

Explanation and explanationism in science and metaphysics*

Juha Saatsi†

20th November 2014

School of Philosophy, Religion, and History of Science
University of Leeds
Leeds LS2 9JT, UK
J.T.Saatsi@leeds.ac.uk

1 Introduction

Contemporary metaphysics is heavily concerned with its methodology and epistemic status. This is for a good reason: in the virtual absence of experiments, predictions, and empirical feedback, it is far from clear how metaphysical theories and views can be rationally justified. These ‘meta-level’ issues are notoriously as impenetrable as they are important, alas. We can make the task more manageable by focusing on the role and status of *explanatory* considerations in metaphysics. Explanatory arguments and inference to the best explanation play a central, perhaps foundational role in metaphysics. The ‘explanationist’ methodology, although far from being universally adhered to, is widespread and commonly adopted in one form or another. While the notion of metaphysical explanation remains murky and nebulous, and while some have expressed qualms about inference to the best explanation altogether, explanationism has been bravely defended by others. I will argue that these vindications of explanationism in metaphysics turn out to be superficial and vexed, *especially in their affiliation to explanationism in science*. Even if we grant the legitimacy to explanationism in

*Forthcoming in *Metaphysics and the Philosophy of Science* (eds.) Matthew Slater and Zanja Yudell

†Acknowledgements: I received very helpful feedback from Anjan Chakravartty, Chris Daly, Steven French, Markus Schrenk, Matthew Slater, Laurie Paul, and others.

science and everyday life, it turns out to be surprisingly difficult to defend explanationism in metaphysics.

The advocates of inference to the best explanation—‘the explanationists’—have much to learn from the study of explanation and explanatory inferences in science. Scientists, like metaphysicians, are also in the business of explaining the world, and explanatory considerations arguably also guide scientific inferences (as the advocates of explanationism are quick to point out). Studies of scientific explanation and explanatory inferences in history and philosophy of science offer a grasp on these issues that has a direct bearing on explanationism in ‘naturalistic’ metaphysics that is ‘continuous’ with scientific theorising. The study of explanation and explanatory inferences in science reveals serious shortcomings in prevalent attempts to vindicate explanationism in metaphysics by reference to continuity between science and naturalistic metaphysics.

2 Explanationism in metaphysics

Explanations and explanatory inferences abound in metaphysics; sometimes it is hard to see how metaphysics could even be done if not by comparing potential explanations. I will not question the credentials of all such explanatory reasoning tout court, but rather focus on explanationists who emphasise *the continuity between metaphysics and science*. Here are some exemplars of this more specific target, in alphabetical order.

Armstrong (1983, 1997) famously argues for his account of natural laws via inference to the best explanation. Armstrong also explicitly likens his argument to explanatory inferences to theoretical entities in science. Even induction by reference to laws “becomes a particular case of the inference to explanatory (‘theoretical’) entities”, and “the law, a relation between universals, is a theoretical entity, postulation of which explains the observed phenomena and predicts further observations” (1983, p. 104). Armstrong ultimately says precious little about the nature and justification of this kind of inference in general, only maintaining that the main explanatory advantage of his laws-as-universals view is *unification*, the relevance of which Armstrong supports by reference to the unificationist theories of scientific explanation.

We avail ourselves in [metaphysics] of whatever apparently good canons of explanation we possess or can develop. The basic insight here seems to be involved in the concept of a good explanation: that it should genuinely unify, and that it should be

genuinely informative. (Armstrong 1983, p. 105)

The premiss required here is that, as argued by Michael Friedman [and Philip Kitcher], in the natural sciences what, before anything else, counts as a good explanation is something that unifies the phenomena. (Armstrong 1997, p. 235)

We can continue the list of like-minded explanationists with Bigelow, Colyvan, Dorato, Ellis, and so forth.¹ All these prominent philosophers—mostly from Australasia, as it happens—have advocated a conception of scientific realism according to which inference to the best explanation can support not only realism about electrons and quarks and the like, but also metaphysical views about e.g. mathematical abstracta, possibilia, laws of nature, the nature of properties and probabilities, and mereology. In the case of Ellis (2009), for instance, the explanationist methodology results in ‘metaphysics of scientific realism’ the ontological count of which includes properties, powers, causes, events, propensities, dispositions, spatiotemporal and numerical relations, all of which are argued for via inference to the best explanation—a methodology unifying metaphysics and science.²

Psillos (2005), an explanationist in philosophy of science, sympathises with Ellis’s explanationism by acknowledging that Ellis’s project “rests on the only workable criterion of reality [...]: something is real if its positing plays an indispensable role in the explanation of well-founded phenomena” (398). As Psillos notes, there is a close connection between this ‘explanatory criterion of reality’, also found in Sellars (1963), and the well-known indispensability argument for mathematical platonism. The latter, as championed by its leading contemporary advocates, is again directly associated with explanationism in science.³

Swoyer (2008) presents an argument for ontological commitment to mathematical and other abstracta that is quite unlike the indispensability argu-

¹See e.g. Bigelow (2010), John Bigelow and Robert Pargetter (1990), Colyvan (2006), Dorato (2012), Ellis (2009).

²According to Ellis the test of a good metaphysical hypothesis is twofold: “it must be consistent with the known facts, and be part of a unifying account of reality that explains the overall structure of what we are able to observe” (2009, p. 123).

³See, for example, Colyvan (2006, p. 229):

“I will take the indispensability argument to be an argument that puts pressure on the marriage of scientific realism and nominalism. It does this because the style of argument [viz. IBE] is one which scientific realists already endorse.”

ment. But it is also a forthright instance of explanationism, with a clear emphasis on methodological similarity with science.

The chief philosophical benefit claimed for ... abstract entities is that they ... help explain otherwise puzzling phenomena ... [IBE] is not some arcane concoction of meta-physicians. We often infer that something exists on the grounds that its existence would explain something that would otherwise be puzzling... Such inferences also seem common in science... (Swoyer 2008, p. 16)

Swoyer reflects on the methodology of metaphysics, and for him explanationism represents the answer to 'how ontology might be possible' (cf. Swoyer 1999). The idea that explanationism and its methodological affinity with science provides an answer to methodological scruples about metaphysics has also been recently defended by Paul (2012), amongst others.⁴

Explanationism is popular in meta-ethics as well. Here an explicit connection to explanationism in science is drawn by Sturgeon (2006), for example.⁵ Examples proliferate; clearly explanationism, with a nod to science, flourishes in metaphysics and beyond.⁶ Every philosopher above has their

⁴See Paul (2012, p. 22):

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 in metaphysical theorizing.

⁵See Sturgeon (2006, p. 243):

The justification of many of our beliefs, not just perceptual beliefs, derives from their explanatory role. Thus, for example, the justification of many scientific beliefs is said to lie in their contributing to good explanations of observed evidence ... and beliefs about others psychological states may explain what we observe of their behaviour. ... [A]lmost everyone agrees that explanatory coherence ... is an epistemic virtue in a set of beliefs, contributing other things equal to its being well justified. And this is thought to be especially so when there is explanatory integration across significantly different categories: beliefs about the past explaining evidence in the present, or beliefs about unobservables explaining what we observe, for example. Now, if evaluative beliefs play a plausible role in explaining other facts we have reason to believe in, facts that are not themselves evaluative, then they will be candidates for justification of a similar sort, by their integrative explanatory role.

⁶In the sphere of epistemology, Pargetter (1984) advocates explanationism as a response to skepticisms about other minds by viewing inferences to other minds as being relevantly similar to "arguments to the best explanation in science."⁷ McLaughlin (2010) makes

critics, of course, but no wholesale assessment has been given of their common methodological denominator: a widespread justificatory appeal to a similarity between metaphysical theorising and the explanatory practices of science. There are obvious questions to be asked here: What kind of similarity is this? What kind of justification does it support? What are the limits of such justification of metaphysical views?

I will next characterise in more detail two ways in which philosophers have attempted to capitalise on the similarity between *naturalistic* metaphysics and science, setting the scene for a more critical discussion to take place in Sections 4 and 5.

3 Naturalistic metaphysics and science

Naturalistic metaphysicians theorise about the natural world—a world also studied by science—by conceiving, assessing and supporting various theories about it. Many of these theories are admittedly quite speculative. I am interested in the brand of naturalistic metaphysics that recommends, as a methodological dictum, turning to science to see how theories are best assessed and supported. Many naturalists have happily followed Quine in accepting the following three broad methodological judgements. First, in science we find a broad methodological injunction to let theoretical virtues guide theorizing, at least regarding matters that are unobservable to our ‘naked senses’.⁸ Secondly, in science we can witness (a degree of) confirmation holism: a theory can be related to evidence only as a part of a wider web of theoretical beliefs, not in and of itself. And thirdly, science (suitably interpreted) has all the answers to questions of ontology: there is no ‘first philosophy’ concerning questions of existence. These three naturalist tenets lead to two broad strategies to support explanationism in metaphysics.

The first strategy capitalises on *methodological similarity* between metaphysics and science in the spirit of the first tenet above. Consider how this tenet is exemplified in the apologia for metaphysics by Sider, Hawthorne, and Zimmermann (2007).

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

comparable use of inference to the best explanation in philosophy of mind, and Biggs (2011) runs a similar line of thought in relation to modal epistemology.

⁸For Quine the same theoretical virtues are also important for justifying our views concerning common sense objects.

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. (p. 7)

Sider et al. stress the idea that in metaphysics, just as in physics, theory choice is governed by theoretical virtues. In effect, they hold that metaphysics is but a *degree* apart from the more speculative reaches of empirical science, and regardless of its highly speculative character the methodological similarity with science nevertheless justifies the rationality and meaningfulness of metaphysics as a theoretical endeavour. This is a very broad idea, of course, and there are many ways to make it more precise. In particular, one can refine it in explanationist terms, leading to a vindication of explanationism in metaphysics. Thus, Swoyer (in Sider et al., 2008), for example, appeals to inference to the best explanation in framing a methodological similarity between his arguments for Platonism, on the one hand, and science on the other.

My suggestion is that we should (re)construe arguments for the existence of abstract entities as inferences to the best over-all available ontological explanation. [...] Inference to the best explanation plays a central role in daily life and, according to many philosophers, in science. [...] [M]any maintain, without inferences to the best explanation science, and much of ordinary life, would be impossible. (p. 15)

This is a good example of the strategy that aims to justify explanationism in metaphysics by reference to a methodological uniformity. (See also Paul, 2012.)⁹

⁹Note that these philosophers are patently not claiming that metaphysical theories are *equally well* supported by evidence as our scientific theories are. Rather, the claim is that the *kind* of justification is arguably similar to theoretical domains of science. Metaphysics is undeniably much more speculative than science, and one should not object to explanationism in metaphysics—as e.g. Ladyman (2007, 2012) partly does—on the grounds that it does not have a *probative force* on a par with the explanatory considerations arguably

As an aside, it is worth noting that a similar line of thought has been popular in the scientific realism debate in philosophy of science. Here naturalism has been taken to imply that an explanationist argument for scientific realism should exemplify the same method as scientists themselves employ: the realist’s inference to the best explanation—empirical success is best explained by approximate truth—should be viewed as a further application of *scientific* explanationism.¹⁰ The realist inference thus arguably exemplifies the same method as various explanatory inferences in science—a fact that arguably provides support for scientific realism via a rule-circular justification of the realist’s inference. (Cf. Psillos, 1999, ch. 4; 2011).¹¹ This, too, is an example of the same strategy for justifying explanationism outside science.

Let’s now move on to the second strategy that capitalises on *confirmational holism*. This flows out of the three tenets of naturalism as follows. In answering questions of ontology naturalists turn to our best scientific theories, and confirmation holism recommends belief in *all* theoretical assumptions that are responsible for the (predictive, explanatory, whatever) successes provide realism-eliciting evidence for a given theory. Our grasp of theoretical virtues further shapes our understanding of those successes. This gives rise to a Quinean ‘indispensability argument’: we ought to have commitment to whatever indispensably contributes to our best theories being the best. This then leads to a justification of metaphysical views when it is argued that *assuming* realism about paradigmatic scientific posits (electrons, quarks, etc.) we should *also* be committed to paradigmatically philosophical

operative in science. For this is just to object to metaphysics’ speculative nature, not to explanationism in metaphysics per se.

¹⁰ As Putman puts it, scientific realism should be “viewed . . . as part of the only scientific explanation of the success of science” (1975, p. 73). Similarly, Psillos maintains that “naturalistic philosophers of science “should employ no methods other than those used by the scientists themselves” (Psillos 1999, p. 78), and Boyd argues that “the epistemology of empirical science is an empirical science.” (Boyd, 1989, p. 13).

¹¹ Note that here, too, a naturalistic philosopher is not committed to claiming that her philosophical theory (about science) can be supported to the same degree that scientific theories themselves are supported. One should not object to explanationism in this context merely on the grounds that it does not have probative force on a par with the explanatory inferences in science. This mistake is made by Frost-Arnold (2010), who argues that the realist explanation (and the related inference to the best explanation) fails to satisfy scientific demands for a good explanation (and good explanatory inference), and therefore fails the tenets of naturalism. This demand is based on too strict a conception of naturalism. The realist explanation can be purely philosophical in the sense of not enjoying the degree of *evidence* enjoyed by paradigmatically good scientific explanations. The mode of inference can be the same in the two cases, while the overall evidence (or ‘epistemological standard’) is not.

assumptions (e.g. existence of abstracta, the presentist view of time) on the basis of their allegedly similar role in contributing to our best theories.

This broad idea of confirmational holism can be further refined in different ways. In particular, the notion of indispensability at stake is sometimes framed in explanationist terms, leading to a vindication of explanationism with respect to metaphysical claims. For example, the advocates of the so-called ‘explanatory indispensability argument’ have recently argued that metaphysical posits (e.g. mathematical and other abstracta) can indispensably contribute to some of our best scientific explanations in a way that is ontologically committing (by scientific realist lights).¹² Such arguments aim to establish that paradigmatically metaphysical views (e.g. platonism) can enjoy a degree of *empirical confirmation* by virtue of the relevant assumptions (e.g. regarding numbers) playing an appropriate explanatory role—a role that is appropriately continuous with those played by paradigmatically scientific assumptions.

* * *

The scene is now set for a more critical discussion. I will not be taking issue with naturalism or the broad idea that science and metaphysics can be viewed ‘of a piece’. Rather, through a closer reflection on explanationism in science I will only take issue with the explanationist spin that has been put on these two strategies for vindicating naturalistic metaphysics. An incentive for this springs from within the naturalistic conception of philosophy. As said, naturalism recommends turning to science to see how theories are best assessed and supported. But exactly *which* scientific disciplines should we turn to in forming a view of the successful scientific methodology and its limits? I think we should follow Quine himself and understand ‘science’ quite broadly, so that all scientific findings concerning science and scientists should be taken on board, including those from the history of science and the relevant areas of psychology.¹³ After all, there’s every reason to expect

¹²See, for example, Colyvan (2006, p. 229):

[IBE] is a special case of the indispensability argument. [T]his is a style of argument that the scientific realist accepts. Mathematical entities surely feature prominently in various explanations.

See also Baker (2009) and Psillos (2011) for arguments in the same spirit.

¹³See e.g. Quine (1995, p. 49):

I use science broadly, including not only the ‘hard sciences’ but also ‘softer sciences’, from psychology and economics through sociology to history.

that these areas of enquiry have potential to inform us of our capacity and reliability in explanatory reasoning. With this in mind, let us now have a closer look at the two strategies for vindicating explanationism, in the light of the study of explanation and explanatory inferences in psychology and history and philosophy of science.

4 On the strategy of methodological similarity

I will now critically examine arguments that defend explanationism in metaphysics by reference to methodological similarity with science (e.g. Swyer 1999, 2008, Paul 2012). I will put aside various differences in details and focus on a common gambit of methodological unity, which can be expressed in general terms as follows.¹⁴

The gambit begins with the premise that explanationism is truth-conducive in (some relevant area of) science, and hence justified in that context. It is then argued that (a given area of) metaphysics is methodologically unified with science in the sense that:

MC1 Both metaphysics and science employ inference to the best explanation.

MC2 We have no reason to think that if explanationism is truth-conducive in science, it is not so in metaphysics.

MC3 We have a positive reason to think that if explanationism is truth-conducive in science, it is also so in metaphysics.

On the basis of these assumptions the gambit concludes that explanationism in (the given area of) metaphysics is truth-conducive, and hence rational and justified also in that context. (Note that MC2 does not entail MC3: even if we cannot see why explanationism would fail to be truth-conducive only in metaphysics, we may not be able to positively argue for its truth-conduciveness either. Both MC2 and MC3 are required to tackle a sceptic who demands a positive argument for the use of IBE in metaphysics.)

As far as I'm concerned, the conclusion follows if MC1–MC3 can be established. But the problem is that MC2 and MC3 have not been established, and there is reason to think that they cannot be established.¹⁵ It is fairly

¹⁴Typically explanationists only offer vague *tu quoque* references to science, scientific realism, and inference to the best explanation therein. I construe them as *aspiring* to the schema presented here.

¹⁵There is reason to be sceptical about MC1 as well. Cf. e.g. Ladyman (2012).

obvious that without support for MC2 and MC3 the gambit reduces to a more or less trivial recognition that theory-choice in different contexts and disciplines can be *described* in similar explanationist terms at some level of abstraction. It is clear that this kind of purely descriptive continuity does not in itself carry any justificatory weight. It is comparable to a folly attempt to justify any old enumerative induction merely on the basis of it being of the same form as some licit enumerative inductions. (Cf. Norton 2003)

Before we analyse the problems with MC2 and MC3 in detail, it is worth noting an ill-considered worry about the gambit. It has to do with an appropriate reading of ‘truth-conducive’ and ‘justification’ in the schema above. One might think that the gambit is problematic if we have reason think that inferences to the best explanation are *much less reliable* in metaphysics than they are in science, or even if we lack a positive reason for thinking that they are equally reliable in both. For example, one might point to differences in the disciplines’ track records—physics vs. metaphysics, say—or argue that our best reasons for thinking that scientific IBEs are reliable do not carry over to metaphysics, raising the spectre of potential unreliability (Cf. Ladyman, 2012). And one might think, in particular, that any such worry about the *relative reliability* in different domains is ipso facto a worry about either MC2 or MC3.

Although there is something to this worry—one cannot wholly divorce a method’s reliability from its justification—it is quite difficult to square it with the fact that metaphysics is admittedly inherently *speculative* in a way that our best science arguably is not. I take it that theorising in metaphysics is generally not taken to be progressive *in the way science is*. For the realist at least, science is systematically latching better and better onto unobservable reality, and scientific theorising is guided or constrained by ‘correspondence principles’ that are grounded in the ideal of discernible continuity in theoretical development. Whatever progress metaphysical theorising makes, it appears to be compatible with the likelihood of much more significant theoretical discontinuities. Accordingly, our degree of confidence is admittedly significantly lower to any particular output of explanatory reasoning in metaphysics. But this need not mean that in metaphysics explanatory loveliness is not in any sense functioning reliably as a guide to inductive likeliness.¹⁶ For instance, it could be that in metaphysics we are much less able to think of a pool of potential alternative explanations from

¹⁶See Lipton (1999) for an explication of the explanationist slogan that ‘explanatory loveliness guides inductive likeliness’.

which we choose the ‘best’, so that we often end up debating over the ‘best of a bad lot’.¹⁷ Still, it could be that inference to the best explanation is quite reliable in picking out the right theory *assuming* that it happens to be included in the pool. In this way the method could be reliable for example ‘in the long run’ (assuming that we eventually manage to conceive of the right theories) and also rational, despite not being reliable in the sense of engendering high degree of confidence in any particular explanatory inference.

Hence, some difference in the method’s relative reliability is compatible with the continuity gambit, because such a difference need not affect the method’s justification, which in inherently speculative disciplines such as metaphysics could be based on the notion of reliability ‘in the long run’, or reliability conditional on not having a ‘bad lot’, or reliability in mere relative ranking of alternatives. Unfortunately we have little reason to think that explanationism is reliable even in such a qualified way. As I will next explain, there are other differences between science and metaphysics that give us reason to think that the (assumed) justification of explanationism in science does *not* carry over to metaphysics.

4.1 Issues with MC2

The vindications of explanationism in metaphysics offer some limited support for MC2—the claim that we have no reason to think that if explanationism is truth-conducive in science, it is not so in metaphysics. This support springs from the somewhat vague notion that naturalistic metaphysics and theoretical science are *of a piece*: metaphysics, like theoretical science, is just further (albeit more abstract or general) theorising about the unobservable world. This naturalist notion that science and metaphysics, as intellectual and theoretical endeavours, are at bottom in the same epistemological boat offers support for MC2 in so far as we have reason to think that there are no differences between these endeavours that would be relevant for the respective reliability of inference to the best explanation.

We *can*, however, point to various sui generis features of science and its explanationist method, and to differences between science and metaphysics, to suggest that MC2 is false: we do have reasons to think that explanation-

¹⁷The ‘bad lot’ objection against explanationism is due to van Fraassen (1989), who employed it against the idea that it is rational to believe that the best explanation is more likely to be (approximately) true than not. The objection loses bite if explanationism is construed more cautiously, pertaining only to a hypothesis’s epistemic probability *relative* to its rival hypotheses. (Cf. Okasha 2000)

ism is not suitably truth-conducive in metaphysics, even if it is (in a sense) truth-conducive in science. For example, we can begin by noting that the truth-conduciveness of explanationism in science can be (and arguably has to be) *qualified* in ways that are not applicable to metaphysics, thus spelling trouble to MC2. The history of science strongly indicates that explanatory inferences in science have (at best) only been reliable in quite limited respects: there's a long and familiar tradition in the history and philosophy of science to point the various gross falsehoods of the best science of yesteryear.¹⁸ Past science has seemingly not been reliable in any significant way regarding its *fundamental ontological posits*, for instance.¹⁹ How should one respond to the challenge this presents to the explanationist? One response is to give up on IBE altogether as a guide to deep unobservable features of the world. According to this response MC2 is irrelevant (even if true) given the failure of its antecedent. This response is clearly quite uninteresting in a current dialectical position, and according to many scientific realists it furthermore throws out the IBE-baby with the bathwater, since there are ways of qualifying the sense in which IBE reliably functions in science.

The situation faced by the advocates of scientific explanationism is (*prima facie*) somewhat paradoxical: again and again in the history of science it is precisely the metaphysical and ontological assumptions concerning the *nature* of gravity and other forces, light, disease, life, genes, etc.—the assumptions that were underwriting the best scientific understanding of the relevant phenomena—that have subsequently turned out to be false.²⁰ The evidence for this claim can be found in the extensive literature on scientific realism (from Laudan (1981) onwards).²¹ Arguably the best realist response

¹⁸See e.g. Laudan 1981, Stanford 2006, Lyons 2006, Vickers 2013.

¹⁹Famous now-rejected ontological posits include: gravitational force, caloric, phlogiston, lumiferous ether, electromagnetic ether, circular inertia, miasma, vortices, vital forces, electron orbits, to name a few. To be fair, no explanationist has alleged that inference to the best explanation is a sure-fire method, even in science, but stressing the method's *defeasibility* is not a satisfactory response to the worry that the method has been demonstrably unreliable with respect to the reality of fundamental scientific posits.

²⁰See Doppelt (2005, 2007) and Saatsi (2012) on the internal tensions that this 'paradoxical' situation can present the realist with.

²¹Consider the optical ether, for just one paradigmatic example. It was the assumed existence of the ether that gave the best and only unified understanding of light phenomena for much of the 19th century. For a nice discussion, see Cordero (2011), who also catalogues famous testimonies of the resulting confidence in the existence of the ether. These include Maxwell:

Whatever difficulties we may have in forming a consistent idea of the constitution of the ether, there can be no doubt that the interplanetary and interstellar spaces are not empty, but are occupied by a material substance

is to accept that the explanationist methodology in science is only reliable in rather limited ways: even if IBE cannot reliably function as a guide to the fundamental nature of things (as the history suggests), it is open for the realist to maintain that scientists, with their IBE-laden methods, nevertheless systematically ‘latch onto’ reality with their theories in ways that largely account for the predictive and instrumental successes of science. So goes a promising realist response to the historical ‘pessimistic induction’ over now-rejected theoretical posits and explanatory mechanisms in ‘fundamental science’.²²

The upshot is that even in the face of the pessimistic historical record there is still room to consider explanationism in science as ‘truth-conducive’, but only in a rather qualified sense. In the light of the history of science it is difficult to support MC2 through the notion that metaphysics and science are methodologically of a piece: the idea that IBE is reliable in science in the way that explanationists would like it to be reliable in metaphysics is arguably premised on a naive view of (the history of) science.²³ Arguably any scientific realism that is sophisticated enough to be tenable in the face of the history of science is highly ‘selective’ regarding our epistemic commitments towards the outputs of the explanationist method, which is only reliable in limited respects. It is a matter of debate how to best characterise the realist commitments and the method’s curtailed reliability. While the currently favoured realist positions vary significantly in detail, they typically incorporate the broad idea that a (radically) false theory can be empirically successful by virtue of *selectively latching onto* some appropriate features of the world. So, even if scientists’ explanatory inferences are not tracking reality in relation to fundamental ontology, a selective realist can save the realist day by claiming that such inferences are nevertheless reliable in ways that fuel scientific progress: they tend to latch onto reality with respect to those aspects of reality that generate the progress and empirical success of

or body, which is certainly the largest, and probably the most uniform body of which we have any knowledge. (Encyclopedia Britannica, 9th ed., vol. 8 [1893], 572; this article first appeared in 1878.)

and Lord Kelvin (previously Thomson):

You can imagine particles of something, the thing whose motion constitutes light. This thing we call the luminiferous ether. That is the only substance we are confident of in dynamics. One thing we are sure of, and that is the reality and substantiality of the luminiferous ether. (Thomson 1889, 310; quoted in Thompson 1910, 1035)

²²See e.g. Harker (2013), Saatsi (forthcoming *b*), and references therein.

²³This point is also made forcefully by Ladyman (2012).

science as measured e.g. in terms increasing empirical adequacy and novel predictions. The debate is subtle and some care is required lest realism collapses into empiricism, but the overall issue with MC2 is clear: the historical track-record impels the realist to adopt thus qualified notion of reliability, but it is rather questionable whether it makes any sense to apply such qualifications to explanatory inferences in metaphysics (since the relevant notion of empirical progress seems inapplicable to metaphysics).

The historical track-record is not the only reason to worry about MC2. There is a challenge to be faced by naturalistically inclined advocates of explanationism in metaphysics and science alike: to account for the truth-conduciveness of inference to the best explanation in a way that coheres sufficiently well with the rest of our naturalistic world view, according to which our capacities and ways of reasoning about the world are products of natural and cultural evolution, and every reliable mechanism of theorising is reliable for a natural reason amenable to a scientific study. I am *not* suggesting that the onus is on the explanationist to actually provide such a naturalistically acceptable account of the workings of explanatory reasoning. But the challenge nevertheless constrains plausible forms of explanationism, because arguably we should at least be able to *conceive of* a potential account that is naturalistically acceptable. Inability to come up with a how-possible account of the truth-conduciveness of explanatory reasoning reflects badly on explanationism, which starts looking suspiciously mysterious. Here we find another difference between science and metaphysics that gives us a reason to worry about explanationism specifically in metaphysics.

To see the difference, consider first the challenge in relation to science. The challenge is to show that it is not inconceivable how we have developed a capacity for explanatory reasoning that tends to latch onto unobservable reality in ways responsible for the empirical success of science. These ways of thus ‘latching onto reality’ come in many forms. In some cases it is easy to conceive of a naturalistically acceptable story, at least in outline. Consider, for example, the explanatory reasoning that leads to knowledge of the sorts of unobservable *causal relations* that underlie sophisticated scientific interventions exemplified by gene manipulation (see e.g. Woodward 2010). Discovering such causal connections through explanatory reasoning need not be mysterious because the relevant mode of explanatory practice—causal-mechanistic explanation—is also exemplified in everyday cognitive situations that we are naturally equipped to tackle. According to the new ‘mechanistic paradigm’ in the philosophy of science much of scientific explanation and associated explanatory reasoning, especially in the life sciences, concern such mechanisms that are not too far-removed from commonplace mechanisms

such as engines and clockworks (see e.g. Bechtel and Abrahamsen 2005). Such causal-mechanistic mode of explanation is, of course, inapplicable to philosophical theorising.

Other modes of scientific explanation and explanatory reasoning call for a different story, but in general there is something specific to science that helps us to conceive how ‘reality tracking’ practice of explanatory inferences could evolve even in much more abstract contexts, such as theoretical physics or mathematical biology. In particular, we can appeal to the myriad ways in which *experimental feedback* in science can conceivably shape scientists’ domain-dependent judgments of explanatory goodness so as to result in a truth-tracking explanationist methodology (e.g. Kuipers 2002). This is compatible with the kind of explanatory pluralism in science that has been recently emphasised by many (e.g. Douglas 2009), as one would indeed expect a degree of divergence in a detailed evolution of explanationism. It is clearly highly questionable whether these conceivable accounts of the (qualified) reliability of explanationism in science can in any way be extended to inferences to the best explanation in metaphysics, where experimental feedback plays no such role in guiding the explanatory practice.

All this points to a challenge to the metaphysical explanationist that is comparable to the one that Field (1989) (following Benacerraf) has famously posed to the mathematical Platonist. All things considered, it appears much more difficult to conceive of a naturalistically acceptable account of the truth-conduciveness of explanationism in metaphysics. The onus is on the explanationist to show that we can conceive of an account of our reliability in explanatory inferences concerning, say, matters of fundamental ontology or modality. Broad references to explanationism in science pay a mere lip service to this challenge given the disanalogies noted above.

As an aside it is worth noting that one of the most prominent scientific realist argument is also equally targeted by the above criticism. The global realist meta-abduction argument (‘no-miracles argument’), as developed by Psillos (1999) and others, is a profoundly problematic exercise in philosophical explanationism. At stake is a ‘global’ reliability of the abductive method of science, which is supported via an inference to the best explanation, a philosophical inference that is continuous with explanationism in science.²⁴ In my view the global no-miracles argument fails to bridge the gap between scientific and philosophical explanationism. The fact that the explanandum at stake is *about* science does nothing to diminish the above worries about

²⁴Cf. Section 3 and note 11.

the continuity gambit exhibited also by this realist argument.²⁵

So far I have surveyed reasons for thinking that explanationism, even if truth-conducive (in a qualified sense) in science, is not so in metaphysics. Next I will turn a critical eye to the positive arguments for thinking that the presumed truth-conduciveness of explanationism in science carries over to metaphysics.

4.2 Issues with MC3

What works for the goose, works for the gander. Or so the advocates of explanationism think: if explanationism is truth-conducive in science, it is so also in metaphysics. Positive arguments to this effect typically point to underlying similarities between scientific and metaphysical explanations and the respective explanatory virtues. But what little has been said of the pertinent similarities turns out to be problematic in the light of current analyses of scientific explanation (in as far as those analyses are not applicable to explanations in metaphysics, in particular).

Swoyer (2008) typifies the slender attempts to support MC3. After noting an obvious disanalogy between scientific and philosophical explanations—only the former typically involve causal mechanisms—Swoyer points to the relevant analogies:

There are a number of [common] explanatory virtues addition to pinpointing causal mechanisms, however, including unification, integration, and redescribing phenomena in a theoretically enlightening way. Seeing a common pattern at work behind seemingly disparate phenomena certainly yields one sort of understanding. By describing the moon, the earth, and projectiles on the surface of the earth as bodies with inertial and gravitational mass, Newton was able to provide a unified account of their motions. Something similar can occur in philosophy. (p. 27)

Swoyer is not alone in pointing to *unification* and explanations' *power to give understanding* as relevant features of good explanations, shared by metaphysical and scientific explanations alike. Typically these explanatory virtues are characterised in extremely general and abstract terms. But

²⁵Since I'm not concerned with defending scientific realism here, I refrain from exploring better vindications of scientific explanationism. See Saatsi (2009) for further discussion and a recommendation that a realist justification of scientific explanationism should proceed 'piecemeal' (as opposed to 'wholesale', as in the case of no-miracles argument), relying more heavily on the role of experimental feedback that shapes our explanatory judgements.

on closer inspection we will see that broad references to these features pay a mere lip service to the literature on scientific explanation, which provides plenty of reason to think that no substantive overarching account of such measures of explanatory goodness can be given to unify metaphysical and scientific theorising as desired.

Understanding. Let's focus on explanations' 'power to give understanding' to begin with. (I will discuss unification separately below.) Is it safe to assume that theories' power to give understanding is a truth-conducive epistemic virtue that can guide theory choice, relating epistemic agents to the unobservable reality regardless of our domain of enquiry, so as to support MC3?²⁶ Such an assumption is often tacitly made, without evidence, by the explanationists.

We should worry about the validity of this tacit assumption. Without clear understanding of 'understanding' (especially in philosophy), one may be inclined to regard 'power to yield understanding' as too subjective and psychological a notion to support a wholesale justification of explanationism (unspecific to a domain of enquiry). This initial worry is much strengthened by the fact that science itself turns out to exhibit a huge variation in its theories' claimed 'intelligibility' across different historical and social contexts and across different domains of theorising (e.g. Kuhn 1977, de Regt & Dieks 2005). Our knowledge of the scientific endeavour indicates that scientists' evaluation of explanations in terms of their provision of understanding can be heavily influenced by various kinds of social and cultural factors; it is not simply a matter of determining some objective feature of explanations.

For example, some excellent scientific theories in the history of science have lacked the power to provide understanding to leading contemporary scientists. For a well-known example, Huygens, Leibniz and much of the scientific community regarded Newton's universal gravitation as involving an unintelligible action at a distance—a sentiment echoed by Newton himself (at times). By comparison, an average physics student today regards Newtonian gravity as a perfectly cogent explanatory posit. Sometimes the intelligibility of a new theory eludes everyone apart from its architects, as when the physicists in the Copenhagen-Göttingen circle around 1926 claimed to understand quantum phenomena with matrix mechanics, while most other physicists outside the circle regarded it unintelligible (de Regt,

²⁶By 'domain of enquiry' I denote not only the subject matter of theoretical enquiry, but it in combination with the myriad psychological and social factors that shape the study of that subject matter. Thus, ancient Chinese astronomy constitutes a different domain of enquiry from 20th c. western astronomy, for example.

2001). There are also clear disagreements across different scientific fields regarding the level of understanding provided by different theories and different theoretical approaches. For instance, it's not uncommon for a group of mathematical biologists to claim that an abstract mathematical theory (e.g. of biological scaling laws) gives them deep understanding of biological phenomena, only to be faced by severe dissent from other camps of biology (see e.g. Keller 2002).

This kind of demonstrable variability in scientific explanations' intelligibility and power to give understanding speaks against the tacit assumption made in support of MC3. In science it can be impossible to objectively rank theories with respect to the level of understanding they provide, given the way in which 'understanding' is interwoven with theorists' social, cultural, and philosophical background.²⁷ Why would 'understanding' be any more objective affair in metaphysics?

To be clear, those who defend explanationism in science must of course also accommodate these data. But the realist arguments to this end can at least appeal to specific features of science, such as experimental feedback and an objective sense of scientific progress. Such arguments, being specific to science, are of no help to an advocate of MC3 however.

The historical and social studies of science cannot be ignored by a naturalistic philosopher who (in the Quinean spirit) views philosophy as continuous and in harmony with our *overarching* vision of science and scientists. Psychology and cognitive science form an ineliminable part of that vision as well. Some philosophers of science have recently turned to these disciplines to explore the relevance of *scientific* studies of explanation and understanding, resulting in increasingly naturalistic perspectives on these topics, yielding hitherto unexplored challenges to explanationism in metaphysics.

Studies of the psychological state of understandings have revealed serious issues with invoking the *sense of understanding* offered by an explanation as

²⁷Such objective ranking eludes us, at least in so far as 'the level of understanding' provided is meant to be meta-cognitively accessible to theorists in question. In the context of specific accounts of scientific explanation we can try to put 'understanding' on a more objective footing by construing 'genuine understanding' (as opposed to mere 'sense of understanding') as grasping of a genuine explanation. Such 'genuine understanding' may bear no relation to agents' own assessment of their understanding, however. If an account of explanation offers a way to measure 'explanatory depth' it can be furthermore associated with the level of genuine understanding. For example, in relation to Woodward's account of scientific explanation, understanding has been construed by some as the ability to draw correct counterfactual inferences, rendering 'the level of understanding' proportional the measure of inferences one can make. (Ylikoski and Kuorikoski, 2010) It is difficult to apply this to philosophical explanations, however.

a reason for accepting it. In his review of these issues, Trout (2007) argues that “the research findings on judgment and decision-making [indicate that] the sense of understanding is a common, but routinely unreliable, index of intellectual achievement.” (p. 566) According to Trout this is not just a matter of reduced reliability; rather, “there is little practical wisdom in relying on the sense of understanding as a cue of a good explanation, or even a potential one.” (p. 574) Trout argues for this by pointing to two well-documented psychological biases (hindsight and overconfidence), as well as other robust psychological phenomena.²⁸

Trout primarily aims his challenge at scientists and scientific realists, but it is obvious that the challenge applies to explanationism at large.²⁹ If the sense of understanding evoked by an explanation is influenced by psychological biases, and we are operating in a domain which does not return any systematic feedback on our explanatory judgments, we ought to worry about the grounds of basing an inference on understanding. Furthermore, as Trout goes on to discuss in a more positive note, there *are* some responses available to a scientific realist (qua explanationist), hinging on the role of explanatory considerations in the empirical progress of science.³⁰ But these responses are difficult to extend to explanationism in metaphysics, where the critical notion of empirical progress—involving predictive and instrumental successes—is simply inapplicable.

In a spirit similar to that of Trout, Lombrozo (2011) reviews the growing literature within psychology and cognitive science that attests to the *instrumental value* of explanations for everyday cognition, with lessons that arguably generalise to science. According to the research reviewed by Lombrozo, explanation can have instrumental value in everyday cognitive tasks of generating predictively useful theories and promoting the discovery of causal structures, for example, by virtue of its intimate cognitive relationship to

²⁸See also Grimm (2009) for a more optimistic vision on the reliability of understanding.

²⁹In some domains scientists have grown acutely aware of these issues (e.g. the problem of just-so-stories in evolutionary biology), in other domains less so (e.g. the problem with the so-called ‘rational addiction theories’ in economics, cf. Rozeberg 2004).

³⁰See Trout (2007, p. 586)

Scientific realists can assign a robust role to objective factors in explanation—such as statistical and causal relevance—and value the contribution of explanation to scientific progress. If your focus is not balanced by a positive account of the sense of understanding in a scientific theory of explanation, then it is easy to portray any criticism of the sense of understanding as the first step toward explanatory nihilism. But explanatory nihilism is surely premature.

learning and causal inferences. According to the picture painted by this research, the cognitive mechanisms shaped for *our* everyday ‘theorising’ and learning naturally carry over to *our* practice of science (where these mechanisms get refined by empirical feedback—e.g. confirmed predictions—that can contextually shape our grasp of explanatory virtues, for example). These cognitive mechanisms also naturally get applied to metaphysics, no doubt. But while this line of thought can to some extent rationalise the explanatory endeavours of science, the talk of the ‘instrumental value of explanation’ in metaphysics is tenuous at best, and in the absence of empirical feedback to guide our explanatory practices it is much more difficult to justify the truth-conduciveness of those practices, especially when they are far-removed from their cognitive origins. (This issue connects to the challenge discussed towards the end of §4.1.)

Unification. Unification is one of the most often cited (alleged) good-making features of both scientific and philosophical explanations.³¹ The advocates of explanationism also often regard unification (*ceteris paribus*) as a universally recognisable objective explanatory virtue of a theory, or even as a constitutive feature of explanation as Friedman and Kitcher have suggested.³² Thus, the thought goes, we can arguably assess philosophical theories with some reliability by virtue of being able to gauge relative degrees of unification.

However, in the light of contemporary philosophy of science the relationship between unification and explanation is much more delicate in ways that threaten to undermine broad appeal to unification in support of MC3. First of all, the attempts (by Friedman, Kitcher) to view unification as a *constitutive* feature of explanation have faced severe criticisms and are highly contentious.³³ And although there is an important seed of truth in the weaker idea that unification is an explanatory *virtue*, the way this idea is developed in different contemporary accounts of explanation offers no solace for MC3.

For example, a well-received recent counterfactual theory of explanation offers a promising way to construe unification as an explanatory virtue. In this account unification comes out, first of all, as one of *several competing*

³¹See Swoyer (1999, 2008), Armstrong 1997, pp. 235–6), and Sider (2007), for example. Some regard the argument for genuine modal realism in Lewis (1986) also critically hinging on unification in a similar spirit. Ditto Sider (2011).

³²Both Friedman and Kitcher also associate unification closely with understanding. Cf. Friedman (1974, p. 15) and Kitcher (1989, p. 432).

³³See e.g. Gijssbers 2007, Barnes 1992, Woodward 2003 (ch. 8).

virtues, and secondly, as a good-making feature of *individual explanations*, not of scientific *theories* per se. Let me elaborate. We can begin, with Hitchcock and Woodward (2003), by recognising that explanatory power is naturally associated with the *generality* of an explanatory regularity employed. The basic intuition harks back to the covering-law model of explanation according to which an explanandum is explained by subsuming it under law-like explanatory regularities. The more these laws cover—the more general they are—the deeper the explanation, according to the intuition. Unificationist accounts of explanation took this idea as their starting point, but arguably misidentified the sense of generality at stake. As Hitchcock and Woodward put it:

[T]raditional approaches to explanation have ... focused on the wrong sort of generality: generality with respect to objects or systems other than the one that is the focus of explanation. The right sort of generality is rather generality with respect to other possible properties of the very object or system that is the focus of explanation. (p. 182)

In their account explanation is a matter of exhibiting a pattern of counterfactual dependence that describes how the explanandum would change under various conditions. Explanations provide such modal information in terms of generalisations that are *invariant* under appropriate changes in some variable values (indexing determinate properties).³⁴ It is a major attraction of the account that it provides the means to capture explanatory depth in general terms, in terms of the *range of invariance* of a generalisation (cf. Hitchcock and Woodward 2003). Unification, furthermore, comes out as an explanatory virtue to the extent it can be linked to generalisations' degree of invariance. Far from being a *distinctive* explanatory virtue, however, unification comes out as one of at least seven different 'dimensions' of explanatory depth, viz. different ways in which generalisations' degree of invariance can vary (ibid.). It can be furthermore argued that some of these different dimensions are *systematically in conflict*, and that scientists' weighing of different dimensions—their assessment of explanatory goodness, in other words—is typically contextual and pragmatic in nature (Ylikoski and Kuorikoski, 2010).

Arguably we can account in these terms for the prevailing general intuition (supporting MC3) that unification is an explanatory virtue (Ylikoski

³⁴For example, Newton's gravitational law is invariant under changes in the specific masses and their relative distance.

and Kuorikoski, 2010). We can also specify more precisely the conditions under which the intuition holds.³⁵ The upshot is that unification is connected to explanatory goodness in a rather specific and less objective way than assumed by the advocates of MC3. Furthermore, the notion that we can compare different *theories* in relation to this explanatory virtue sits somewhat uncomfortably in the Woodwardian framework, which is primarily concerned with analysing the depth of *individual explanations*. Thus, we can make good sense of the notion that an explanation of Uranus’s orbit in terms of general relativity is better than in terms of Newtonian mechanics. Or that the latter explanation does not improve if it cites the ideal gas law in conjunction with Newton’s law, even though it (artificially) ‘unifies’ two types of phenomena in the covering-law sense. But for this we do not need the notion that Einstein’s *theory* is more explanatory than Newton’s by virtue of being more unifying. The notion that general relativity as a theory has more ‘explanatory power’ than Newtonian mechanics may or may not make derivative sense on the basis general relativity in general furnishing explanations that are individually better in the relevant sense.³⁶ But such derivative (and seemingly unneeded) notion of explanatory power is intelligible (at best) when the individual explanations can in the first place be compared in the Woodwardian sense. This particular analysis of explanatory unification thus bestows no general notion of unification-qua-explanatory-virtue applicable outside the province of Woodward’s account. Thus, in as far as Woodward’s account captures the unificatory virtue of scientific explanations, and in as far this account fails to capture the sense in which metaphysical theories explain, one does not find in science support for the dictum that in general the most unified theories—whether scientific or metaphysical—are the likeliest to be true since they (*ceteris paribus*)

³⁵The explanatory power of an explanation is increased by integrating it into a unifying theory *only if* it increases the amount of explanatory modal information provided. According to Ylikoski and Kuorikoski (2010, p. 214):

[Such] integration [should] be more than formal compatibility or conceptual coherence. The bodies of knowledge should come together in a manner that allows new relevant inferences to be made about the phenomenon. The requirement of relevance makes the idea of integration a local notion: it depends on one’s epistemic aims. This contextual notion is quite different from the global notion of unification advocated by the supporters of unification accounts of explanation.

Furthermore, arguably such integration can often be achieved only by compromising other dimensions of explanatory depth, for example by idealising away fine-grained detail and accuracy.

³⁶Cf. Ylikoski and Kuorikoski 2010.

explain the best.

I have flagged Woodward's account as an important example of contemporary philosophy of science that sheds light on scientific explanation in a way that considerably complicates the simple picture underlying MC3, according to which scientists appeal to unification as a theoretical virtue much like metaphysicians do, via inference to the best explanation. Woodward's account is not the only game in town, of course, but other notable contemporary accounts of explanation are equally hostile to the idea that unification is an explanatory virtue of a *theory*. The 'kairitic account' of Strevens (2008), for example, values generality in some very specific ways in locating its explanatory difference-makers, but it does not associate any kind of explanatory virtue with unification. So also in light of this recent account of scientific explanation there is reason to worry about the commonplace idea that metaphysical theories can be appraised in terms of their unifying explanatory power simply 'because scientific explanations are so appraised.'

Of course, none of the above discussion goes to show that unification is not a *theoretical* virtue in science. But even if it is—even if more unified theories are generally better theories, *ceteris paribus*, by the lights of science—the question is whether unification should be thought of as an *epistemic* or merely *pragmatic* virtue. Here again it is natural to turn to the history of science, which indicates that unification is often best thought of as a pragmatic virtue, not as an epistemic guide to what the world is like. For example, in her book-length study of unification, explanation, and realism in science, Morrison (2000) points to various theoretical unifications in physics and biology that are not explanatory (as she argues), or at least were achieved by means of mathematical structures and theoretical constructs that do not reflect reality.³⁷ The truth is that the historical record attested by Morrison is a mixed-bag: sometimes theoretically virtuous unification is a guide to ontology (e.g. Maxwell on electricity and magnetism), but often it is not (e.g. the electroweak theory, or Kaluza-Klein theory).³⁸ Given

³⁷Morrison (2000, p. 232) sums up her take-home message thus:

[N]o single account of theory unification can be given. A philosophical consequence of that claim is that unity should not be linked to truth or increased likelihood of truth; unification cannot function as an inference ticket.

There can be no doubt that unity exists in science, that unified theories have been enormously successful and that unity is a goal pursued by many practicing scientists in a variety of fields. But nothing about a corresponding unity in nature follows from those facts...

³⁸See also Karaca 2012, Maudlin 1996.

this lesson from the history, it is only right that in connection with some exemplar cases scientists themselves have not been epistemically swayed by mere unification, but have demanded *independent evidence*, as in the case of Maxwell's theory the vindication of which had more to do with the independent evidence for electromagnetic waves from Hertz's experiments conducted some years after Maxwell's unification.

* * *

Contemporary philosophy of scientific explanation paints a fascinating naturalistic picture the details of which challenge in various ways the *strategy of methodological similarity*. This strategy, recall, aims to justify explanationism in metaphysics on the basis of its role and (assumed) epistemic status in science. For the sake of the argument I have granted that explanationism in science is rational and reliable, although obviously the sense of reliability has to be qualified in quite specific ways to fit the picture painted above. Can scientific realists actually defend a sufficiently robust, yet suitably qualified reliability of explanationism in science? Whatever the answer, we have reviewed a number of relevant aspects of explanation and explanationism in science that collectively throw considerable doubt on the abstract vindications of metaphysical explanationism.

The explanationist can retreat: instead of trying to justify *metaphysical* instances of inference to the best explanation, she can argue that some metaphysical views get support directly through *scientific* instances of inference to the best explanation. This is the gambit of confirmational holism to which I now turn.

5 On the strategy of confirmational holism

Naturalistic metaphysics sometimes engages in exploring the ontological commitments of science, as in the case of the so-called indispensability arguments. I will now examine the connection between these arguments and scientific explanationism.

The best known indispensability argument views mathematical realism as a *corollary* of scientific realism (Colyvan 2006). Briefly and roughly put, in the Quinean holistic picture of science the empirical justification for theoretical assumptions 'bleeds over' to (applied) mathematics. As Psillos (2012, p. 53) puts it:

[I]ndispensability arguments capitalise on the strengths of scientific realism, and in particular of the no-miracles argument

(NMA), in order to suggest that a) the reality of mathematical entities (in their full abstractness) follows from the truth of (literally understood) scientific theories; and b) there are good reasons to take certain theories to be true.

An explanationist can also add that empirical justification is ‘mediated’ via inference to the best explanation, as this is how we can arguably construe the relation between our best theories and their empirical support. Furthermore, the argument for scientific realism can itself be an instance of inference to the best explanation, as in the case of the no-miracles (NMA) argument:

The epistemic optimism characteristic of scientific realism is based on NMA. The argument, roughly put, is that empirical success (suitably regimented so as to include novel predictions and the like) offers good reasons to believe in the truth of theories, since it is best explained by the claim that theories are true. Thus conceived, NMA is blind to a distinction between abstract entities and concrete ones insofar as commitment to both types is implied by the truth of (literally understood) scientific theories. (Ibid.)

The original (Quine-Putnam) indispensability argument, recited by Psillos, concerns ontological commitment of *literally true* theories. What should we make of the fact that in light of the above discussion we are not justified in taking explanatory inferences in science as delivering literal truths? If mathematical realism is but a corollary of scientific realism, presumably it matters what kind of scientific realism it is rational for us to maintain? Taking on board our complete understanding of science and its explanatory endeavours, we must admit that the extent of our ability to answer ontological questions with science, naturalistically, depends on whether or not, and in what sense, scientific explanationism is reliable. The history of science indicates that the reliability of explanationism is curtailed: explanationism allows us to track the unobservable reality only in limited respects (cf. §4.1). Such curtailed reliability can still generate considerable empirical successes (including novel predictions and the like), if the outputs of explanatory inferences selectively latch onto appropriate features of the world—namely those features that entail the right predictions etc.³⁹ Hence, the sort of no-miracles argument alluded to by Psillos above is *over*-optimistic, since the

³⁹For example, although Fresnel’s ether theories of optics or Newtonian gravity are far from being literally true and their explanatory posits do not exist, these theories arguably selectively latch onto reality so as to generate hugely impressive empirical successes.

best realist explanations of the empirical success of science need not be in terms of the *literal* truth of theories. In the history of science the realist must repeatedly explain the empirical success of theories that are only partially veridical, in a way that is compatible with the curtailed reliability of the scientific method.

We have reason to worry about indispensability arguments, since we have reason to worry that realist explanations of empirical success fail to support mathematical realism.⁴⁰ It is a considerable, hitherto unsettled question whether there is any support for mathematical realism (or other paradigmatically metaphysical views, such as presentism) to be found in the best scientific realist arguments. I am skeptical, but I won't argue for this here. My present point is that our overarching understanding of scientific explanation and explanationism can challenge the strategy of confirmational holism according to which the scientific realist should by her own lights be committed to paradigmatically metaphysical views.

The recent literature on the indispensability argument displays a broad consensus according to which scientific realist commitments cannot be read off from our best theories construed as true *simpliciter*. Philosophers are nowadays much less concerned with mathematics' *unqualified* indispensability in science—the unvarnished Quinean notion that our best scientific theories simply quantify over mathematics. The focus has rather shifted to mathematics' indispensability for scientific *explanations*.⁴¹ This shift presents a further explanationist twist on the strategy of confirmational holism.⁴² The champions of the new 'explanatory indispensability argument' envisage a more direct route from scientific realism to mathematical

⁴⁰I think Hawley (2006) is methodologically on the right track in thinking that in naturalistic metaphysics “we have reason to believe metaphysical claims to the extent that they are *genuinely involved* in generating empirical success” (p. 460, my emphasis). Hawley argues that the indispensable role of metaphysical assumptions in science can be revealed by their involvement in the best realist explanations of some empirical successes of science. But while I broadly agree with Hawley's methodological pronouncement, I see no grounds for her optimistic belief that “there are actual cases in which the involvement of a metaphysical claim in an empirically successful scientific theory provides some reason to think that the claim is true.” (p. 456)

⁴¹See, for example, Colyvan (2012).

⁴²Baker (2005) puts it thus:

A crucial plank of the scientific realist position involves inference to the best explanation (IBE) to justify the postulation in particular cases of unobservable theoretical entities. . . . [T]he indispensability debate only gets off the ground if both sides take IBE seriously, which suggests that explanation is of key importance in this debate. (225)

realism, based on the notion that (for a scientific realist) ontological commitment and scientific explanation are directly connected. Realist commitment to mathematics allegedly follows from the admission that mathematics plays a genuine ‘explanatory role’ in science.

Although the focus on explanation admittedly enhances the indispensability argument, our best understanding of scientific explanation yet again considerably complicates the key issue at stake. The (burgeoning) literature on the explanatory indispensability argument has by and large taken for granted the connection between explanatory indispensability and ontological commitment, without any reference to a particular conception of explanation to underwrite this connection. Prima facie innocent assumptions concerning explanations’ ontological commitments turn out to harbour various complications, however, when examined in relation to specific accounts of explanation. In particular, the literature on the indispensability argument is rife with references to ‘mathematics’ (indispensable) explanatory role,’ but the key notion of ‘explanatory role’ has been mostly left unanalysed. This is a major shortcoming: the notion of ‘mathematics’ explanatory role’ must be examined in relation to different analyses and conceptions of explanation to properly judge whether mathematics plays *the kind of* ontologically committing explanatory role that matters for the indispensability argument. This is critical because in the context of different accounts of explanation we can delineate *different kinds* of explanatory roles. The admission that mathematics is in *a* sense ‘genuinely explanatory’, or is ‘playing an explanatory role’, does not by itself at all imply that it is playing the *right* kind of explanatory role. Saatsi (forthcoming **a**) draws some critical distinctions between different types of explanatory roles—in connection with some leading ‘ontic’ accounts of explanation in particular—pointing to various hitherto unappreciated challenges faced by the strategy of confirmational holism. A properly informed analysis of scientific explanation is again required in order to see whether scientific practice can underwrite a naturalistic argument to a metaphysical conclusion (viz. mathematical realism).

6 Coda

Swoyer, a keen advocate of explanationism in metaphysics, engages in a contemporary debate on abstract objects with Dorr (Sider et al. 2007). The two debaters largely agree on the methodology of their domain of enquiry: metaphysical theories are to be evaluated on their explanatory merits, where those merits are judged like the explanatory merits of scientific theories. In

general, Dorr grants that:

There is something very attractive about the idea that we should try to make progress in philosophy by learning from the disciplines in which progress is most manifest, namely the sciences. More specifically, *the proposal is that we should take as our starting point the large and impressive body of case-by-case epistemological judgments shared by all scientific realists*. We then decide what we ought to believe about controversial philosophical questions in accordance with whichever epistemological theory does the best job of accounting for and systematizing these data. (2007, p. 43, my emphasis)⁴³

I agree. But as far as there are any judgments shared by all scientific realists, these judgments raise considerable challenges to explanationism in metaphysics. Inference to the best explanation in metaphysics is badly in need of justification that is properly grounded in our best understanding of the nature of scientific theorising and explanation, and the role and limits of inference to the best explanation in science.

References

- Armstrong, D. (1985). *What is a Law of Nature?* Cambridge University Press.
- Armstrong, D. (1997). *A world of states of affairs*. Cambridge University Press.
- Baker, A. (2005). Are there genuine mathematical explanations of physical phenomena? *Mind* 114, 223–238.
- Baker, A. (2009). Mathematical explanation in science. *British Journal for the Philosophy of Science* 60(3), 611–633.
- Barnes, E. (1992). Explanatory unification and scientific understanding. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1992, pp. 3–12.
- Bechtel, W. and A. Abrahamsen (2005). Explanation: The Mechanist Alternative. *Studies in History and Philosophy of the Biological and Biomedical Sciences* 36, 421–441.

⁴³This is how Dorr understands naturalism.

- Bigelow, J. (2010). Quine, mereology, and inference to the best explanation. *Logique et Analyse* 53(212), 465–482
- Bigelow, J. and R. Pargetter (1990). *Science and necessity*. Cambridge: Cambridge University Press.
- Biggs, S. (2011). Abduction and modality. *Philosophy and Phenomenological Research* 83(2), 283–326.
- Cameron, R. and R. P. Cameron (2008). Turtles all the way down: regress, priority and fundamentality. *The Philosophical Quarterly* 58(230), 1–14.
- Chang, H. (2003). Preservative Realism and Its Discontents: Revisiting Caloric. *Philosophy of Science* 70, 902–912.
- Colyvan, M. (2006). Scientific Realism and Mathematical Nominalism: A Marriage Made in Hell. In C. Cheyne and J. Worrall (Eds.), *Rationality and Reality: Conversations with Alan Musgrave*, pp. 225–37. Springer.
- Colyvan, M. (2012). Road work ahead: Heavy machinery on the easy road. *Mind* 121(484), 1031–1046.
- Cordero, A. (2011). Scientific realism and the divide et impera strategy: The ether saga revisited. *Philosophy of Science* 78(5), 1120–1130.
- De Regt, H. W. (2001). Spacetime visualisation and the intelligibility of physical theories. *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 32(2), 243–265.
- De Regt, H. W. and D. Dieks (2005). A Contextual Approach to Scientific Understanding. *Synthese* 144, 137–170.
- Doppelt, G. (2005). Empirical success or explanatory success: What does current scientific realism need to explain? *Philosophy of Science* 72(5), 1076–1087.
- Doppelt, G. (2007). Reconstructing scientific realism to rebut the pessimistic meta-induction. *Philosophy of Science* 74(1), 96–118.
- Dorato, M. (2012). Mathematical biology and the existence of biological laws. In *Probabilities, Laws, and Structures*, pp. 109–121. Springer.
- Dorr, C. (2008). There are no abstract objects. In T. Sider, J. Hawthorne, and D. Zimmerman (Eds.), *Contemporary Debates in Metaphysics*, pp. 32–64. Blackwell.

- Douglas, H. E. (2009, October). Reintroducing prediction to explanation. *Philosophy of Science* 76(4), 444–463.
- Ellis, B. (2009). *The Metaphysics of Scientific Realism*. Acumen.
- Field, H. (1989). *Realism, mathematics, and modality*. Oxford: Blackwell.
- Friedman, M. (1974). Explanation and Scientific Understanding. *Journal of Philosophy* LXXI, 5–19.
- Frost-Arnold, G. (2010). The No-Miracles argument for realism: Inference to an unacceptable explanations. *Philosophy of Science* 77(1), 35–58.
- Gijssbers, V. (2007). Why unification is neither necessary nor sufficient for explanation. *Philosophy of Science* 74(4), 481–500.
- Grimm, S. (2009). Reliability and the sense of understanding. In K. E. Henk W. de Regt, Sabina Leonelli (Ed.), *Scientific Understanding: Philosophical Perspectives*, pp. 83–99. University of Pittsburgh Press.
- Harker, D. (2013). How to split a theory: Defending selective realism and convergence without proximity. *The British Journal for the Philosophy of Science* 64(1), 79–106.
- Harman, G. (1965). The Inference to the Best Explanation. *Philosophical Review* 74, 88–95.
- Hawley, K. (2006). Science as a guide to metaphysics. *Synthese* 149, 451–470.
- Jones, R. (1991). Realism about what? *Philosophy of Science* 58, 185–202.
- Karaca, K. (2012). Kitcher’s explanatory unification, Kaluza–Klein theories, and the normative aspect of higher dimensional unification in physics. *The British Journal for the Philosophy of Science* 63(2), 287–312.
- Keller, E. F. (2002). *Making sense of life: Explaining biological development with models, metaphors, and machines*. Harvard University Press.
- Kitcher, P. (1981). Explanatory unification. *Philosophy of Science*, 507–531.
- Kuhn, T. S. (1977). Concepts of cause in the development of physics. In *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago: Chicago University Press.

- Kuipers, T. A. (2002). Beauty, A Road to the Truth. *Synthese* 131, 291–328.
- Ladyman, J. (2007). Does Physics Answer Metaphysical Questions? *Royal Institute of Philosophy Supplement* 82(61), 179–201.
- Ladyman, J. (2012). Science, metaphysics and method. *Philosophical Studies*.
- Laudan, L. (1981). A Confutation Of Convergent Realism. *Philosophy of Science* 48, 19–49.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
- Lipton, P. (1999). *Inference to the best explanation*. London: Routledge.
- Lombrozo, T. (2011). The instrumental value of explanations. *Philosophy Compass* 6(8), 539–551.
- Lyons, T. D. (2006, September). Scientific realism and the stratagema de divide et impera. *British Journal for the Philosophy of Science* 57(3), 537–560.
- Maudlin, T. (1996). On the Unification of Physics. *Journal of Philosophy* 93(3), 129–144.
- McLaughlin, B. P. (2010). Consciousness, type physicalism, and inference to the best explanation. *Philosophical Issues* 20(1), 266–304.
- Morrison, M. (2000). *Unifying Scientific Theories*. Cambridge: Cambridge University Press.
- Norton, J. D. (2003). A Material Theory of Induction. *Philosophy of Science* 70, 647–670.
- Okasha, S. (2000, December). van Fraassen’s critique of inference to the best explanations. *Studies In History and Philosophy of Science* 31(4), 691–710.
- Pargetter, R. (1984). The scientific inference to other minds. *Australasian Journal of Philosophy* 62(2), 158–163.
- Paul, L. (2012). Metaphysics as modeling: the handmaiden’s tale. *Philosophical Studies*, 1–29.
- Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. London: Routledge.

- Psillos, S. (2005). Scientific Realism and Metaphysics. *Ratio* 18, 385–404.
- Psillos, S. (2011). The scope and limits of the no-miracles arguments. In F. Stadler (Ed.), *The Philosophy of Science in a European Perspective*, Volume 2. Springer.
- Psillos, S. (2012). Anti-nominalistic scientific realism: a defence. In A. Bird, B. Ellis, and H. Sankey (Eds.), *Properties, Powers and Structures: Issues in the Metaphysics of Realism*. New York: Routledge.
- Putnam, H. (1975). *Mathematics, Matter, and Method*, Volume Vol. 1. Cambridge: Cambridge University Press.
- Rogeberg, O. (2004). Taking absurd theories seriously: Economics and the case of rational addiction theories. *Philosophy of science* 71(3), 263–285.
- Saatsi, J. (2009). Form vs. content-driven arguments for realism. In P. Magnus and J. Busch (Eds.), *New Waves in Philosophy of Science*, pp. 8–28. Basingstoke: Palgrave.
- Saatsi, J. (2012). Scientific realism and historical evidence: Shortcomings of the current state of debate. In S. O. S. de Regt, Henk W.; Hartmann (Ed.), *EPSA Philosophy of Science: Amsterdam 2009*, pp. 329–340. Dordrecht: Springer.
- Saatsi, J. (Forthcoming a). On the ‘indispensable explanatory role’ of mathematics. *Mind*.
- Saatsi, J. (Forthcoming b). On historical inductions, old and new.
- Sellars, W. (1963/1991). *Science, Perception and Reality*. Atascadero, CA: Ridgeview Publishing Company.
- Sider, T. (2007). Parthood. *The Philosophical Review* 116(1), 51–91.
- Sider, T. (2011). *Writing the Book of the World*. Oxford University Press.
- Sider, T., J. Hawthorne, and D. Zimmerman (2008). *Contemporary debates in metaphysics*. Blackwell.
- Stanford, K. P. (2006). *Exceeding Our Grasp*. Oxford: Oxford University Press.
- Stanford, P. K. (2003). No Refuge for Realism: Selective Confirmation and the History of Science. *Philosophy of Science* 70(5), 913–925. PSA Proceedings.

- Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard: Harvard University Press.
- Sturgeon, N. (2006). Moral explanations defended. In J. Dreier (Ed.), *Contemporary debates in moral theory*, Volume 6. Wiley-Blackwell.
- Swoyer, C. (1999). Explanation and inference in metaphysics. *Midwest Studies in Philosophy* 23, 100–131.
- Swoyer, C. (2008). Abstract entities. In T. Sider, J. Hawthorne, and D. W. Zimmerman (Eds.), *Contemporary Debates in Metaphysics*, pp. 11–31. Blackwell.
- Trout, J. D. (2007). The psychology of scientific explanation. *Philosophy Compass* 2/3, 564–591.
- Vickers, P. (2013). A confrontation of convergent realism. *Philosophy of Science* 80(2), pp. 189–211.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- Woodward, J. (2010). Causation in biology: stability, specificity, and the choice of levels of explanation. *Biology and Philosophy* 25(3), 287–318.
- Woodward, J. and C. Hitchcock (2003). Explanatory Generalizations, Part 2: Plumbing Explanatory Depth. *Noûs* 37(2), 181–199.
- Ylikoski, P. and J. Kuorikoski (2010). Dissecting explanatory power. *Philosophical Studies* 148, 201–219.