Explanation and Evidence
(For Routledge Handbook of Evidence)
-Kevin McCain and Ted Poston-

Explanation and evidence are related in one way that is uncontroversial: evidence can confirm or disconfirm explanations. One explanation of Sally’s cold is that she has a virus; another is that she has a bacterial infection. The available evidence confirms the virus explanation because the evidence supports that colds are caused by viruses, not bacteria. A more interesting question concerns whether explanatory facts themselves provide evidence. That is to say, do we get evidence for \( p \) simply by realizing that \( p, \text{if true, would explain some data?} \) Explanationists answer this question affirmatively, maintaining that explanatoriness is evidence of truth (with various qualifications, of course).

What exactly is explanationism though? William Lycan (2002 p. 417) offers a helpful taxonomy of explanationist views:

- **Weak Explanationism:**
  explanatory inferences (inferences to the best explanation) from a given set of premises can epistemically justify a conclusion.

- **Sturdy Explanationism:**
  Weak Explanationism + explanatory inferences can justify conclusions without being derived from some other more basic form of ampliative inference.

- **Ferocious Explanationism:**
  Sturdy Explanationism + explanatory inference is the only basic form of ampliative inference.

- **Global Explanationism:**
  Ferocious Explanationism + all inferences and reasoning, including deductive, is derived from explanatory inference.

The common commitment is that explanatory inference is justifying. That is to say, there are instances where it is reasonable to infer \( p \)'s truth on the basis of the fact that \( p \) provides the best explanation of some data. This amounts to the view that \( p \)'s explaining the data provides evidence for \( p \)'s truth. Our focus is this common commitment to inference to the best explanation (IBE).

1. **IBE: A First Pass**

In simplest terms the idea behind inference to the best explanation (IBE) is that explanatory virtues are a guide to truth. In slogan form: “the best explanation is true” or “it is rational to believe the best explanation”. Such slogans can be convenient when thinking about IBE, but without qualification they are false. It is more appropriate to say, “the explanation that would, if true, provide the deepest understanding is the explanation that is likeliest to be true.” More schematically, we might understand IBE in this way:

1. \( F_1, F_2, \ldots, F_n \) are facts in need of explanation.
2. Hypothesis \( H \) explains the \( F_i \).
3. No available competing hypotheses would explain the \( F_i \) as well as \( H \) does.

---

1 Lycan terms this option “Holocaust Explanationism”. We prefer the more accurate label “Global Explanationism”.
2 Lipton (2004) p. 61
\[ \text{H is true.}^3 \]

Let us clarify a few points about this formulation of IBE. First, what does it mean to say that H explains a fact or set of facts? Often an explanation is thought to be something that provides understanding.\(^5\) So, when H explains F, H provides understanding of F. The most straightforward way to make sense of this idea is that H answers a “why-question”—why did F occur, why is F true, or why does F have the features that it does? \(^6\) Explanations construed as answers to why-questions can be understood both causally and non-causally. Although there is a vast literature on the nature of explanation, for present purposes the intuitive idea that an explanation involves answering various why-questions should suffice.\(^7\)

Second, the fact that a hypothesis would explain some facts does not entail that the hypothesis is the actual explanation of those facts. Instead, it means that the hypothesis is a potential explanation of those facts. In other words, the hypothesis, if true, would provide understanding of the facts. Hence, IBE says that we should infer that the best potential explanation is the actual explanation.

Third, (3) is where the “best” part of inference to the best explanation comes into play. Claiming that no available hypothesis explains the facts as well as a given hypothesis means that the superior hypothesis is the best available explanation. Many things have been identified as explanatory virtues: various kinds of simplicity, explanatory power (the range of phenomena explained and/or how illuminating the explanation is), coherence with background information, non-ad hocness, conservatism, and raising fewer unanswered questions are among the more commonly discussed virtues.\(^8\) The hypothesis with the best complement of explanatory virtues is the best explanation. And so, this hypothesis is the explanation that, if true, would provide the most understanding of the facts to be explained.

2. The case for IBE

Let us review the case for thinking that explanatory reasoning provides evidence. One core intuition and three intersecting lines of argument support the goodness of explanatory reasoning. The core intuition is that we have some inductively justified beliefs. Consider the belief that the mail has been delivered today. We have lots of good evidence that the mail is delivered on a daily basis, but this evidence is consistent with the possibility that the mail was not delivered today. Sometimes odd things happen. Even so, we are justified in believing that the mail has been delivered today.

Given the correctness of this core intuition, the philosophical challenge is to explain how inductive justification works. In this context there are three intersecting arguments that support IBE. The first argument may be called The Ubiquity of Explanatory Reasoning. The basic idea is that explanatory reasoning is commonplace, found across a spectrum of contexts.\(^9\) Moreover,
explanatory reasoning occurs quite early on in cognitive development. Here’s a simple example. A fourth-grader observes that her jacket is not in its cubby hole. Her teacher asks her where her jacket is. She answers “I must have left it at home.” This is an example of explanatory reasoning. The fourth-grader infers that the jacket is at home from the evidence that the jacket is not in the cubby hole. If this were the only thing that was going on, it would not be a good inference. But this is not the case. The fourth-grader is sensitive to some explanatory dimensions in this case. Namely, that the hypothesis that she left it at home explains better than a range of alternative hypotheses.

The second argument for IBE comes from the history of science. While the first argument appeals to the commonplace use of explanatory reasoning, this argument appeals to its sophisticated use in the sciences. Darwin (1872/1962: 476) in the recapitulation of his argument in *Origins* writes:

> It can hardly be supposed that a false theory would explain, in so satisfactory a manner as does the theory of natural selection, the several large classes of facts above specified. It has recently been objected that this is an unsafe method of arguing; but it is a method used in judging of the common events of life, and has often been used by the greatest natural philosophers.

Darwin appeals to the explanatory power of his case for evolution and notes as well that it is method of arguing that has wide application.

Darwin is not alone in his use of explanatory reasoning. The independent discoveries of Neptune by John Couch Adams and Urbain Leverrier were the result of explanatory reasoning, as was Joseph John Thomson’s discovery of the electron. The Ptolemaic model of the solar system was replaced by heliocentric model for explanatory reasons.\(^{10}\) Examples of this sort are easily multiplied.\(^{11}\) As Clark Glymour (1984: 173) has rightly noted, “One can find such arguments [inferences to the best explanation] in sociology, in psychometrics, in chemistry and astronomy, in the time of Copernicus, and in the most recent of our scientific journals.”

The final argument for IBE is an argument against enumerative induction. Some philosophers maintain that IBE works only if grounded in enumerative induction.\(^{12}\) Enumerative induction works by listing instances of a general claim and then, given a sufficient number of those instances, inferring the truth of the general claim. For instance, consider the generalization that *all ravens are black*. This generalization is not deductively implied by any body of observable evidence because it is possible that there are unobserved ravens (perhaps in the future) that are not black. The method of enumerative induction, though, licenses inference to the generalization when enough instances are surveyed.

There are several standing problems with enumerative induction. One is that there is no means to determine what counts as a “sufficient number of instances.” This problem, however, isn’t as serious as the other two problems. A second problem is that there must be some reason to think that the instances are representative of the population. Suppose we have a million positive instances, a large number on reasonable view. But suppose further that all these instances are only taken from one of several relevant populations. In such a case the positive instances do not support

---

\(^{10}\) See Douven (2017a) for further discussion of these.

\(^{11}\) See Kampourakis (2014) and Trout (2016) for several more examples. Trout goes so far as to argue that not only does explanatory reasoning play a significant role in particular scientific discoveries, but it has driven the success of modern science as a whole.

\(^{12}\) See, for example, Fumerton (1980).
the generalization because there is a viable alternative explanation that the positive instances only reflect this particular subgroup.

A third problem is that some generalizations are not of the right kind to receive support from instances. Consider the predicate *grue* which means that “x is green and observed prior to 2025 or x is blue.” The generalization *all emeralds are grue* has a very impressive record. All the instances observed so far fit with the claim that *emeralds are grue*. Yet the instances do not support this; rather the instances support that *emeralds are green*. The discussion of the issues here are rich but one common thread is that enumerative induction cannot to the work needed for a theory of inductive confirmation, but explanatory reasoning can do the work.\(^\text{13}\)

2. **Bad Lots and Probable Disjunctions**

Although IBE is intuitively plausible, and it is widely employed in science and everyday life, a number of objections have been pressed against it. One of the most prominent objections is Bas van Fraassen’s (1989: 143) “Best of a Bad Lot”:

> To believe is at least to consider more likely to be true, than not. So to believe the best explanation requires more than an evaluation of the given hypothesis. It requires a step beyond the comparative judgment that this hypothesis is better than its actual rivals…For me to take it that the best of set \(X\) will be more likely to be true than not, requires a prior belief that the truth is already more likely to be found in \(X\), than not.

The concern is that when we infer the truth of a hypothesis because it is the best explanation “our selection may well be the best of a bad lot.” Hence, being the best explanation of a set of competing explanations does not warrant an inference the truth of a hypothesis since the truth may not be among the limited number of hypotheses that are available.

A related objection is what McCain and Poston (forthcoming) have termed the “Disjunction Objection”\(^\text{14}\). Like the Best of a Bad Lot, this objection highlights the fact that being the best available explanation does not seem sufficient to warrant inference. However, the two objections are distinct. The heart of the Best of a Bad Lot objection is the concern that the true hypothesis may not be among the set of available hypotheses. However, the Disjunction Objection can arise even if it is certain that the true hypothesis is among the hypotheses considered. Grant that IBE is correct in the sense that the best explanation is the most likely explanation to be true. Does it follow that the best explanation is more likely than not to be true? It does not. Consider a situation where there are only three possible hypotheses: H1, H2, and H3. In such a case we know that one of these three is correct, hence there is no worry about a bad lot. Nevertheless, should we infer that the best explanation among the three is true? It does not seem so because it could be that while H1 is the best explanation, and so more likely than H2 and more likely than H3, it is not more likely than *either H2 or H3*. For example, suppose we know that either Harry, Ron, or Hermione missed potions class. Given their dispositions and relevant facts the best explanation is that Harry missed class. But even so the competing explanations have some non-trivial merit, and so it’s more likely that either Ron or Hermione missed class. Thus, according to Disjunction Objection, IBE would license

\(^{13}\) For more on how explanatory reasoning can solve the first two problems for enumerative induction see Harman (1965), Lycan (1988), and McCain (2016). See Poston (2014) discussion of how IBE provides a solution to the New Riddle of Induction, i.e. solves the grue problem.

\(^{14}\) See also Fumerton (1995) and van Fraassen (1989) for brief discussions of this objection.
inference to a claim that is likely to be false.

3. IBE: Reformulated

We have seen two serious challenges for the sufficiency of being the best explanation warranting inference. Even though both the Best of a Bad Lot and the Disjunction Objection pose challenges to IBE\(^\text{15}\), they can be avoided by thinking more clearly about how explanatory reasoning works. Rather than abandoning IBE, perhaps the best route is to refine our earlier formulation to the following:

1. F\(_1\), F\(_2\), …, F\(_n\) are facts in need of explanation.
2. Hypothesis H explains the F\(_i\).
3. No available competing hypotheses would explain the F\(_i\) as well as H does.
4. H is a good explanation in light of our background information.

\[ \therefore \text{H is true.}\] \(^\text{16}\)

The key insight to this formulation of IBE is that explanatory reasoning warrants inference to a conclusion only if the explanatory hypothesis is sufficiently good.\(^\text{17}\) Explanatory reasoning involves two stages. First, there is a selection of plausible hypotheses that offer explanations of the relevant data. These hypotheses fit with the background evidence and offer incompatible explanations of the data. In actual applications of IBE in both mundane and sophisticated contexts this stage tracks the formulation of plausible hypotheses. Outlandish hypotheses are dismissed out of hand, but nonetheless dismissed on explanatory grounds. Second, the plausible hypotheses are then assessed in terms of their explanatory merits vis-à-vis the particular data. McCain and Poston (forthcoming) put the point this way: “IBE is a multi-track inference. Or, perhaps more simply, we must realize that IBE says to infer the truth of the best explanation only if that explanation is sufficiently good in light of the background evidence. Proper appreciation of the role of background evidence in IBE dispels the Disjunction Objection.” Plausibly, this sort of restriction also blocks the Best of a Bad Lot objection. If there is a restriction on how good a hypothesis must be in order to be inferred, even if it is the best explanation, it seems that a hypothesis inferred to be true will not be simply part of a bad lot. As a result, understanding IBE as limited to hypotheses that are sufficiently good appears to save a plausible and widespread inference rule from serious objections.

4. Indifference, Truth, and Bayesianism

We’ve seen that appreciating the role of background information in explanatory reasoning leads to a formulation that avoids two significant objections to the acceptability of inferences of this sort. Unfortunately, this is not enough to put IBE in the clear. In this section we will briefly discuss three additional challenges confronting IBE that arise even when it properly restricts inferred hypotheses to those that are sufficiently good in the light of background information.

\(^{15}\) Though see Schupbach (2014) for reasons to think that the Best of a Bad Lot objection at least fails to pose a genuine problem.

\(^{16}\) This formulation is similar to McCain & Poston’s (forthcoming).

\(^{17}\) See Lipton (2004) and Musgrave (1988) for more on the importance of restricting IBE in this way.
4.1 Unconceived alternatives

The first of these three challenges is the problem of unconceived alternatives. Here’s how van Fraassen puts this challenge:

I believe, and so do you, that there are many theories, perhaps never yet formulated but in accordance with all evidence so far, which explain at least as well as the best we have now. Since these theories can disagree in so many ways about statements that go beyond our evidence to date, it is clear that most of them by far must be false. I know nothing about our best explanation, relevant to its truth-value, except that it belongs to this class. So I must treat it as a random member of this class, most of which is false. Hence it must seem very improbable to me that it is true (1989: 146).

Thus, van Fraassen claims that the fact that H is the best available explanation of some facts does not justify us in thinking that H is true because H is simply one of a very large number of mostly false hypotheses (most of which we have not thought of yet) that explain the relevant facts equally well.

Some have argued against van Fraassen on this point by maintaining that the way in which we come to form hypotheses, at least in the empirical sciences, precludes our best explanations from being merely members of a large set of mostly false hypotheses even if there are unconceived alternatives to our best explanations. They point out that we have a large amount of background information that goes into our theory formation practices and that this background information increases the likelihood that our best explanations are correct. It may be that appeal to background information again protects IBE.

4.2 Truth

Another challenge that remains for IBE concerns the relationship between explanatory virtues and truth. Peter Lipton (2004, p. 144) has termed the objection that can arise from this challenge “Voltaire’s Objection” (this has also been called the “Truth Demand”):

Why should the explanation that would provide the most understanding if it were true be the explanation that is most likely to be true? Why should we live in the loveliest of all possible worlds? Voltaire’s objection is that, while loveliness may be as objective as you like, the coincidence of loveliness and likeliness is too good to be true. It would be a miracle if using explanatory considerations as a guide to inference were reliably to take us to the truth.

The worry here is that explanatory considerations are not be linked to the truth. And, if explanatory considerations in general are not linked to the truth, then it is not rational to infer that the hypothesis that best explains a given body of evidence is true.

There have been attempts to defend the truth connectedness of specific explanatory virtues, such as simplicity, as well as attempts to defend the idea that explanatory considerations more

---

19 For further discussion of this problem, particularly its application to the debate over scientific realism, see Devitt (2011) and Stanford (2006).
20 See McCain (2016) for discussion.
generally are indicative of the truth. However, it is a matter of controversy how successful such attempts are.

The challenge to show that a method of inference is truth-conducive is not unique to IBE. Suppose, for instance, that a challenge arises as to whether memory is reliable. One way of meeting this challenge is to argue that memory is an instance of another form of inference, say schmemory, and schmemory is reliable. The problem with this move is that memory seems to be a basic source of belief, i.e. it is not derived from any other source. This leads to the other way to meet this challenge: argue that memory is a basic form of inference and that basic forms of inference do not require non-circular justifications. Applying this in the context of explanatory inference then, the idea is that if explanatory inference is not dependent on another form of justifying inference, it may still be good even if we are unable to justify it in a non-circular fashion. It may be that given the ubiquity of explanatory reasoning if it must be non-circularly justified in order to justify, then skepticism follows.

4.3 Bayesianism

The third challenge remaining for IBE concerns how to reconcile it with Bayesianism. Some have argued that IBE is incompatible with Bayesianism, and consequently IBE should be abandoned. Others argue that IBE and Bayesianism are compatible. And, yet others have argued that although IBE and Bayesianism are incompatible, this does not pose a problem for IBE because it is a reliable inference method.

A crucial issue in this debate is whether explanatory reasoning is good independently of grounding in some other kind of reasoning. In other words, should we accept “sturdy explanationism” in addition to “weak explanationism”? While there is considerable evidence that people use IBE in their actual reasoning, one might think that the normative work of correct reasoning comes from Bayesianism. Hence, one might think that while weak explanationism is plausible, sturdy explanationism is not because IBE is derived from a more fundamental sort of ampliative inference, e.g., Bayesian inference. Consequently, some hold that IBE is, at best, a heuristic for realizing Bayesian inference because explanatoriness is evidently irrelevant once the appropriate probabilities are taken into consideration. Some philosophers respond that the goodness of IBE is independent of any grounding in Bayesian inference. In fact, some go so far as

21 See Swinburne (1997) and White (2005) for defenses of the idea that simplicity is a sign of truth. See Bird (2017), McCain (2018b), and McCain & Poston (2014), (2017) for reasons to think that explanatory virtues in general are relevant to evidence for a proposition’s truth.
22 Sober (2015) and Roche (2018) put forward important challenges to the idea that simplicity is evidence of truth. Rinard (2017) and Roche & Sober (2013), (2014), (2017) argue that explanatory virtues are unrelated to the likelihood of truth.
23 See McCain (2016) for more on why the justification of IBE may be rule-circular without thereby being problematic. See Poston (2011) for a related discussion of how explanationism addresses circularity concerns in the context of the problem of the criterion. For a classic discussion of the difference between rule-circularity and premise-circularity see Van Cleve (1984).
24 See, for example, van Fraassen’s (1989) Dutch book argument against IBE. For discussion of this argument see Christensen (2004) and Douven (2013).
28 See Roche & Sober (2013), (2014), (2017) for arguments for the conclusion that explanatory considerations are irrelevant to confirmation. For responses on behalf of IBE see Climenhaga (2017) and McCain & Poston (2014), (2017).
to say that the grounding relation goes in the other direction—Bayesian inference is grounded in IBE. One line of argument for this is that Bayesian inference lands in skepticism if it isn’t grounded in explanatory inference. The idea is that Bayesian inference requires an initial probability function that supports inductive learning, and explanatory inference provides such a probability function.²⁹ The debate over this issue is unsettled, and there remains work to be done in order to establish that being the best explanation is non-derivative evidence.

5. Beyond IBE

We have seen that despite the plausibility of explanatory reasoning, sturdy explanationism is controversial. Weak explanationism, without sturdy explanation, is controversial itself since it requires that explanatory reasoning is grounded in a deeper form of inductive reasoning. That being said, before closing it is worth briefly taking a look at some of the challenges facing either ferocious or global explanationism. Explanationists of these stripes understand inductive justification or justification simpliciter in terms of explanatory considerations.³⁰ Let us focus on global explanationism. On this view justification is a matter of explanatory relations between one’s evidence and a proposition. So, roughly, S is justified in believing that when p is part of the best explanation of S’s evidence. Explanationism of this sort is initially plausible and worth serious consideration.

Despite the plausibility of explanationism as a theory of epistemic justification, it faces some significant challenges of its own. One challenge concerns beliefs about the future. It seems intuitive that at least some beliefs about the future are justified, but some argue that it is difficult to see how explanationism can adequately account for this justification.³¹ Consider, for instance, the evidence that supports the justified belief that it will snow before Christmas. The objection is that the claim it will snow before Christmas does not explain the evidence because the evidence is temporally prior to the putative fact that it snows before Christmas.

Another challenge concerns the explananda that explanationism relies on to generate justification. As Richard Fumerton (2018: 337-338)³² puts the concern:

Part of the problem is that one isn’t even in a position to seek explanations until one has found oneself with a potential explanandum—something that needs explaining. If the proposition that I’m in pain is an explanation, what does it explain? Is it supposed to be explaining the fact that I believe that I’m in pain, or the fact that it seems to me that I’m in pain, or the fact that I’m bleeding profusely and screaming profanities at the top of my lungs? All this seems wrong—almost comically wrong.

Clearly, in many cases of introspection (whether it is a pain or some other mental state) we have justification for believing that we are in a particular mental state/having a particular experience.

---

²⁹ For example Huemer (2009) and Poston (2014) argue that in order to avoid skeptical problems explanatory reasoning is needed to establish initial priors. Similarly, McCain & Poston (2014) maintain that IBE is needed in order to license projecting from observed to unobserved cases.

³⁰ See Harman (1973), (1986) and Lycan (1988) for defenses of these forms of explanationism. More recently, see McCain (2014a) and Poston (2014).


³² Also see Goldman (2011) for presentation of this challenge.
However, it may be difficult for an explanationist to say exactly what our introspective beliefs are explaining.\textsuperscript{33}

Finally, one might worry that explanationists have things backwards. That is, one might think that explanatory considerations do not determine evidential support, but rather, evidential support determines which explanation is the best.\textsuperscript{34} Of course, if this is correct explanationism fails to provide an account of epistemic justification. So, explanationists face a challenge of explaining why their picture captures the correct direction of evidential fit. That is, they face the challenge of showing that it is explanatory considerations that make it so that \( p \) fits the evidence rather than \( p \)'s fitting the evidence making it so that \( p \) is the best explanation.\textsuperscript{35}

Like weak and sturdy explanationism, ferocious and global explanationism faces numerous challenges. That should not lead us to think that explanationism in any of its varieties is to be abandoned. Perhaps the best explanation of the controversy surrounding explanationism is that although a hypothesis’s being the best explanation is evidence of the hypothesis’s truth, there is much work left to be done when it comes to the details.

References


---

\textsuperscript{33} See McCain (2018a) for a response to Fumerton on this point.

\textsuperscript{34} See Steup (2018).

\textsuperscript{35} See McCain (2018a) for a response to Steup on this point.


Roche, W. & Sober, E. 2013. Explanatoriness is evidentially irrelevant, or inference to the best explanation meets Bayesian confirmation theory. Analysis, 73: 659-668.


