost. While hypothetical explanatory power may not warrant believing a theory, it can suggest that the theory is worthy of further development. In biology, this development comes precisely in the form of seeking to provide independent evidence showing that the causes invoked by the model exist and are competent to produce the requisite effects. The metaphysician, by contrast, appears to be left with merely pursuit-worthy theories without any means by which to pursue them further. She can, of course, uncover further virtues and vices of her models, but will still be left with a merely pursuit-worthy theory in the end.

[13] Paul has offered this line of defense in personal communication.

[14] At the end of her paper, Paul argues for the importance of the co-existence of multiple competing models in metaphysics. This is, in one sense, a form of epistemic anti-realism, for it requires that none of these models be so definitively supported as to achieve consensus. But the reason for this, on Paul's view, is our inability to know which model actually maximizes the virtues (since different models enjoy different virtues, and it is not obvious how to weight the virtues). She is thus still committed to a methodological realism about IBE: if we knew which model maximized the virtues, we could infer to its truth.

The realist metaphysician is therefore ill-advised to use IBE as a "second-best" method. I see four ways forward for the defender of metaphysics. First, she may show that, despite the limitations surveyed above, some analog of the *vera causa* ideal is met in metaphysics. Second, she may show that the use of IBE in sciences other than biology more closely resembles its use in metaphysics. This task is not accomplished in Paul's articles (2010, 2012), but neither is it ruled out of contention by the arguments provided here, which concern only biology. I consider this strategy further in the next section. Third, she may continue relying on IBE to adjudicate between competing metaphysical models, but accept that the theoretical virtues are not truth-conducive, though this would require abandoning Paul's realist aspirations.[15]

Finally, the metaphysician may abandon the use of IBE and seek some other methodology.

7 When is it Truth-Conducive to Favor Parsimonious Theories?

In the foregoing discussion, I have taken for granted that, if successful biology relied on the theoretical virtues, this would show the virtues to be truth-conducive, not just in biological reasoning, but in reasoning in any discipline. This assumption is crucial to Paul's argument, for the path by which the success of science purportedly vindicates the methodology of metaphysics runs through showing that the theoretical virtues are truth-conducive in such a general manner. I have been able to grant this assumption for the sake of argument because I have been arguing that biology relies on the theoretical virtues for a more limited set of tasks than metaphysicians require. If that is correct, then it does not matter whether theoretical virtues are generally truth-conducive when asked to serve these limited roles: Paul still does not get what she needs.

However, it is clear from discussion with Paul that, of the four responses considered above—show that metaphysicians can meet the *vera causa* ideal, show that IBE is used (and reliable) in scientific disciplines other than biology, use IBE in metaphysics at the cost of giving up realist aspirations, or abandon the use of IBE in metaphysics—she favors the second. I am in favor of the attempt by metaphysicians to study the methodologies of the various sciences in order to determine what can be usefully transported between disciplines. However, if this avenue is to be pursued, the assumption that locally successful reliance on the theoretical virtues provides evidence that they are generally truth-conducive cannot be accepted for the sake of argument. It must be scrutinized.

[15] Benovsky (2016, Chap. 7) defends something like this view. He argues that the theoretical virtues do not by themselves decide between competing theories. They must be weighted, and what decides how a philosopher weights the virtues is that philosopher's aesthetic taste. As a result, Benovsky defends a kind of instrumentalism about metaphysics, according to which the primitive postulates of metaphysical theories are not taken to exist, but merely to serve to systematize our concepts. He writes: "What we do when we say that there is a substratum is not to say what there is in the world, rather, we introduce a new theoretical concept that allows [us] to systematize, organize, and understand the concepts of material object and property in such a way that we have a satisfactory answer to the questions we started with" (Benovsky 2016, p. 126).

Specifically, there is a substantial issue that threatens any argument that the success of science shows the theoretical virtues to be truth-conducive. The worry is that, on close inspection, the justification for relying on a particular theoretical virtue in a particular discipline will turn out to be local to that discipline. On this view, appeals to, say, simplicity may turn out to be mere surrogates for patterns of reasoning justified by our knowledge of local features of the domain under study. Even if a genuinely trans-disciplinary framework for understanding simplicity can be recognized,[16] the justification for thinking it truth-conducive may not be trans-disciplinary. If justification is local in this way, then the successful reliance on a theoretical virtue in a particular scientific context cannot support reliance on that virtue in metaphysics, *unless* it can be shown that the justification transfers across contexts. I illustrate these worries by considering one of the most commonly invoked theoretical virtues: simplicity (parsimony). In doing so, I build upon recent work by Sober (2015) and Norton (manuscript).

Specifically, I consider the case of cladistic parsimony. On a first glance, parsimony reasoning in phylogenetic systematics seems like a prime case that metaphysicians might cite as an instance of the reliable use of IBE in the sciences.[17] However, when we consider the manner in which parsimony analysis is justified, we find that it fits the pattern diagnosed by Norton (manuscript): "Good invocations of simplicity are really veiled references to background facts or assumptions whose content functions to license the relevant inductive inference." The justified use of cladistic parsimony rests on substantial assumptions about the nature of the evolutionary process, and not on any assumption that simplicity is a generally truth-conducive theoretical virtue.

Parsimony analysis is one means of reconstructing phylogenies. Described qualitatively, it works as follows.[18] First, the taxa whose phylogenetic relationships one hopes to discover are scored on a variety of characters (that can come in 2+ character states), such that their similarities and differences are noted. It is hypothesized that similarities are more likely to be due to shared descent than to convergent origin.[19] This amounts to favoring the simpler hypothesis, since shared descent implies that a particular character state arose only once, while convergent evolution implies at least two origins for that character state. The data thus encode numerous hypotheses of shared descent. Unfortunately, no single phylogenetic tree can preserve the truth of every such hypothesis. Thus, second, parsimony analysis is performed. The goal is to find the phylogenetic tree that maximizes preservation of

[16] Sober (2015, Chap. 2) identifies two such frameworks for understanding appeals to simplicity in scientific inference.

[17] In fact, the reliability of parsimony analysis is controversial, and there are other ways of reconstructing phylogenies (see Wiley and Lieberman 2011, Chaps. 6–7; Sober 2015, Chap. 3). Here I will skip over this controversy, as it does not affect my point: the reasons for being skeptical of parsimony analysis are just as local as the reasons to think it justified.

[18] There are many algorithms for actually conducting parsimony analysis, all of which make different assumptions and so are justified only insofar as those assumptions hold. For discussion, see Wiley and Lieberman (2011, Chap. 6).

[19] Wiley and Lieberman (2011, p. 153): "The sharing of character states is always evidence that those taxa that share a character state are related unless the weight of other evidence dictates that they are of independent origin."

these initial hypotheses of relationship, i.e. the tree that minimizes the number of independent character state origins.[20] Parsimony analysis in this sense favors the simplest phylogenetic tree relative to the dataset.

What justifies the use of parsimony analysis is a set of substantial assumptions about the evolutionary process.[21] Specifically, evolution must proceed in such a way that a single origin of a character is in fact more likely than multiple independent origins. However, under the right conditions (a combination of an evolutionarily labile character and shared selection pressures), homoplasy (convergent evolution of the same character state) may be a likelier hypothesis than homology (shared origin). So long as those conditions are met only rarely for the characters considered in the dataset, there is no problem, but as homoplasies accumulate, the ability of parsimony analysis to approximate the true phylogeny disintegrates. One solution is to attempt to determine which characters are unreliable indicators of phylogeny and exclude them from consideration. Another solution is to differentially weight characters in a manner reflective of their differential propensity to homoplasy.[22] (Indeed, exclusion of characters is simply the extreme case of differential weighting in which the excluded characters are given zero weight.)

Character selection and weighting also involves the assumption of the evolutionary independence of the characters considered. That is, if two taxa share five independent character states, that can be explained by at minimum five independent evolutionary origins, one for each character state. By contrast, if these character states are not independent, possibly only a single evolutionary origin needs to be postulated. Wiley and Lieberman (2011, p. 189) provide an example of why this assumption matters:

> Consider a hypothetical case in which one clade is supported by 10 unique and unreversed synapomorphies and the alternative is supported by two unique and unreversed synapomorphies. What if the 10 unique and unreversed synapomorphies corroborating the first clade are not independent of each other while the 2 synapomorphies supporting an alternative clade are independent? Then the 10 synapomorphies really represent a single syanopomorphy and the alternative group would represent more support.

In some cases, independence can be assessed simply by considering the tree topology. However, this cannot always be accomplished, and in that case information about the developmental and/or genetic basis of the character states is required. For instance, evidence that a single genetic mutation produced all ten

[20] Wiley and Lieberman (2011, p. 153): "The tree with the fewest number of independent origins of shared characters is the preferred solution. This is the maximum parsimony principle."

[21] One area of research in contemporary phylogenetic systematics involves constructing artificial datasets stipulated to result from a particular phylogeny and testing how well various algorithms succeed in recovering that phylogeny. For the phylogeny in question to produce the dataset analyzed, particular assumptions about the evolutionary process must hold. This thus allows systematists to test the manner in which different evolutionary assumptions affect the reliability of different algorithms.

[22] Wiley and Lieberman (2011, p. 196) are explicit that weighting depends on evolutionary assumptions: "Although the investigator may think that no evolutionary assumptions are invoked when he or she decides to treat all characters as equally weighted and unordered, he or she has, in fact, made an explicit evolutionary assumption that each transformation has the same information content."

synapormorphies supporting the first clade would suggest that they can be reduced to a single independent origin and so should count as only a single synapomorphy.

The justification of the use of parsimony analysis in phylogenetic systematics is thus a localized version of a classic thought: we are justified in operating under the assumption that nature is simple only if nature is in fact simple. Only now the goal is not to show that the vast totality of nature is simple through and through, but merely that the evolutionary process is simple. Or, more precisely: the goal is to show that the evolutionary process is simple enough to support favoring the phylogeny that minimizes independent evolutionary origins. Or, still more precisely: the goal is to show that aspects of the evolutionary process are simple enough to support favoring the phylogeny that minimizes independent evolutionary origins, and to include only data reflective of those aspects in our dataset (as well as to appropriately weight those data that are included). It should be clear that this sort of justification affords no comfort to the metaphysician, for she needs simplicity to be truth-conducive in general. The justification of cladistic parsimony, by contrast, shows only that simplicity is (sometimes) truth-conducive in the context of discovering phylogenies. It is difficult to imagine how this particular justification for favoring simpler theories could transfer to metaphysical reasoning.[23]

The upshot is not that Paul and other metaphysicians should not look to the use of IBE in the sciences as part of an attempt to vindicate the use of IBE in metaphysics. The upshot is that, if such an argument is to succeed, one must show not only that IBE is used in some science, but also that the justification for relying on the theoretical virtues invoked transfers between disciplines. The argument that I have offered here does not show that this condition cannot be met, but it does show that surface appeals to theoretical virtues in scientific inference can be misleading. The only way to establish methodological continuity between metaphysics and the sciences is to discover what actually justifies scientific inferences and to show that this justification transfers to metaphysics. That is the task awaiting the metaphysician who wishes to rescue Paul's argument.

8 Conclusion

In biological IBE, theoretical virtues perform a limited amount of evidential work. In metaphysical IBE, the theoretical virtues are asked to perform substantially more evidential work. The use of IBE in metaphysics is thus not vindicated by its use in biology.

[23] Benovsky (2016, pp. 84–87) asks whether the simplicity of nature can justify favoring simpler metaphysical theories and concludes that it cannot, that we cannot generally assume nature to be simple. This is correct, and it is instructive to see that cladistic parsimony is justified only because it relies on an assumption with a narrower scope, about the simplicity of the evolutionary process. Benovsky ultimately claims that "the requirement for parsimony and simplicity comes from *us* rather than from the metaphysical reality." It should be obvious from my arguments that this conclusion holds only if a local justification for favoring simpler theories in metaphysics, analogous to the justification of cladistic parsimony, cannot be found.

The defense provided for the premises of this argument has been programmatic, and the case is not conclusive. It remains possible that the use of IBE outside of biology resembles its use in metaphysics sufficiently to support Paul's argument. Only close examination of the use of IBE in the sciences can support that claim. In particular, it must be established that the theoretical virtues relied upon are the same in metaphysics and the science in question, and it must be shown that the justification that that virtue is truth-conducive in the sciences transfers to metaphysics.

Such concrete examination of scientific reasoning is lacking in Paul's article, which proceeds at a high level of abstraction: science is successful, scientists judge theories by weighing their theoretical virtues, therefore these virtues are truth-conducive in science, therefore these virtues are truth-conducive.

No persuasive argument can be made at this level of abstraction. Any defense of metaphysicians' methodology along the lines offered by Paul must carefully compare the role of the theoretical virtues in scientific and metaphysical reasoning. Perhaps such comparison will reveal scientifically respectable forms of IBE that do vindicate the methodology of metaphysics. It behooves the metaphysician to look for them.

Acknowledgements The author thanks Zina Ward, David Colaço, Liam Kofi Bright, Eden McQueen, Adrian Currie, Jack Samuel and John Norton for comments on drafts and L.A. Paul for discussion.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflict of interest.

References

Arthur, W. (2001). Developmental drive: An important determinant of the direction of phenotypic evolution. *Evolution and Development, 3*(4), 271–278. doi:10.1046/j.1525-142x.2001.003004271.x.

Benovsky, J. (2016). *Meta-metaphysics: On metaphysical equivalence, primitiveness, and theory choice*. Berlin: Springer.

Brössel, P. (2015). On the role of explanatory and systematic power in scientific reasoning. *Synthese, 192*(12), 3877–3913. doi:10.1007/s11229-015-0870-6.

Darwin, C. (1988). *The works of Charles Darwin, volume 20: Variation of animals and plants under domestication, volume iI*. New York, Chesham: NYU Press.

Darwin, C. (2001). *On the origin of species: A facsimile of the first edition. facsimile edition*. Cambridge, Mass: Harvard University Press.

Dawes, G. W. (2013). Belief is not the issue: a defence of inference to the best explanation. *Ratio, 26*(1), 62–78. doi:10.1111/j.1467-9329.2012.00537.x.

Delpino, G. G. F. (1869). On the Darwinian theory of pangenesis. *Scientific Opinion, 2*, 365–367. **(391–393, 407–408)**.

Doppelt, G. D. (2011). From standard scientific realism and structural realism to best current theory realism. *Journal for General Philosophy of Science/Zeitschrift Für Allgemeine Wissenschaftstheorie, 42*(2), 295–316.

Dyke, H., & Maclaurin, J. (2013). What shall we do with analytic metaphysics? A response to McLeod and Parsons. *Australasian Journal of Philosophy, 91*(1), 179–182. doi:10.1080/00048402.2012.762029.

Gayon, J. (1998). *Darwinism's struggle for survival: Heredity and the hypothesis of natural selection*. Cambridge, New York: Cambridge University Press.

Hawley, K. (2006). Science as a guide to metaphysics? *Synthese, 149*(3), 451–470. doi:10.1007/s11229-005-0569-1.

Herschel, J. F. W. (2010). *Preliminary discourse on the study of natural philosophy*. Chicago: University of Chicago Press.

Hodge, M. J. S. (1992). Darwin's argument in the origin. *Philosophy of Science, 59*(3), 461–464.

Jenkin, F. (1867). [Review of] The origin of species. *The North British Review, 46* (June), 277–318. Accessed online: http://darwin-online.org.uk/content/frameset?pageseq=1&itemID=A24&viewtype=text.

Lennox, J. G. (1991). Darwinian thought experiments: A function for just-so stories. In T. Horowitz & G. Massey (Eds.), *Thought experiments in science and philosophy* (pp. 223–246). Savage, MD: Rowman and Littlefield.

Lennox, J. G. (2009). Darwin, Philosopher. *Metascience, 18*(1), 121–124. doi:10.1007/s11016-009-9252-3.

Lewens, T. (2015). Backwards in retrospect. *Philosophical Studies, 172*(3), 813–821. doi:10.1007/s11098-014-0370-4.

Lewes, G. H. (1868). Mr. Darwin's Hypotheses. *Fortnightly Review*, 3 (N.S.), 353-373, 611-628; *Fortnightly Review*, 4 (N.S.), 61-80, 492-509. Accessed online: http://darwin-online.org.uk/content/frameset?pageseq=1&itemID=A604&viewtype=text.

Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London: Routledge.

Lipton, P. (2007). Précis of inference to the best explanation, 2nd edition. *Philosophy and Phenomenological Research, 74*(2), 421–423. doi:10.1111/j.1933-1592.2007.00027.x.

Maclaurin, J., & Dyke, H. (2012). What is analytic metaphysics for? *Australasian Journal of Philosophy, 90*(2), 291–306. doi:10.1080/00048402.2011.587439.

Markosian, N. (2008). Restricted composition. In T. Sider, J. Hawthorne, & D. W. Zimmerman (Eds.), *Contemporary debates in metaphysics* (pp. 341–363). Oxford : Blackwell Publishing.

McLeod, M., & Parsons, J. (2013). Maclaurin and Dyke on analytic metaphysics. *Australasian Journal of Philosophy, 91*(1), 173–178. doi:10.1080/00048402.2012.730534.

Mivart, G. J. (2009). *On the genesis of species* (Reissue ed.). Cambridge: Cambridge University Press.

Norton, J. D. (manuscript). Simplicity as a surrogate. http://www.pitt.edu/~jdnorton/papers/material_theory/6. Simplicity.pdf. Accessed November 8, 2016.

Novick, A., Scholl, R. (manuscript). Presume it not: True causes in the search for the basis of heredity.

Park, S. (2015). Accepting our best scientific theories. *Filosofija Sociologija, 26*(3), 218–227.

Paul, L. A. (2010). A new role for experimental work in metaphysics. *Review of Philosophy and Psychology, 1*(3), 461–476. doi:10.1007/s13164-010-0034-z.

Paul, L. A. (2012). Metaphysics as modeling: The handmaiden's tale. *Philosophical Studies, 160*(1), 1–29. doi:10.1007/s11098-012-9906-7.

Sober, E. (2015). *Ockham's razors: A user's manual*. Cambridge: Cambridge University Press.

Thagard, P. R. (1978). The best explanation: criteria for theory choice. *The Journal of Philosophy, 75*(2), 76–92. doi:10.2307/2025686.

van Fraassen, B. C. (2002). *The empirical stance*. New Haven: Yale University Press.

Wiley, E. O., & Lieberman, B. S. (2011). *Phylogenetics: theory and practice of phylogenetic systematics* (2nd ed.). Hoboken: Wiley-Blackwell.