



Title	Explanatory Rivals and the Ultimate Argument
Authors(s)	Dellsén, Finnur
Publication date	2016-09
Publication information	Dellsén, Finnur. "Explanatory Rivals and the Ultimate Argument." Wiley, September 2016. https://doi.org/10.1111/theo.12084 .
Publisher	Wiley
Item record/more information	http://hdl.handle.net/10197/8175
Publisher's statement	This is the author's version of the following article: Finnur Dellsen (2016) 'Explanatory Rivals and the Ultimate Argument' Theoria, 82(3) : 217-237 which has been published in final form at http://dx.doi.org/10.1111/theo.12084 .
Publisher's version (DOI)	10.1111/theo.12084

Downloaded 2026-05-01 23:48:06

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd_oa)



© Some rights reserved. For more information

Explanatory Rivals and the Ultimate Argument

Finnur Dellsén

Penultimate draft – please cite the published version in *Theoria*

Abstract

Although many aspects of Inference to the Best Explanation (IBE) have been extensively discussed, very little has so far been said about what it takes for a hypothesis to count as a *rival* explanatory hypothesis in the context of IBE. The primary aim of this paper is to rectify this situation by arguing for a specific account of explanatory rivalry. On this account, explanatory rivals are (roughly speaking) complete explanations of a given explanandum. When explanatory rivals are conceived of in this way, I argue that IBE is a more plausible and defensible rule of inference than it would otherwise be. The secondary aim of the paper is to demonstrate the importance of accounts of explanatory rivalry by examining a prominent philosophical argument in which IBE is employed, viz. the so-called *Ultimate Argument* for scientific realism. In short, I argue that a well-known objection to the Ultimate Argument due to Arthur Fine fails in virtue of tacitly assuming an account of explanatory rivalry that we have independent reasons to reject.

1 INTRODUCTION

Inference to the Best Explanation (IBE) is an ampliative form of inference in which one accepts a hypothesis on the grounds that it is a better explanation than any rival hypothesis. Many aspects of IBE have been extensively

discussed, e.g. what an *explanation* is and what makes one explanation *better* than another. (See, e.g., Hempel, 1965; Thagard, 1978; Kitcher, 1981; Salmon, 1984; Lipton, 2004; Poston, 2014) However, very little has so far been said about what it takes for a hypothesis to count as a *rival* hypothesis in the context of IBE – i.e. be such that a hypothesis may be inferred according to IBE just in case it is the best among hypotheses of that kind.

This is surprising since the concept of rival explanatory hypotheses occupies such a central role in all of the most prominent characterizations of IBE. Consider Harman (1965):

In making this inference one infers, from the fact that a certain hypothesis would explain the evidence, to the truth of that hypothesis. In general, there will be several hypotheses which might explain the evidence, so one must be able to reject all such alternative hypotheses before one is warranted in making the inference. Thus one infers, from the premise that a given hypothesis would provide a “better” explanation for the evidence than would any other hypothesis, to the conclusion that the given hypothesis is true. (Harman, 1965, 89)

Or consider Thagard (1978):

To put it briefly, inference to the best explanation consists in accepting a hypothesis on the grounds that it provides a better explanation of the evidence than is provided by alternative hypotheses. (Thagard, 1978, 77)

Finally, consider Lipton (2004):

According to Inference to the Best Explanation, our inferential practices are governed by explanatory considerations. Given our data and our background beliefs, we infer what would, if true, provide the best of the competing explanations we can generate of those data [...] (Lipton, 2004, 56)

All of these authors agree that IBE licenses an inference to a hypothesis just in case the hypothesis is better (*qua* explanation) than some other hypotheses

that would, if true, explain the same facts. (Harman and Thagard refer to these other hypotheses as “alternative hypotheses”, while Lipton refers to them as “competing hypotheses”. I will refer to them as “explanatory rivals” in order to emphasize the role such hypotheses are meant to play in IBE.)

Unfortunately, however, these authors include no systematic discussion about what it takes to be an explanatory rival (i.e. “competing” or “alternative” hypothesis). My primary aim in this paper is to argue for a specific conception of what explanatory rivals are – in other words, I will argue for an account of what kinds of hypotheses are being compared in an IBE. Since this is part of what IBE is, I shall in effect argue for a partial view of the very nature of IBE itself. Specifically, I argue that a hypothesis and its explanatory rivals should be mutually exclusive, and then suggest how to model explanatory rivals in the context of IBE given this constraint by conceiving of explanatory rivals as complete potential explanations of a given explanandum.

One might wonder why we should care about explanatory rivals in the first place. What does it matter what structural properties explanatory rivals have or don’t have? Well, one reason to care about explanatory rivals is that without an account of explanatory rivals we have only an incomplete account of IBE itself. Thus, insofar as we care about what IBE is, we ought to care about explanatory rivals. A perhaps more tangible reason to care, which will be discussed in the next section, is that the issue of what counts as an explanatory rival is crucial to evaluating what is perhaps the most prominent use of IBE in the philosophical literature, viz. the so-called *Ultimate Argument* for scientific realism. Indeed, the secondary aim of this paper is to defend the Ultimate Argument against a well-known objection due to Arthur Fine. I shall argue, in brief, that Fine’s objection rests on an implausible conception of explanatory rivals – one that we have independent reasons to reject.

The structure of the paper is as follows. In order to motivate the issue of what explanatory rivals are, I start by examining Fine’s objection to the Ultimate Argument in section 2. In section 3, I then make a number of preliminary points about IBE in order to set up the discussion to follow. Section 4 starts by arguing that explanatory rivals should be mutually exclusive (§4.1), then presents an apparent counterexample to this requirement (§4.2), and finally presents an account of explanatory rivals in terms of complete explanations

that avoids the counterexample while satisfying the requirement (§4.3). Section 5 revisits the Ultimate Argument in light of the discussion in section 4, arguing that Fine’s objection to the argument rests on an objectionable assumption about explanatory rivals.

2 THE ULTIMATE ARGUMENT FOR SCIENTIFIC REALISM

The so-called *Ultimate Argument* for scientific realism has a venerable history. Versions of the argument have been articulated by Maxwell (1962), Smart (1963), Putnam (1975), Boyd (1980), McMullin (1987), Musgrave (1988), Lipton (1994), Leplin (1997), and Psillos (1999). In its simplest form, the argument runs as follows:¹

(U1) Scientific theory T is empirically successful.

(U2) The best explanation for T ’s empirical success is that T is true.²

(U3) Thus, by IBE, (it is reasonable to believe that) T is true.³

A common objection to this argument challenges the inferential step from (U1) and (U2) to (U3). According to this objection, it is *not* reasonable to believe the best explanation of T ’s empirical success. (Laudan, 1981; van Fraassen, 1989) In effect, this objection challenges IBE itself, seen as a rule of inference which takes one to conclusions that are epistemically justified. The objection I want to consider, however, grants the realist that if (U1) and (U2) are both true (U3) would follow, but instead challenges (U2).

The basic idea behind this objection, which is due to Fine (1986, 1991),⁴

¹I shall later suggest that this version of the argument is a bit *too* simple (see §5).

²Here and throughout, I am using the truth predicate in the minimal or deflationary sense in which the claim that T is true is equivalent to T itself. Thus I am not assuming that any particular theory of truth plays any essential role in the Ultimate Argument. (Levin, 1984)

³It is in some ways more accurate to call this an argument-*schema*, since it applies to a given theory T (e.g. the atomic theory of matter), arguing that (it is reasonable to believe that) T is true.

⁴A similar objection is also often attributed to van Fraassen (1980), though it is unclear

is that T 's truth is never the best explanation for T 's empirical success since one can always construct a better rival explanation. Fine takes this to be established by a “metatheorem” which states that “to every good realist explanation there corresponds a better instrumentalist one.” (Fine, 1986, 154) If Fine is right, it would follow that no IBE to a given theory being true (as opposed to empirically adequate) is warranted – even on the assumption that the conclusions of IBEs are generally epistemically justified. According to Fine’s objection, then, IBE simply does not apply to any scientific theories, for there is always a rival instrumentalist explanation that makes (U2) false.

The instrumentalist explanation to which Fine appeals is the proposition that the given theory T is *instrumentally reliable*, i.e. “lead[s] only to instrumentally correct conclusions”. (Fine, 1986, 153) Unfortunately, Fine does not specify precisely what he means by “instrumental reliability”. However, this notion is clearly quite similar to van Fraassen’s (1980) better-known notion of “empirical adequacy” on which a theory is empirically adequate (roughly) just in case it is correct in all its claims about the observable aspects of the world.⁵ Although there may well be some differences between Fine and van Fraassen’s notions, they will not be important in what follows.⁶ I will therefore substitute Fine’s “instrumental reliability” for van Fraassen’s “empirical adequacy”

whether van Fraassen anywhere makes an objection of this sort. Van Fraassen did object to (U2) with his well-known “evolutionary” argument, but that argument is quite different from the objection I am attributing to Fine here. (van Fraassen, 1980, 39-40)

⁵This rough characterization of empirical adequacy will do for the purposes of this paper. For a much more precise characterization, see (van Fraassen, 1980, chapter 3).

⁶What’s important here is only that Fine needs there to be some distance between T 's success and the rival explanation he offers, for otherwise the explanans and explanandum below will be the same. Of course, Fine also needs there to be some distance between his rival explanation and the realist’s explanation, i.e. that T is true. (Leplin, 1988) Both the claim that T is instrumentally reliable and the claim that T is empirically adequate constitute such intermediate claims since they are both stronger than the claim that T is successful but weaker than the claim that T is true.

It is also worth noting that there are important differences between van Fraassen’s constructive empiricism and the instrumentalism to which Fine is referring in his objection. While an instrumentalist would hold that accepting a scientific theory does not involve believing any part of that theory, van Fraassen’s empiricism holds that accepting a theory does involve the belief that the theory is empirically adequate (but not the belief that it is true). However, since both instrumentalism and van Fraassen’s constructive empiricism hold that the realist is wrong to claim that accepting a theory involves believing that it is true, the

in the rest of the paper, referring to Fine’s instrumentalist explanation as the proposition that a given theory T is empirically adequate. Let us call this claim T_E .

Now, Fine’s “metatheorem” essentially claims that T_E (that T is empirically adequate) is always a better explanation of T ’s empirical success than T itself, and thus that (U2) is false for any T . Fine is somewhat opaque on how he takes this claim to be established. He does say:

Since no further work is done by ascending from that intermediary [i.e. T_E] to the realist’s ‘truth’, the instrumental explanation has to be counted as better than the realist one. In this way the realist argument leads to instrumentalism. (Fine, 1986, 154)

Even granting (for the sake of the argument) that the realist explanation does “no further work”, there is still a question of why the instrumentalist explanation is necessarily *better* than the realist one. After all, the realist might reply, the explanation provided by T might be considered deeper, less *ad hoc*, and perhaps even simpler, than the explanation provided by T_E .⁷ Although Fine does not address this point explicitly, he seems to be alluding to the fact that T is a strictly stronger proposition than T_E – i.e. that T entails T_E and not vice versa.⁸ Now, by one of the most elementary theorems of the Kolmogorov axioms of probability, it follows that T_E is necessarily more likely to be true than T . So if explanatory goodness and likelihood of being true go hand in hand – i.e. if a more probable explanation is necessarily better than a less probable explanation of the same facts – then T_E must be a better explanation than T . This holds for any T and T_E , irrespective of their relative depth, *ad hoc*-ness, and simplicity.

What can realists say in response to this argument? It may be tempting to respond by rejecting the link between probability and explanatory goodness according to which a more probable explanation is necessarily better than its

two positions both stand opposed to scientific realism.

⁷For a response to Fine along these lines, see (Leplin, 1988, 1997).

⁸This is suggested by his calling the reasoning behind the metatheorem “a structural reason”. (Fine, 1986, 154)

less probable counterpart (henceforth: *the probability-explanatoriness link*). Realists might simply want to say that sometimes less probable theories are better explanations than their more probable counterparts; thus, in such cases, we can infer by IBE to comparatively improbable explanations. However, this move is less appealing for the realist than it may seem at first, for it would leave realists in an uncomfortable position vis-à-vis the status of IBE itself. This is so for two closely related reasons.

First, rejecting the probability-explanatoriness link would require a conception of explanatory goodness that conflicts with recent attempts to reconcile Bayesianism and IBE. Van Fraassen (1989) famously argued that IBE and Bayesianism are incompatible, concluding that IBE should therefore be abandoned. Against this, however, several authors have argued that IBE is in fact compatible with, or even supplementary to, Bayesianism. (Okasha, 2000; McGrew, 2003; Lipton, 2004; Weisberg, 2009; Henderson, 2014) Simplifying quite a bit, this approach is based on the idea that explanatory considerations may come into play in determining prior probabilities such that, at the end of the day, more explanatory hypotheses receive a higher posterior probability than less explanatory hypotheses. There is a fairly wide consensus that something along these lines must be the correct conception of IBE, yet without the probability-explanatoriness link this entire research project is doomed from the outset. Given the plausibility and fertility of construing scientific inferences along Bayesian lines, this would be a serious blow to the defensibility of IBE itself.

Second, it is not clear how to understand IBE's *epistemic* status once we reject the probability-explanatoriness link. As Lipton notes, "Inference to the Best Explanation is supposed to describe strong inductive arguments, and a strong inductive argument is one where the premises make the conclusion likely." (Lipton, 2004, 60) To sever the connection between explanatory goodness and likelihood of being true would compromise IBE's claim to be a rule of ampliative inference whose conclusions are epistemically justified. And that, of course, is the very ground upon which the Ultimate Argument itself is built – if IBEs aren't epistemically justified then (U3) does not follow from (U1) and (U2). In other words, although rejecting the probability-explanatoriness link might undermine Fine's "metatheorem", it would constitute a mere pyrrhic

victory in virtue of undermining the epistemic status of IBE itself.

Let us take stock. I have argued that realists should be uncomfortable with rejecting the claim that probability and explanatory goodness go hand-in-hand – the probability-explanatoriness link – without compromising the epistemic status of IBE itself and its compatibility with Bayesianism. However, given the probability-explanatoriness link, it looks hard to reject Fine’s claim that for any T , T_E is a better explanation than T for T ’s empirical success. And this, in turn, seems to show that IBE cannot possibly warrant any inferences to the truth of theories concerning unobservable entities, for an inference to the truth of any theory (however successful) would be undermined by the fact that the truth doesn’t provide a better explanation than its empirical adequacy. This conclusion should be more than a little surprising, since IBE was originally introduced by Harman (1965) partly in order to make sense of ampliative inferences to unobservable entities – something enumerative induction famously cannot do. Of course, anti-realists may take all of this to show the hopelessness of an IBE-based defense of scientific realism – this seems to be Fine’s conclusion – but the realist will start to suspect that she has become the victim of an epistemological rope-trick of sorts.

In what follows, I argue that this suspicion is indeed correct. The problem with Fine’s argument is not the probability-explanatoriness link, but the assumption (left implicit by Fine) that T must be a better explanation than T_E in order for T to be inferred by IBE. Put differently, I shall argue that T and T_E are not explanatory rivals in the context of a plausible version of IBE. This will be done by giving general reasons for construing explanatory rivals as mutually exclusive propositions, and then presenting an account on which that is true.⁹ I will then return to the Ultimate Argument, arguing that no version of Fine’s objection can be made given this account of explanatory rivalry.

⁹It will be important in what follows that since my reasons for construing explanatory rivals as mutually exclusive propositions are quite general and have nothing in particular to do with scientific realism or the Ultimate Argument, they do not beg any questions against Fine or other anti-realists.

3 SOME PRELIMINARY POINTS ON IBE

This section makes several necessary preliminary points about IBE and explanatory rivalry. This will serve to clarify exactly what kinds of inferences we will be concerned with and (relatedly) exactly what “explanatory rivals” is meant to refer to. First of all, I will be following the practice of using “IBE” to refer both to the rule of inference (what is sometimes referred to as “the principle of inference to the best explanation”) and to particular inferences made in accordance with this rule. Context will reveal which of the two is being referred to in a particular case.

Second, I will be assuming that IBE is a rule of epistemic evaluation as opposed to a psychological process that agents either do, or should, explicitly go through in their reasoning. Here I am following Lycan who construes IBE as a normative theory of justification which distinguishes “those beliefs that are justified (to whatever degree) from those that are not, regardless of whether the canons ever are or could be used explicitly in human doxastic practice.” (Lycan, 1985, 141) Putting the point slightly differently, I am not conceiving IBE as a psychological theory of any sort – neither as a theory of what our psychological process of reasoning is really like, nor as a theory of what our psychological process of reasoning ought to be like. It is a theory only of what theories are epistemically justified, and why. What psychological processes do or should cause us to accept these theories is another matter.

Third, some authors formulate IBE as a rule of inference that warrants inferring the best of the *available* rival explanations of some data, where an available explanation is one that has been considered or formulated. On this conception, IBE operates on a space of hypotheses that has already been formulated and warrants inferring the hypothesis in that space that best explains the relevant data. One problem with this formulation of IBE is that it seems to warrant inferring an awful explanation in cases where one simply hasn’t thought of any better explanation. Alternatively, one can formulate IBE as a rule that warrants inferring the best explanation *in logical space*, but this might seem to make IBE unnecessarily “idealistic” since it might seem to require agents to be theoretically omniscient. Lipton (1993) suggests a way to deal with this dilemma: On Lipton’s view, IBE operates not only by choosing

the best of already considered rival explanations, but also by guiding the construction of hypotheses in the first place. Thus, for Lipton, in a successful IBE the set of available rival explanations already contains the best explanation in logical space. Of course, there is no *guarantee* that a given IBE will be successful in this way, but then again proponents of IBE are keen to emphasize that inferences made in accordance with IBE are meant to be fallible.¹⁰ At any rate, I will adopt Lipton’s view in this respect and thus assume in what follows that in *successful* IBEs the best explanation overall will be contained in the set of the best available explanations.

Fourth, and perhaps most importantly, there is a distinction to be made between *direct* and *indirect* IBEs. Suppose I come home one day to discover that the outdoor lock has been broken and that all of my most prized possessions are gone. I infer by IBE that someone broke into my apartment, since that is the best explanation of the state of the apartment. From that I can infer, deductively, that someone other than me has been in my apartment. Yet the second conclusion – that someone has been in my apartment – presumably does not explain why my outdoor lock is broken and my most prized possessions are gone, even though that conclusion was also, in a sense, reached by IBE. It is worth noting that many of Harman’s examples in his seminal “Inference to the Best Explanation” (Harman, 1965) are inferences where the conclusion merely follows from something that purports to explain the data. For example, Harman claims that one may infer by IBE from “All observed *As* are *Bs*” to “The next observed *A* will be *B*”. (Harman, 1965, 91) But surely, that the next observed *A* will be *B* couldn’t possibly by itself explain why all observed *As* are *Bs*. Rather, what explains that all observed *As* are *Bs* is that *all As* are *Bs*, from which it follows that *the next* observed *A* will be *B*. Moreover, Harman uses the point that IBEs “make use of lemmas” in this way to argue that IBE can account for how knowledge can be undermined by false lemmas in Gettier-cases. (Harman, 1965, 91-95)

¹⁰The risk that, in a given IBE, the best of the available explanations is not also the best of all explanations in logical space is of course an important problem for scientific realists who endorse IBE-based epistemologies of science. Indeed, some anti-realists argue that there are positive reasons to think that this situation is typical. (van Fraassen, 1989; Stanford, 2006) I will not be discussing this problem here, though see my (Dellsén, ms).

To accommodate Harman’s conception of IBE, we can allow both *direct* and *indirect* IBEs – the difference being that the conclusion of a direct IBE must in some sense explain the evidence from which one is inferring,¹¹ whereas the conclusion of an *indirect* IBE must only follow from a hypothesis that explains the evidence. Since every proposition follows from itself, *any* IBE (be it direct or indirect) can be characterized as an inference to H from E based on the fact that H follows from the hypothesis that provides the best explanation of E . This is illustrated in Figs. 1 and 2.

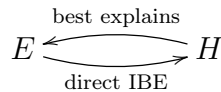


Figure 1: Direct IBE from E to H .

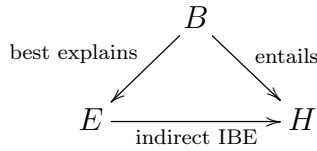


Figure 2: Indirect IBE from E to H .

Having clarified what I take IBE to be in these four respects, we can now finally define the explicandum, viz. what I call “explanatory rivals” (what others call “alternative” or “competing” hypotheses). I shall use “explanatory rival” to refer to any hypothesis in a set S such that an IBE from some evidence E to some hypothesis H is warranted just in case H follows from the hypothesis in S that provides the best explanation of E . Roughly, then, explanatory rivals are the things that are being compared in an IBE – the things among which the best explanation must be the best. To illustrate: In Fig. 1, H is an explanatory rival. In Fig. 2, however, it is B (not H) that is an explanatory rival (assuming of course that $H \neq B$).

¹¹I say that the conclusion of a direct IBE must *in some sense* explain the evidence from which one is inferring, because I shall later suggest that we should allow for the possibility that the conclusion of a direct IBE from some evidence E is the “null-explanation” that

Note that this definition picks out explanatory rivals solely in virtue of the role they play in an IBE, and is thus as neutral as possible between different conceptions among proponents of IBE of what properties explanatory rival should be said to possess. The definition is thus also noncommittal on which hypotheses are explanatory rivals, i.e. which hypotheses are contained in the set S . The role of this definition is merely to enable us to pick out the explicandum – referred to as “explanatory rival” – in order to later be able to argue for a particular explication of that explicandum. It is then the role of an *account* of explanatory rivals to specify which hypotheses count as explanatory rivals for some particular evidence E . I will soon argue for such an account, and thus take a stand on which hypotheses get to go into S for a given E , but I want to be clear that the current definition is not meant to take a stand on the issue. Indeed, if the definition did take a stand on that issue I could be accused of begging the question by defining “explanatory rival” in a way that guarantees that some part of my account is true. Relatedly, note also that “explanatory rivals” is a technical term defined functionally in virtue of playing a certain role in IBEs. Accordingly, its definition is not meant to capture the intuitive or common meaning of the expression (nor is it meant to capture the intuitive or common meaning of “alternative” or “competing” hypothesis). I am stressing this point here in order to preempt “counterexamples” consisting of theories for which it is common or intuitive to use the term “explanatory rivals” but do not satisfy the requirements argued for below.

So to present an account of explanatory rivalry is not an attempt at conceptual analysis. Instead, it is an attempt to spell out one aspect of IBE – and thus a way of doing (normative) epistemology. Accordingly, the desiderata for an account of explanatory rivalry have to do not with whether it will seem intuitive to say that some theories are explanatory rivals, but whether the account makes IBE into a maximally plausible and defensible rule of inference. This point will be important in what follows, where I will argue that the structural properties that I attribute to explanatory rivals prevent IBE from having undesirable features. Accordingly, I shall often speak not about what explanatory rivals *are* but about what they *should be* (if IBE is to be

maximally plausible).

4 EXPLANATORY RIVALS SHOULD BE MUTUALLY EXCLUSIVE

This section first argues that explanatory rivals should be *mutually exclusive* – or, as I shall often say, *pairwise incompatible*. (§4.1) I then discuss an apparent counterexample to this requirement, from which I draw an important lesson about the structure of explanatory rivals. (§4.2) Finally, I present an account of explanatory rivalry on which such rivals are, roughly, proposals for complete explanations of the relevant explanandum. (§4.3)

4.1 THE CASE FOR MUTUAL EXCLUSIVITY

Recall that we defined the set of explanatory rivals as the set of propositions that are being compared in an IBE, such that one may infer (by IBE) any hypothesis that follows from the best explanation in that set. Now, clearly only one hypothesis in a set can be a better explanation than all other hypotheses in the set – this follows from the more general principle that only one element in a set can be better (in some particular way) than all the other elements in that set. So if IBE warrants an inference (from some particular evidence) to one member of the set of explanatory rivals then it cannot also warrant an inference (from the same evidence) to another member of that set. So we can say that the inferability of an explanatory rival (by IBE) *entails* that no other explanatory rival is inferable (by IBE) – under no circumstances can two explanatory rivals both be inferable (by IBE) from the same evidence.

Now, what has this got to do with explanatory rivals being incompatible? Well, consider what it could be about explanatory rivals that makes it true that if IBE warrants inferring one of them then it cannot warrant inferring another. Incompatibility springs to mind: Suppose H_1 and H_2 are incompatible. Then on pain of accepting incompatible hypotheses, one can only infer one of H_1 and

H_2 , not both. Conversely, if two theories are *compatible*, then surely the fact that inferring one of them is warranted by IBE should not by itself prevent one from being warranted in inferring the other by IBE. After all, the two theories may both be extremely plausible in light of one's overall evidence, and their compatibility ensures that they may both be true. This strongly suggests that only in the case of incompatible hypotheses should the fact that one hypothesis is inferable *entail* that the other isn't.¹²

Putting these points together, we have that unless one requires explanatory rivals to be incompatible, being warranted by IBE to infer that one hypothesis is true could *entail* that one is not warranted by IBE to infer that a compatible hypothesis is true – even though the two hypotheses can be as well supported by one's evidence as you like. Putting the point slightly differently, the inferability of one hypothesis (by IBE) would prevent another perfectly compatible (and arbitrarily well-supported) hypothesis from being inferable (by IBE). This seems absurd. By requiring explanatory rivals to be pairwise incompatible, as I suggest we do, we avoid this undesirable consequence and instead get that only incompatible theories are such that the inferability of one entails that another isn't inferable (by IBE).

Of course, this doesn't show that it's impossible not to require explanatory rivals to be pairwise incompatible. It's not incoherent to have a conception of IBE according to which explanatory rivals may be compatible. What it shows is that IBE is more plausible when we require explanatory rivals to be mutually exclusive, since such a requirement avoids the absurd consequence that the inferability of one hypothesis may prevent an (arbitrarily well-supported) compatible hypothesis from being inferable. Thus, what it shows is that proponents of IBE, whose project it is to make IBE maximally plausible, *should* make this requirement on explanatory rivalry, since that makes IBE as a whole more plausible than it would otherwise be.

there is no explanation for E at all.

¹²Of course, this is not to say that whenever two hypotheses aren't both inferable (by IBE) it is because they are incompatible with one another. Rather, it is to say that whenever there is no possibility, whatever one's epistemic situation might be, that two hypotheses are both inferable, it is because they are incompatible. To deny this is to say that under no possible circumstances – whatever one's epistemic situation might be – could two perfectly

Now, this structural requirement on explanatory rivals is admittedly rather abstract, and one might want to see what this “means in practice”. What would explanatory rivals have to look like if they must be pairwise incompatible? I will address this question below, when I have developed an account of explanatory rivalry based on the constraint that explanatory rivals be pairwise incompatible. I’ll motivate my account in a slightly roundabout way, viz. by first considering a *prima facie counterexample* to explanatory rivals being pairwise incompatible. This example shows, I argue, that explanatory rivals should be construed not just as any old explanatory hypotheses, but instead as *proposals for complete explanations* (in a sense to be explicated below).

4.2 A SUGGESTIVE EXAMPLE

Consider the following case due to Lipton (2004):

When my computer did not work, I did not infer that the fuse was blown, since I noticed that the computer was unplugged. These two explanations are logically compatible, since they both could be true, but the plug explanation is known on independent grounds to be correct, and it takes away any reason I would have had to infer that the fuse has blown. Once I accept the plug explanation, there is nothing left for the fuse to explain. (Lipton, 2004, 193)

This example appears to show that explanatory rivals can be pairwise compatible. After all, the plug hypothesis (H_P) is compatible with the fuse hypothesis (H_F). Each hypothesis appears to provide an explanation of the fact that the computer did not work (E_C). Moreover, it seems that being warranted in inferring one of these hypotheses by IBE would prevent one from being warranted in inferring the other. Thus, it may seem, H_P and H_F must be explanatory rivals.¹³

compatible hypotheses both be inferable by IBE from the same evidence.

¹³This is by no means an isolated example, of course. One hypothesis may stand opposed to another for the purposes of IBE because once we accept one hypothesis the other is rendered *explanatorily superfluous*. And, generally, one hypothesis can render another explana-

However, note that although H_P and H_F are certainly compatible, H_P and H_F cannot both be the *complete* explanation for E_C , where a complete explanation consists of all the explanatory information relevant to the explanandum.¹⁴ That is, the following are *not* compatible:

R_P : H_P is the complete explanation for E_C .

R_F : H_F is the complete explanation for E_C .

To see why R_P and R_F are not compatible, note that if the plug being out is the complete explanation for the computer not working, then the fuse being blown doesn't at all explain why the computer doesn't work. Thus, *a fortiori*, the fuse being blown is not the complete explanation for the computer not working. Conversely, if the fuse being blown is the complete explanation for the computer not working, then (by the same token) the plug being out cannot also be the complete explanation.

To a very rough first approximation, I am suggesting that the explanatory rivals in Lipton's case take the form of R_P and R_F , and thus that Lipton's case is no counterexample to the requirement that explanatory rivals be pairwise incompatible. To see why this is only an approximation, however, note that, strictly speaking, neither H_P nor H_F provide a *complete* explanation for E_C , and thus R_P and R_F are both false, strictly speaking. H_P only explains E_C together with auxiliary assumptions such as the hypothesis that the computer needs electricity in order to work (H_E), that the only available source of electricity is through an a nearby outlet (H_O), and so forth. (Similarly for H_F .) Generally, a *complete* explanation will normally appeal to a rather large set of hypotheses (many of which it would be infelicitous to mention in most explanatory contexts).

We can accommodate this point by modifying R_P and R_F as follows:

R_{P+} : The set $\{H_P, H_E, H_O, \dots\}$ (which does not include H_F) is the complete explanation for E_C .

torily superfluous without being incompatible with it.

¹⁴I'll say much more about what I mean by "complete explanation" below.

R_{F+} : The set $\{H_F, H_E, H_O, \dots\}$ (which does not include H_P) is the complete explanation for E_C .

So a more precise version of my suggestion is that explanatory rivals take the general form of R_{P+} and R_{F+} . On this view, what really undermines an IBE to H_F in Lipton’s case is the fact that R_{P+} provides a better explanation than R_{F+} in light of one’s evidence – which in this case includes the fact that “I noticed that the computer was unplugged”. This undermines an IBE to H_F because R_{P+} entails that H_F is not part of the explanation for E_C and thus is not inferable (by IBE) from E_C . In this way, construing the explanatory rivals in this case as R_{P+} and R_{F+} *explains* how H_P undermines H_F .¹⁵

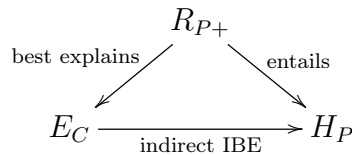


Figure 3: Lipton’s example construed as an indirect IBE.

One might object to this that R_{P+} and R_{F+} are not plausible candidates as conclusions of an IBE because we normally infer *hypotheses* and not *that some hypotheses explain some evidence*. My reply to this objection is twofold: First, recall that I am not conceiving of IBE as a set of epistemic instructions that is, or could be, used explicitly in human reasoning (see also Lycan, 1985). Thus facts about how ordinary agents’ explanatory reasoning actually proceeds are not directly relevant here. Second, recall that an IBE may warrant an inference to a hypothesis that merely *follows from* an explanatory rival and is not itself an explanatory rival. In other words, IBEs can be *indirect* as opposed to *direct*. So although R_{P+} and R_{F+} may not be the sorts of things one normally infers by IBE, it does not follow that they cannot be explanatory rivals. On the view that I am proposing, H_P may be inferred by IBE given that R_{P+} is a better explanation than any other explanatory rival (including R_{F+}), since

¹⁵More generally, my account has it that the fact that one hypothesis is known on independent grounds to be true may undermine an IBE to a compatible hypothesis by virtue of it being reasonable to believe that the best complete explanation that appeals to the first

H_P follows from R_{P+} . (See Fig. 3.) Thus it is the fact that R_{P+} is better than R_{F+} and other explanatory rivals of that form that makes H_P inferable by IBE.

4.3 EXPLANATORY RIVALS AS COMPLETE EXPLANATIONS

Generalizing this proposal, we get that explanatory rivals, relative to some explanandum E , take the following form:

R_X : X is the complete explanation for E .

where X is some set of hypotheses.¹⁶ Thus the set of explanatory rivals relative to a given explanandum E is the set of all propositions of the form R_X , where X ranges over all sets of hypotheses. Roughly, then, my view is that explanatory rivals are different proposals for the complete explanation of a given explanandum.

What is a “complete explanation”? There is a limit to how much can be said about this without assuming a specific account of explanations. What can be said without assuming such an account is that a complete explanation of an explanandum E consists in all the explanatory information relevant to E . Complete explanations can be contrasted with partial explanations, where a partial explanation provides some but not all explanatory information relevant to an explanandum. For example, a partial explanation for why the car crashed into a tree might be that the road was icy, but it is not a complete explanation since a lot of explanatorily relevant information is left out (e.g. that the car was going quite fast, that the tires were not studded, etc.). A complete explanation does not leave out any explanatory information relevant to the explanandum.

What more we say about complete explanations will depend on what ac-

hypothesis does not also appeal to the second hypothesis.

¹⁶Note that I am conceiving of X not as a hypothesis but as a *set* of hypotheses. The set may be a singleton set, a set containing multiple hypotheses, or even the the empty set (more on that below). This *permits* (but does not require) that explanations are sets of several claims (as opposed to just one claim), as it arguably is in the DN-model of Hempel and Oppenheim (1948). (However, when the set in question is a singleton set, I have omitted the curly brackets, i.e. written “ H ” instead of “ $\{H\}$ ”.)

count of explanations we adopt. To illustrate, let us see how the notion can be fleshed out in the two most widely accepted accounts of scientific explanations in the recent literature. Consider first David Lewis's (1986) causal account of explanations, according to which one explains an event by providing some information about the event's causal history. Given Lewis's causal account, a complete explanation of an event can be said to consist in *all* of the information about the event's causal history – information that would specify the event's entire causal history. Indeed, Lewis says something along these lines himself in clarifying his view:

Among the true propositions about the causal history of an event, one is maximal in strength. It is the whole truth on the subject – the biggest chunk of explanatory information that is free of error. We might call this the *whole* explanation of the explanandum event, or simply *the* explanation. (Lewis, 1986, 219)¹⁷

A complete explanation, given the causal account of explanation, would thus be what Lewis calls “the (whole) explanation”.

Consider next Philip Kitcher's (1981) unificationist account of explanation. On this account, an explanation of some fact E is roughly given by a derivation of E using an argument pattern that is (roughly speaking) part of the most unifying set of argument patterns available. Given this account, we may say that a complete explanation specifies the *entire* derivation of E using the requisite argument pattern; a partial explanation, by contrast, merely specifies part of the derivation (perhaps the contextually relevant part). So, for example, on the unificationist account a partial explanation of why the planetary orbits are elliptical could have been that Newton's inverse square law of gravitation is true.¹⁸ A complete explanation, by contrast, would have to add to Newton's law the various facts about the solar system that are, strictly speaking, required in order for the derivation to go through (e.g. that the

¹⁷Similarly, Brad Skow uses the notion of a “complete causal explanation” in a recent defense of the causal account of explanations. (Skow, 2014)

¹⁸Since Newton's law of gravitation is not true, strictly speaking, this is not an actual partial explanation of the planetary orbits being elliptical. But it is a potential partial

planets and the sun are spherical, that the effects of non-gravitational forces are negligible, and that the planets are sufficiently far away from each other for the gravitational forces between them to be negligible as well).

In sum, then, the situation regarding how to describe complete explanations is this: We can only say so much about what a complete explanation is without assuming an account of explanations. The most we can say is that an explanation is complete when it provides all the explanatory information relevant to the explanandum. However, once we assume a particular account of explanations, such as the causal account or the unificationist account, we can elaborate substantially on this basic idea, e.g. by equating a complete explanation with a specification of the explanandum's entire causal history (the causal account) or the entire derivation of the explanandum using the requisite argument pattern (the unificationist account). That said, since I do not want to commit myself to any particular account of explanations in this paper – and since nothing in what follows turns on this – from now on I'll adopt the less informative description and just say that a complete explanation of E provides all the information that is explanatorily relevant to E .

Now, unless one still holds out hope for the principle of sufficient reason, one will accept that some facts do not have any explanation. For example, the fundamental laws of nature, whatever they are, presumably cannot themselves be explained by anything else. Moreover, it's at least plausible that some facts due to chance have no explanation (e.g. why an uranium atom decayed at some particular moment rather than a nanosecond later). Fortunately, this conception of explanatory rivals provides a place for explanatory rivals that say that there is no explanation of a given E . For one set of potentially explanatory hypotheses X is the empty set, \emptyset . The corresponding explanatory rival says that *nothing* is the explanation of E , i.e. that E has no explanation. Thus, on this way of partitioning explanatory rivals, there will always be one explanatory rival – call it “the null rival” – that does *not* provide an explanation of the phenomena in question. This is a welcome consequence because if we thought, for whatever reason, that there was no explanation of some phenomena, then surely we should not infer *any* explanation, not even

the best.^{19,20}

Let us take stock. I have argued that, on the most plausible version of IBE, explanatory rivals are mutually exclusive. Contrary to first appearances, this does not conflict with cases in which an IBE to some hypothesis H is undermined by the existence of a compatible explanation, roughly since the compatible explanation also forms a part of a *complete* explanation that may be better than any complete explanation that appeals to H . This strongly suggests a natural account of explanatory rivalry where the set of explanatory rivals relative to some evidence E is formed by all the different proposals for the complete explanation for E .

5 REVISITING THE ULTIMATE ARGUMENT

Let us now return to Fine's objection to the Ultimate Argument. Recall that Fine's objection rested on the claim that the truth of a theory, T , is never a better explanation than the theory's empirical adequacy T_E . Thus, Fine

explanation, in the sense that it would have been a partial explanation if it had been true.

¹⁹Bird makes a similar point:

The evidence may not require an explanation at all. [...] The subject must have some evidence that rules out the null hypothesis, that there is no explanation. In some cases the null hypothesis may indeed be true. We may take one lesson of quantum indeterminacy to be just that. (Bird, 2005, 14)

One might think that this betrays the slogan "Inference to the Best *Explanation*" in that the non-explanatory explanatory rival would of course not be an explanation at all. But then again the slogan can only be fully respected by requiring that one should infer to an explanation even if one knows full well that there is no such explanation. That, I submit, would be a poor rule of inference by anyone's standard. Better then to admit that the slogan is only an approximation to the rule of inference we should be interested in.

²⁰Note that since every fact either has, or does not have, an explanation, it follows that at least one proposition of the form R_X must be true. After all, if a given fact E has an explanation, then there will be some non-empty set of potentially explanatory hypotheses X for which the corresponding R_X will be true. If however E does not have an explanation, then there is still a set X for which a hypothesis of the form R_X is true, viz. the empty set \emptyset . So whether or not E has an explanation, there will be some explanatory rival (relative to E) that is true on this account. In other words, the set of explanatory rivals relative to E will always be collectively exhaustive on this account. Now, since any proposition of the form R_X is incompatible with any other such proposition, we also have that on this account explanatory rivals are pairwise incompatible. In sum, then, we have here an account that

argued, T is never the best explanation of T 's success, and so IBE could never warrant inferring T on the basis of its success. If Fine's argument is correct, then it would mean that any theory that goes beyond empirical adequacy – i.e. any theory for which $T \neq T_E$ – is unferrable by IBE, and thus that the Ultimate Argument is utterly impotent as an argument for scientific realism.

The reply to this objection that I now want to explore starts by pointing out that the objection presupposes that T and T_E are explanatory rivals, i.e. such that an IBE warrants inferring one just in case it is a better explanation than the other. But is this a claim that proponents of the Ultimate Argument are obliged to accept? Clearly not, since they may appeal to the account of explanatory rivalry presented at the end of the previous section where explanatory rivals are, among other things, guaranteed to be pairwise incompatible. Since T and T_E are compatible, they would not be explanatory rivals on this account. So the relative explanatory goodness of T and T_E would be *irrelevant* to whether T is inferable by IBE.

Moreover, and perhaps more importantly, the previous section also argued that explanatory rivals would have to be incompatible *on the most plausible version of IBE*. Since T and T_E are compatible hypotheses, they would not be explanatory rivals according to this requirement. The argument for this requirement was a general one that appealed to the fact that unless we require explanatory rivals to be pairwise incompatible, the inferability (by IBE) of one rival will entail that a compatible rival of arbitrary plausibility is not inferable (by IBE). This, I argued, is something proponents of IBE would do well to avoid, and thus that they should require explanatory rivals to be incompatible. Note that this argument appeals to a general consideration that has nothing in particular to do with scientific realism or the Ultimate Argument, so this response does not simply beg the question against Fine. On the contrary, the argument constitutes an independent reason to reject Fine's presupposition that T and T_E are explanatory rivals.

Now, one might ask what it would take for T to be inferable by IBE from T 's empirical success if explanatory rivals must be incompatible. Well, on the account of explanatory rivalry presented in the previous section, T would have

to follow from the best proposal for a *complete* explanation of T 's success. In effect, this is to say that on a more precise reconstruction, the Ultimate Argument goes as follows:²¹

(U1) Scientific theory T is empirically successful.

(U2') The best proposal for a complete explanation of T 's empirical success is or entails that T is true.

(U3) Thus, by IBE, (it is reasonable to believe that) T is true.

Now, one might think that a modified version of Fine's objection succeeds against this precisified version of the Ultimate Argument. Consider the following pair of propositions that *would* count as explanatory rivals on the account presented in section 4:

R_T : T is the complete explanation for T 's empirical success.

R_E : T_E is the complete explanation for T 's empirical success.

Of course, R_T is a proposal for a complete explanation for T 's success that entails that T is true (and thus counts as one of the proposals referred to in (U2')), whereas R_E is a proposal on which only T 's empirical adequacy is explanatorily relevant for T 's empirical success. Here we have what would (on my account) be a pair of explanatory rivals relative to T 's success. One might think Fine's objection can be resurrected in this way.

However, note that precisely because R_E and R_T are pairwise incompatible, R_E is not merely a weaker version of R_T . R_E effectively denies that anything beyond the empirical adequacy of T explains T 's success, and thus takes a stand on the explanation of T 's success that R_T does not. Specifically, R_E holds that the "explanatory buck" stops at T 's empirical adequacy, and thus that T 's empirical adequacy is a brute unexplainable fact. This is important because it means that R_T does not entail R_E , and thus that there is no "metatheorem"

ensures in a natural way that explanatory rivals form a partition of logical space.

²¹The only difference between this and the original formulation of the Ultimate Argument is that (U2) has been precisified in light of the account of explanatory rivalry presented in

according to which a rival of the R_E -form is necessarily a better explanation of T 's success than one of the R_T -form. R_T and R_E are distinct and incompatible claims about the (complete) explanation for T 's success, and thus neither can be proved to be better by virtue of being more modest than the other.

It is true, of course, that R_E is *in one way* more modest than R_T , since it doesn't posit any unobservable entities. But R_T is more modest *in another way*, since it doesn't posit an unexplainable regularity in T 's empirical adequacy. If R_E is to be a better (or at least as good an) explanation as R_T for any successful theory T then it would have to be that positing a brute regularity (i.e. a regularity that has no explanation) always explains a theory's success better (or at least as well) as positing the existence of some unobservable entities. Apart from being wildly implausible, we find no hint of an argument for such a claim in Fine's writings. Indeed, it seems contrary to at least the spirit (if not the letter) of Fine's (1996) own non-realist position – the Natural Ontological Attitude (NOA) – to positively assert that the success of a theory could not possibly be explained by reference to the unobservable entities it posits. After all, that is itself a claim – albeit a *negative* claim – about the unobservable aspects of the world and how they are related to what we actually observe, but Fine's non-realism is supposed to take a non-committal stance on issues of this sort.

I thus conclude that the Ultimate Argument can be saved from Fine's objection and related considerations, in that there is no proof that an explanatory rival that merely appeals to a theory's empirical adequacy is always at least as good an explanation as an explanatory rival that appeals to its truth. That's good news for realists who rely on the Ultimate Argument, of course. However, note that this does not mean that a rival that appeals to the truth of a theory is necessarily better than one that merely appeals to its empirical adequacy. What I have argued for is the weaker claim that an explanatory rival that appeals to a theory's empirical adequacy is not necessarily a better explanation than a rival that appeals to its truth. For all I have argued, it may be that *in some cases* the empirical adequacy of a theory provides the best complete explanation of its success, e.g. because one has some special reason to think

that there could be no explanation for something beyond a theory's empirical adequacy. Thus it is at least an open possibility that something like Fine's objection undermines an Ultimate Argument to the truth of *some* theories. Put differently, the Ultimate Argument may well support only a *piecemeal* realism on which only some empirically successful theories are taken to be true.

6 CONCLUSION

I have argued for three main claims in this paper: (i) First, I argued that on the most plausible version of Inference to the Best Explanation (IBE), explanatory rivals are mutually exclusive. I argued that imposing this structural requirement makes IBE more plausible than it would otherwise be. (ii) Partly by considering a potential counterexample to this requirement, I spelled out an account of explanatory rivalry that satisfies it in an effortless way. On this account, explanatory rivals are (roughly speaking) complete explanations of a given explanandum. (iii) Finally, I then used these results to criticize a prominent anti-realist objection to the so-called Ultimate Argument for scientific realism. In short, this objection turns out to presuppose a simplistic account of explanatory rivalry – one that, I argued, we have independent reasons to reject.

ACKNOWLEDGEMENTS

This paper was presented at the Hebrew University/UNC Workshop on Scientific Explanations in Chapel Hill, North Carolina, on May 21-22, 2014. I would like to thank the participants at the workshop for valuable feedback, especially my commentator Jarrett Leplin. I would also like to thank members of the Philosophy of Science Graduate Student Reading Group at UNC Chapel Hill, who read and commented on my paper at a very early stage. Finally, I would especially like to thank Marc Lange, Huginn Thorsteinsson, and two anonymous referees for this journal.

REFERENCES

- Bird, A. (2005). Abductive Knowledge and Holmesian Inference. In Gendler, T. S. and Hawthorne, J., editors, *Oxford Studies in Epistemology*, pages 1–31. Oxford University Press, Oxford.
- Boyd, R. (1980). Scientific Realism and Naturalistic Epistemology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, pages 613–662.
- Dellsén, F. (ms). Realism and the Absence of Rivals. Unpublished manuscript.
- Fine, A. (1986). Unnatural Attitudes: Realist and Instrumentalist attachments to Science. *Mind*, 95:149–179.
- Fine, A. (1991). Piecemeal Realism. *Philosophical Studies*, 61:79–96.
- Fine, A. (1996). *The Shaky Game*. University of Chicago Press, Chicago, IL, second edition edition.
- Harman, G. (1965). The Inference to the Best Explanation. *The Philosophical Review*, 74:88–95.
- Hempel, C. G. (1965). *Aspects of Scientific Explanation and Other Essays*, chapter Aspects of Scientific Explanation. Free Press, New York, NY.
- Hempel, C. G. and Oppenheim, P. (1948). Studies in the Logic of Explanation. *Philosophy of Science*, 15:135–147.
- Henderson, L. (2014). Bayesianism and Inference to the Best Explanation. *British Journal for the Philosophy of Science*, 65:687–715.
- Kitcher, P. (1981). Explanatory Unification. *Philosophy of Science*, 48:507–531.
- Laudan, L. (1981). A Confutation of Convergent Realism. *Philosophy of Science*, 48:19–49.

- Leplin, J. (1988). Surrealism. *Mind*, 97:519–524.
- Leplin, J. (1997). *A Novel Defense of Scientific Realism*. Oxford University Press, Oxford.
- Levin, M. (1984). What Kind of Explanation is Truth. In Leplin, J., editor, *Scientific Realism*, pages 124–139. University of California Press, Berkeley, CA.
- Lewis, D. (1986). Causal Explanation. In *Philosophical Papers, volume II*, pages 214–240. Oxford University Press, Oxford.
- Lipton, P. (1993). Is the Best Good Enough? *Proceedings of the Aristotelian Society*, 93:89–104.
- Lipton, P. (1994). Truth, Explanation, and the Best Explanation. In Derksen, A. A., editor, *The Scientific Realism of Rom Harré*. Tilburg University Press, Tilburg.
- Lipton, P. (2004). *Inference to the Best Explanation*. Routledge, London, 2nd edition.
- Lycan, W. G. (1985). Epistemic Value. *Synthese*, 64:137–164.
- Maxwell, G. (1962). The Ontological Status of Theoretical Entities. In Feigl, H. and Maxwell, G., editors, *Minnesota Studies in the Philosophy of Science*, volume 3, pages 3–15. Minnesota University Press.
- McGrew, T. (2003). Confirmation, Heuristics, and Explanatory Reasoning. *British Journal for the Philosophy of Science*, 54:553–567.
- McMullin, E. (1987). Explanatory Success and the Truth of Theory. In Rescher, N., editor, *Scientific Inquiry in Philosophical Perspective*. University Press of America, Lantham.
- Musgrave, A. (1988). The Ultimate Argument for Scientific Realism. In Nola, R., editor, *Relativism and Realism in Science*, pages 229–252. Kluwer Academic Publishers, Dordrecht, Boston.

- Okasha, S. (2000). Van Fraassen's Critique of Inference to the Best Explanation. *Studies in the History and Philosophy of Science*, 31:691–710.
- Poston, T. (2014). *Reason and Explanation: A Defense of Explanatory Coherentism*. Palgrave Macmillan, New York, NY.
- Psillos, S. (1999). *Scientific Realism: How Science Tracks Truth*. Routledge, London.
- Putnam, H. (1975). *Mathematics, Matter and Method*. Cambridge University Press, Cambridge.
- Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press, Princeton, NJ.
- Skow, B. (2014). Are There Non-Causal Explanations (of Particular Events). *British Journal for the Philosophy of Science*, 65:445–467.
- Smart, J. J. C. (1963). *Philosophy and Scientific Realism*. Routledge & Kegan Paul, London.
- Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press, Oxford.
- Thagard, P. R. (1978). The Best Explanation: Criteria for Theory Choice. *The Journal of Philosophy*, 75:76–92.
- van Fraassen, B. C. (1980). *The Scientific Image*. Clarendon, Oxford.
- van Fraassen, B. C. (1989). *Laws and Symmetry*. Clarendon Press, Oxford.
- Weisberg, J. (2009). Locating IBE in the Bayesian Framework. *Synthese*, 167:125–143.