G.Hon and S.S. Rakover (eds.), *Explanation: Theoretical Approaches and Applications*, 93-120 (Kluwer, 2001).

IS EXPLANATION A GUIDE TO INFERENCE? A REPLY TO WESLEY C. SALMON

PETER LIPTON
Cambridge University

Introduction

Earlier in this volume, Wesley Salmon has given a characteristically clear and trenchant critique of the account of non-demonstrative reasoning known by the slogan `Inference to the Best Explanation'. As a long-time fan of the idea that explanatory considerations are a guide to inference, I was delighted by the suggestion that Wes and I might work together on a discussion of the issues. In the event, this project has exceeded my high expectations, for in addition to the intellectual gain that comes from the careful study of his essay, I have benefited enormously from the stream of illuminating emails and faxes that Wes has sent me during our collaboration. Doing philosophy together has been an education and a pleasure.

Salmon's essay would place Inference to the Best Explanation beyond the pale of acceptable philosophical accounts of inference. According to Salmon, Inference to the Best Explanation has serious internal difficulties and compares very unfavourably with Bayesian approaches to these matters. My aim in the following remarks is irenic. I hope to show that a number of the claimed difficulties either are not really difficulties or are avoidable. In some cases, the avoidance will require a mild re-interpretation of the account that lies behind the slogan `Inference to the Best Explanation'; in others, it will require admitting limits to the scope of the account. For I accept at the outset that Inference to the Best Explanation cannot possibly be the whole story about the assessment of scientific hypotheses. For me, the interesting idea is simply that we sometimes decide how likely a hypothesis is to be correct in part by considering how good an explanation it would provide, if it were correct. This is the idea of explanatory considerations providing a guide to inference, and this is the idea that I will here promote.

As I read him, Salmon's most central objection is that the preference for one hypothesis over another is not based, as Inference to the Best Explanation would have it, on a judgement of which hypothesis would be most explanatory -- what I call the `loveliest' explanation -- but rather on a judgement of which hypothesis is likeliest to be correct, where this judgement of likeliness or probability is determined on Bayesian grounds. This is a fundamental objection. One could still maintain that what is inferred is the best explanation, where `best' just means likeliest, but this would in my view take away almost all the interest in the explanationist programme. It is hardly surprising that what we infer are often explanations, since our search is often a search for explanations. And given that we are looking for an explanation, neither is it surprising that we prefer ones with high probability. The exciting idea behind the slogan is that we use explanatory considerations as a guide to judgements of likeliness. This is what Salmon's objection casts into doubt, and the basis of this scepticism is his view that our guide to likeliness is to be found elsewhere, in Bayes' Theorem. (A list of `Some Key Concepts' at the end of this essay may help to keep track of some of the attributes of explanation and of inference as they fly by in the main text.)

Any proper articulation and defence of Inference to the Best Explanation must meet the `guiding challenge'; the challenge of showing that loveliness is a guide to likeliness. I cannot attempt this in any detail here, but I do have a strategy for defusing Salmon's objection. Central to this strategy will be arguments that explanationism and Bayesianism may be compatible and indeed complementary, because of the role that explanatory considerations might play in the actual mechanism by which enquirers `realise' Bayesian reasoning. As we will see, explanatory considerations may help enquirers to determine prior probabilities, to move from prior to posterior probabilities, and to determine which data are relevant to the hypothesis under investigation. Explanatory considerations may also play a crucial role in scientists' expression of their preference for hypotheses that promise to be fertile, that is to explain phenomena in addition to those directly under scrutiny at the time of inference. I will not succeed here in positively establishing a major role for explanatory considerations in Bayesian inference, but I aim to show that such a role is possible and even plausible, and in so doing promote the philosophical project of properly articulating the relationship between these two approaches to inference. Perhaps this will move Salmon to allow Inference to the Best Explanation back within the pale of philosophical accounts of inference worth developing. Having recently seen a magnificent production of the musical `Oklahoma!', I am moved to sum up my aim by saying that the Bayesian and the Explanationist should be friends.

The Idea of Inference to the Best Explanation

Science depends on judgments of the bearing of evidence on theory. Scientists must judge whether an observation or the result of an experiment supports, disconfirms, or is simply irrelevant to a given hypothesis. Similarly, scientists may judge that, given all the available evidence, a hypothesis ought to be accepted as correct or nearly so, rejected as false, or neither. Occasionally, these evidential judgments can be made on deductive grounds. If an experimental result strictly contradicts a hypothesis, then the truth of the evidence deductively entails the falsity of the hypothesis. In the great majority of cases, however, the connection between evidence and hypothesis is non-demonstrative or inductive. In particular, this is so whenever a general hypothesis is inferred to be correct on the basis of the available data, since the truth of the data will not deductively entail the truth of the hypothesis. It always remains possible that the hypothesis is false even though the data are correct.

One of the central aims of the philosophy of science is to give a principled account of these judgements and inferences connecting evidence to theory. In the deductive case, this project is well-advanced, thanks to a productive stream of research into the structure of deductive argument that stretches back to antiquity. The same cannot be said for inductive inferences. Although some of the central problems were presented incisively by David Hume in the eighteenth century (1777, Sec. 4), our current understanding of inductive reasoning remains remarkably poor, in spite of the intense efforts of numerous epistemologists and philosophers of science.

The model of Inference to the Best Explanation is designed to give a partial account of many inductive inferences, both in science and in ordinary life. One version of the model was developed under the name `abduction' by Charles Sanders Pierce early in this century (1931, 5.180-5.212, esp. 5.189), and the model has been considerably developed and discussed over the last twenty-five years (e.g. Harman, 1965; Thagard, 1978, Lipton 1991). Its governing idea is that explanatory considerations are a guide to inference, that scientists infer from the available evidence to the hypothesis which would, if correct, best explain that evidence. Many inferences are naturally described in this way. Darwin inferred the hypothesis of natural selection because, although it was not entailed by his biological evidence, natural selection would provide the best explanation of that evidence. When an astronomer infers that a galaxy is receding from the earth with a specified velocity, she does this because the recession would be the best explanation of the observed red-shift of the galaxy's spectrum. When a detective infers that it was Moriarty who committed the crime, he does so because this hypothesis would best explain the fingerprints, blood stains and other forensic evidence. Sherlock Holmes to the contrary, this is not a matter of deduction. The evidence will not entail that Moriarty is to blame, since it always remains possible that someone else was the perpetrator. Nevertheless, Holmes is right to make his inference, since Moriarty's guilt would provide a better explanation of the evidence than would anyone else's.

Inference to the Best Explanation can be seen as an extension of the idea of `self-evidencing' explanations, where the phenomenon that is explained in turn provides an essential part of the reason for believing that the explanation is correct. The galaxy's speed of recession explains why its spectrum is red-shifted by a specified amount, but the observed red-shift may be an essential part of the reason the astronomer has for believing that the galaxy is receding at that speed. Self-evidencing explanations exhibit a curious circularity, but this circularity is benign. The recession is used to explain the red-shift and the red-shift is used to confirm the recession; this reciprocal relationship may leave the recession hypothesis both explanatory and well-supported. According to Inference to the Best Explanation, this is a common situation in science: hypotheses are supported by the very observations they are supposed to explain. Moreover, on this model, the observations support the hypothesis precisely because it would explain them. Inference to the Best Explanation thus partially inverts an otherwise natural view of the relationship between inference and explanation. According to that natural view, inference is prior to explanation. First the scientist must decide which hypotheses to accept; then, when called upon to explain some observation, she will draw from her pool of accepted hypotheses. According to Inference to the Best Explanation, by contrast, it is only by asking how well various hypotheses would explain the available evidence that she can determine which hypotheses merit acceptance. In this sense, Inference to the Best Explanation has it that explanation is prior to inference.

Salmon characterises the Bayesian approach to inference that he favours as an instance of the natural view, where inference is prior to explanation. Thus from the Bayesian standpoint, Inference to the Best Explanation gets things backwards (29). (Page numbers refer to Salmon's essay in this volume.) It is not clear to me, however, that this need be so. To see why, we must say a bit more about how the slogan `Inference to the Best Explanation' is best understood.

The first distinction we need is between actual and potential explanation, where a potential explanation is something that satisfies all the conditions on actual explanation, with the possible exception of truth. Thus all actual explanations are potential explanations, but not conversely. Stories of alien abduction might explain certain observations -- to that extent they are potential explanations, but they are not actual explanations because they are not true. (My assumption that actual explanations must be true is not uncontroversial: cf. e.g. Cartwright 1983, esp. chs. 2 & 8.) According to Inference to the Best Explanation, we infer that what would best explain our evidence is likely to be true, that is, that the best potential explanation is likely to be an actual explanation. So at the level of actual explanation, there is no conflict of direction between Inference to the Best Explanation and Bayesianism: on both views we must first work out what is the likeliest explanation before we offer an account as an actual explanation. Nor is it clear to me that there need be a conflict at the level of potential explanation. The distinctive claim of Inference to the Best Explanation is that we use judgements of the quality of potential explanations as a guide to likeliness or probability. Although this will require further discussion below, this seems to be closely related to the Bayesian transition from prior to posterior probability. And if considerations of explanatory quality are in fact one way that inquirers make that transition, then the two approaches to inference may be seen to flow in the same direction.

Another useful distinction concerns the two different sorts of problem that an account of induction in science might purport to solve. The problem of description is to give an account of the principles that govern the way scientists weigh evidence and make inferences. The problem of justification is to show that those principles are sound or rational, for example by showing that they tend to lead scientists to accept hypotheses that are true and to reject those that are false. One popular application of Inference to the Best Explanation, for example, has been the attempt to mount a philosophical inference to the best explanation in aid of scientific realism, arguing that the truth of certain scientific theories, and so the reliability of scientific methods, would be the best explanation of their predictive successes. But while Inference to the Best Explanation has been applied both to descriptive and to justificatory problems, both Salmon and I focus here primarily on the former. The central issue for us is not whether inferences governed in part by explanatory considerations would be a good way to think, but whether scientists do think that way.

The difficulties of the descriptive problem are sometimes underrated, because it is supposed that inductive reasoning follows a simple pattern of extrapolation, with `More of the Same' as its fundamental principle. Thus we predict that the sun will rise tomorrow because it has risen every day in the past, or that all ravens are black because all observed ravens are black. This picture of enumerative induction has however been shown to be strikingly inadequate as an account of inference in science. On the one hand, a series of formal arguments, most notably the raven paradox and the new riddle of induction, have shown that the enumerative model is wildly over-permissive, treating virtually any observation as if it were evidence for any hypothesis (Hempel 1965, ch. 1; Goodman 1983, ch. 3). On the other hand, the model is also much too restrictive to account for most scientific inferences. Scientific hypotheses typically appeal to entities and processes not mentioned in the evidence that supports them and often themselves unobservable and not merely unobserved, so the principle of More of the Same does not apply. For example, while the enumerative model might account for the inference that a scientist makes from the observation that the light from one galaxy is red-shifted to the conclusion that the light from another galaxy will be red-shifted as well, it will not account for the inference from observed red-shift to unobserved recession.

The best-known attempt to account for these `vertical' inferences that scientists make from observations to hypotheses about often unobservable entities and processes is the Hypothetico-Deductive model. According to this model, scientists deduce predictions from a hypothesis (along with various other auxiliary premises) and then determine whether those predictions are correct. If some of them are not, the hypothesis is disconfirmed; if all of them are, the hypothesis is confirmed and may eventually be inferred. Unfortunately, while this model does make room for vertical inferences, it remains (like the enumerative model) far too permissive, counting data as confirming a hypothesis which are in fact totally irrelevant to it. For example, since a hypothesis (H) entails the disjunction of itself and any prediction whatever (H or P), and the truth of the prediction establishes the truth of the disjunction (since P also entails (H or P)), any successful prediction will count as confirming any hypothesis, even if P is the prediction that the sun will rise tomorrow and H the hypothesis that all ravens are black.

What is wanted is thus an account that permits vertical inference without permitting absolutely everything, and Inference to the Best Explanation promises to fill that bill. Inference to the Best Explanation sanctions vertical inferences, because an explanation of some observed phenomenon may appeal to entities and processes not themselves observed; but it does not sanction just any vertical inference, since obviously a particular scientific hypothesis would not, if true, explain just any observation. A hypothesis about raven coloration will not, for example, explain why the sun rises tomorrow. Moreover, Inference to the Best Explanation discriminates between different hypotheses all of which would explain the evidence, since the model only sanctions an inference to the hypothesis which would best explain it.

Plugging Accounts of Explanation into an Account of Inference

Inference to the Best Explanation thus has the advantages of giving a natural account of many inferences and of avoiding some of the limitations and excesses of other familiar accounts of non-demonstrative inference. If, however, it is to provide a serious model of induction, Inference to the Best Explanation needs to be developed and articulated, and this has not proven an easy thing to do. More needs to be said, for example, about the conditions under which a hypothesis explains an observation. Explanation is itself a major research topic in the philosophy of science, but the standard models of explanation yield disappointing results when they are plugged into Inference to the Best Explanation. For example, the best-known account of scientific explanation is the Deductive-Nomological model, according to which an event is explained when its description can be deduced from a set of premises that essentially includes at least one law. This model has many flaws. Moreover, it is virtually isomorphic to the Hypothetico-Deductive model of confirmation, so it would disappointingly reduce Inference to the Best Explanation to a version of hypothetico-deductivism. (Another `covering-law' model, Inductive-Statistical explanation (Hempel 1965, 381-393), would however yield something closer to the more promising Bayesian approach to inference of the sort that Salmon champions.)

Salmon and I thus speak in one voice in rejecting the Deductive-Nomological model of explanation on its own terms and finding it unsuitable in the context of Inference to the Best Explanation (5-8). For me, however, that dual failure is vaguely comforting, because many of the flaws of the Deductive-Nomological model of explanation and those of the Hypothetico-Deductive model of confirmation seem to `line up'. Here is one example.

Applied to the explanations of laws, the Deductive-Nomological model notoriously counts as explanatory the worthless deduction of a law from the conjunction of itself and another unrelated law. (This is a difficulty for the Deductive-Nomological model that Hempel (1965, 273 n33) acknowledged in his early work on explanation.) There is a parallel problem for the Hypothetico-Deductive model of confirmation, since that model wrongly allows that a conjunct confirms an arbitrary conjunction. The encouraging general thought is that the *actual* explanatory relation (which the Deductive-Nomological model fails to capture) may correspond well enough with the *actual* relation of inductive support to underwrite the explanationist programme. As Salmon emphasises, however, there is a great controversy over how to characterise the explanatory relation, and he goes on to suggest that without an adequate account of explanation, Inference to the Best Explanation has no clear meaning (10).

I am not convinced that this last point follows. Whether or not explanatory considerations are a guide to inference does not depend on whether we have an adequate account of explanation, any more than our use of a grammar to understand our language depends on our ability to give an adequate explicit account of the structure of that grammar. Moreover, if in fact we do use explanatory considerations as a guide to inference, to say that we do so seems to me not a meaningless or even a trivial claim, even in absence of an account of explanation, because we have some semantic grip on the concept of explanation in the absence of such an account. (Indeed if we did not, the project of generating the account would not get off the ground.) At the same time, I agree entirely that a really satisfying version of Inference to the Best Explanation would have to provide an articulated account of explanation, and that this is no easy thing to do.

The most important requirement on an articulated account of explanation suitable for insertion into Inference to the Best Explanation is the most obvious, but also the most difficult to meet. The account should be correct: it should capture our actual explanatory practices. If it does not do this, the resulting account of inference will not really be inference to the best explanation, but inference to something else. Are there any further constraints? What I have in mind is the question of whether there are certain general features an account of explanation might have which would make it unsuitable for use in an account of inference, even if it were a correct and otherwise legitimate account of explanation. Two features of this sort come to mind. One suggests that certain `epistemic' accounts of explanation may be unsuitable; the other casts doubt on the suitability of certain `ontic' accounts. In both cases, the worry is that, however attractive such accounts may be on their own, they would produce vicious circles if plugged into an account of inference.

The first worry is that an account of explanation might itself be framed in inferential terms, so that plugging it into Inference to the Best Explanation results in a pernicious circularity. Consider for example an account according to which to explain is to provide the reason to believe the explanandum, the description of the phenomenon to be explained. Plugging this into Inference to the Best Explanation seems to leave us in the uncomfortable situation where reasons are running simultaneously in both directions. If H is our inferred explanation, and E is what is explained, then Inference to the Best Explanation says that E is our reason for believing H, while the embedded account of explanation says that H is our reason for believing E. This does seem pernicious, quite unlike the self-evidencing explanations mentioned above. There is nothing vicious in itself about having E be our reason for believing H while H explains E; but things do seem illicit if we insist that what it means for H to explain E is that H be our reason for E.

This particular epistemic account of the explanatory relation does seem unavailable to the proponent of Inference to the Best Explanation, but it is important not to overstate the constraint. It certainly does not rule out all epistemic accounts of explanation. Moreover, the account that does appear to produce a vicious circularity is narrower than it may at first appear. Those familiar with Carl Hempel's seminal work on scientific explanation may identify it with him, since Hempel does sometimes write of good explanations providing understanding by providing reasons for believing that the phenomenon to be explained does actually occur. As Salmon has pointed out to me, however, Hempel need not be taken to claim that a good explanation has to be our *only* reason for believing the explanandum. Perhaps our reason for believing the explanandum is that we have observed it. Hempel may still require that a good explanation provides *another* reason. And this seems sufficient to break the circle, since when we infer H from E by means of an inference to the best explanation, we will also have grounds for E quite apart from H.

A final attenuating point to make about the unsuitability of some epistemic accounts of explanation for use in Inference to the Best Explanation is that, for me at least, it is really no additional constraint at all, since such an account of explanation would also already be unsuitable for the simpler reason that it would be false. To identify explanations with one's reasons for belief would entail that, for each thing we have reason to believe is the case, we *already* understand why it is the case. The starting point of enquiry into explanation, however, is the gap between knowing that something is the case and understanding why it is. When we ask a why-question -- Why is the sky blue? Why does the same side of the moon always face the earth? Why were you late? -- we almost always already have a reason to believe that the explanandum is true, so any account of explanation that identifies explanation with reason for belief would make understanding why things are as they are found to be appear much easier than it really is.

The second feature an account of explanation might have that would debar it from a role in Inference to the Best Explanation is that it might yield an account of inference that is not `epistemically effective'. What are the features that a hypothesis has when it is judged to be well-supported by the available evidence? What are the symptoms by which the likeliness of a hypothesis is judged? If explanatory features are among these symptoms, then these features must be available to the enquirer, and available to her before she makes her inference. According to `ontic' accounts of explanation, however, it is features of the world, rather than epistemic states, that do the explanatory work, so the worry arises that we would already need to know H before we could determine whether H explains E, in which case it appears we could not use explanatory considerations as a guide to whether H is true.

Here too, however, the extent of the constraint is easy to exaggerate. Note first of all that, insofar as epistemic effectiveness is a problem, it may apply to epistemic as well as to ontic accounts of explanations. If one requires of an explanation that it provide a reason to believe the explanandum, and one requires that the reasons themselves be known, then here too we would have to know H before we could know whether it explains E. Indeed the situation would arise on any account of explanation that requires that the explanation be true. But this is the reason why I stressed above the importance of the distinction between actual and potential explanation. An account of Inference to the Best *Actual* Explanation would indeed be circular, but an account of Inference to the Best *Potential* Explanation need not be, whether the account of explanation one uses is ontic or epistemic, because a judgement of the quality of a potential explanation does not require a prior judgement as to its truth value. As we will see, however, the issue of epistemic effectiveness is important, especially as it bears on the relationship between Inference to the Best Explanation and Bayesianism.

Explanatory Virtues

Inference to the Best Explanation does not then appear to rule out most approaches to explanation; but this does not make it any easier to come up with an adequate account of the explanatory relation. The difficulty of articulating Inference to the Best Explanation is compounded when we turn to the question of what makes one explanation better than another. To begin with, the model suggests that inference is a matter of choosing the best from among those explanatory hypotheses that have been proposed at a given time, but this seems to entail that at any time scientists will infer one and only one explanation for any set of data. As Salmon points out, however, this is not promising, since scientists will sometimes infer more than one explanation and will sometimes refuse to infer at all. Thus Salmon rightly emphasises `the multiplicity of legitimate explanations' (9): a single phenomenon can be correctly explained in different ways. This is very plausible, and just what one would expect from almost any model of explanation on the market. From the point of view of a causal model, for example, one event will have many explanations because one event has many causes. Just how extensive we should say the multiplicity of explanation is will depend on how finely we characterise phenomena or explananda. If we make them contrastive, for example, the degree of multiplicity will go down, since two explanations of P will turn out to be one explanation of P rather than Q while the other will be an explanation of P rather than R. Thus while both a drought and the absence of food reserves are causes of a famine in India, only the drought explains why India had a famine that year rather than the year before (when there was no drought), and only the absence of food reserves explains why there was a famine in India that year rather than in Egypt that year (which had food reserves). More generally, since what counts in practice as a legitimate explanation will depend on the interests of the inquirer, accounts of explanation that incorporate such factors into the explanandum will have more explananda to go around, and so fewer explanations for each of them (cf., van Fraassen, 1980, ch.5). But these finer points of the taxonomy of explanations are not crucial for present purposes, since they will at most affect the extent of multiplicity we find, not its existence.

The multiplicity of explanation appears an embarrassment for Inference to the Best Explanation, because of the uniqueness implied by the superlative term `best'. If there are several correct explanations of the data, but Inference to the Best Explanation only sanctions an inference to one of them, then Inference to the Best Explanation would seem incorrectly to place the others beyond our ken. That is the threat to Inference to the Best Explanation, but we should not take the uniqueness suggested by the slogan literally. Inference to the Best Explanation is meant to tell us something about how we choose between *competing* explanations: we are to choose the best of these. But among compatible explanations we need not choose. Thus inferences to the best explanation may support inferences to several causes of an event, each of which provides a legitimate explanation.

Inference to the Best Explanation must also take account of the important point that sometimes the best is not good enough, a point Salmon emphasises near the end of his essay (33-33). Sometimes the correct response to the data is agnosticism: the evidence is not strong enough to support inference to any of the available explanations. In such a case, as Salmon says, inference to the best available explanation would be a mistake and, as he suggests, an account of inference that sanctions such an inference would be mistaken. This objection from weak evidence is the flip side of the objection from the multiplicity of explanations. Both objections focus on the `best' in `Inference to the Best Explanation'. Where the multiplicity of explanations suggests that Inference to the Best Explanation sanctions too few inferences -- since `best' entails uniqueness -- the cases of weak evidence suggests that it would sanction too many -- since `best' entails existence. The correct response to the objection of weak evidence is once again to drop the entailment. Correctly construed, Inference to the Best Explanation only sanctions inference where the best is good enough.

So I agree with Salmon that scientists are sometimes agnostic, unwilling to infer any of the available hypotheses, and they are also sometimes happy to infer more than one explanation from the same data set, when the explanations are compatible. `Inference to the Best Explanation' must thus be glossed by the more accurate but less memorable phrase, `inference to the best of the available competing explanations, when the best one is sufficiently good'. But under what conditions is this complex condition satisfied? How good is `sufficiently good'? Even more fundamentally, what are the factors that make one explanation better than another? Standard models of explanation are virtually silent on this point. This does not suggest that Inference to the Best Explanation is incorrect but, unless we can say more about explanation, the model will remain relatively uninformative.

Fortunately, some progress has been made in analyzing the relevant notion of the best explanation. We may begin by considering a basic question about the sense of `best' that the model requires. Does it mean the most probable explanation, or rather the explanation that would, if correct, provide the greatest degree of understanding? In short, should Inference to the Best Explanation be construed as inference to the *likeliest* explanation, or as inference to the *loveliest* explanation? A particular explanation may be both likely and lovely, but the notions are distinct. For example, if one says that smoking opium tends to put people to sleep because opium has a `dormative power', one is giving an explanation that is very likely to be correct but not at all lovely: it provides very little understanding. At first glance, it may appear that likeliness is the notion Inference to the Best Explanation ought to employ, since scientists presumably only infer the likeliest of the competing hypotheses they consider. As I have already indicated, however, this is probably the wrong choice, since it would severely reduce the interest of the model by pushing it towards triviality. Scientists do infer what they judge to be the likeliest hypothesis, but the main point of a model of inference is precisely to say how these judgements are reached, to give what scientists take to be the symptoms of likeliness. If Inference to the Best Explanation is along the right lines, explanations that are lovely will also be likely, but it should be in terms of loveliness that the inference is made. For to say that scientists infer the likeliest explanations is perilously similar to saying that great chefs prepare the tastiest meals, which may be true, but is not very informative if one wants to know the secrets of their success. Like the dormative power explanation of the effects of opium, `Inference to the Likeliest Explanation' would itself be an explanation of scientific practice which provides only little understanding.

The model should thus be construed as `Inference to the Loveliest Explanation'. Its central claim is that scientists take loveliness as a guide to likeliness, that the explanation that would, if correct, provide the most understanding, is the explanation that is judged likeliest to be correct. This at least is not a trivial claim, but it raises three general challenges, all of which figure in Salmon's critique:

1. The first challenge is to identify the explanatory virtues, the features of explanations that contribute to the degree of understanding they provide.

2. The second challenge is to show that these aspects of loveliness match judgements of likeliness, that the loveliest explanations tend also to be those that are judged likeliest to be correct.

3. The third challenge is to show that, granting the match between loveliness and judgments of likeliness, the former is in fact the scientists' guide to the latter.

Identifying the Explanatory Virtues

To begin with the challenge of identification, there are a number of plausible candidates for the explanatory virtues, including scope, precision, mechanism, unification and simplicity. Better explanations explain more types of phenomena, explain them with greater precision, provide more information about underlying mechanisms, unify apparently disparate phenomena, or simplify our overall picture of the world. Some of these features, however, have proven surprisingly difficult to analyse. There is, for example, no uncontroversial analysis of unification or simplicity, and some have even questioned whether these are genuine features of the hypotheses deployed in scientific explanations, rather than artifacts of the way they happen to be formulated, so that the same hypothesis will count as simple if formulated in one way but complex if formulated in another (cf. Sober 1988).

A different but complementary approach to the problem of identifying some of the explanatory virtues focuses on the contrastive structure of many why-questions. As I have already observed, a request for the explanation of some phenomenon often takes a contrastive form: one asks not simply `Why P?', but `Why P rather than Q?'. What counts as a good explanation depends not just on fact P but also on the foil Q. Thus the increase in temperature might be a good explanation of why the mercury in a thermometer rose rather than fell, but not a good explanation of why it rose rather than breaking the glass. Accordingly, it is possible to develop a partial account of what makes one explanation of a given phenomenon better than another by specifying how the choice of foil determines the adequacy of contrastive explanations. Although many explanations both in science and in ordinary life specify some of the putative causes of the phenomenon in question, the structure of contrastive explanation shows why not just any causes will do. Roughly speaking, a good explanation requires a cause that `made the difference' between the fact and foil. Thus the fact that Smith had untreated syphilis may explain why he rather than Jones contracted paresis (a form of partial paralysis), if Jones did not have syphilis; but it will not explain why Smith rather that Doe contracted paresis, if Doe also had untreated syphilis. Not all causes provide lovely explanations, and an account of contrastive explanation helps to identify which do and which do not (cf. van Fraassen, 1980, ch. 5; Lipton, 1993).

Matching Explanatory and Inferential Virtues

Assuming that a reasonable account of the explanatory virtues is forthcoming, the second challenge to Inference to the Best Explanation concerns the extent of the match between loveliness and judgments of likeliness. If Inference to the Best Explanation is along the right lines, then the lovelier explanations ought also in general to be judged likelier. Here the situation looks promising, since the features we have tentatively identified as explanatory virtues seem also to be inferential virtues, that is, features that lend support to a hypothesis. Hypotheses that explain many observed phenomena to a high degree of accuracy tend to be better supported than hypotheses that do not. The same seems to hold for hypotheses that specify a mechanism, that unify, and that are simple. The overlap between explanatory and inferential virtues is certainly not perfect, but at least some cases of hypotheses that are likely but not lovely, or conversely, do not pose a particular threat to Inference to the Best Explanation. As we have already seen, the explanation of opium's soporific effect by appeal to its dormative power is very likely but not at all lovely; but this is not a threat to the model, properly construed. There surely are deeper explanations for the effect of smoking opium, in terms of molecular structure and neurophysiology, but these explanations will not compete with the banal account, so the scientist may infer both without violating the precepts of Inference to the Best Explanation.

The structure of contrastive explanation also helps to meet this matching challenge, because contrasts in why-questions often correspond to contrasts in the available evidence. A good illustration of this is provided by Ignaz Semmelweis's nineteenth-century investigation into the causes of childbed fever, an often fatal disease contracted by women who gave birth in the hospital where Semmelweis did his research. Semmelweis considered many possible explanations. Perhaps the fever was caused by `epidemic influences' affecting the districts around the hospital, or perhaps it was caused by some condition in the hospital itself, such as overcrowding, poor diet, or rough treatment. What Semmelweis noticed, however, was that almost all of the women who contracted the fever were in one of the hospital's two maternity wards, and this led him to ask the obvious contrastive question and then to rule out those hypotheses which, though logically compatible with his evidence, did not mark a difference between the wards. It also led him to infer an explanation that would explain the contrast between the wards, namely that women were inadvertently being infected by medical students who went directly from performing autopsies to obstetrical examinations, but only examined women in the first ward. This hypothesis was confirmed by a further contrastive procedure, when Semmelweis had the medics disinfect their hands before entering the ward: the infection hypothesis was now seen also to explain not just why women in the first rather than in the second ward contracted childbed fever, but also why women in the first ward contracted the fever before but not after the regime of disinfection was introduced. This general pattern of argument, which seeks explanations that not only would account for a given effect, but also for particular contrasts between cases where the effect occurs and cases where it is absent, is very common in science, for example wherever use is made of controlled experiments (Cf. Hempel, 1966, ch. 2; Lipton, 1991, ch. 5).

Loveliness a Guide to Likeliness

This leaves the challenge of guiding. Even if it is possible to give an account of explanatory loveliness (the challenge of identification) and to show that the explanatory and inferential virtues coincide (the challenge of matching), it remains to be argued that scientists judge that an hypothesis is likely to be correct *because* it is lovely, as Inference to the Best Explanation claims. Thus a critic of the model might concede that likely explanations tend also to be lovely, but argue that inference is based on other considerations, having nothing to do with explanation. For example, one might argue that inferences from contrastive data are really applications of Mill's method of difference, which makes no explicit appeal to explanation, or that precision is a virtue because more precise predictions have a lower prior probability and so provide stronger support as an elementary consequence of the probability calculus.

The defender of Inference to the Best Explanation is here in a delicate position. In the course of showing that explanatory and inferential virtues match up, he will also inevitably show that explanatory virtues match some of those other features that competing accounts of inference cite as the real guides to inference. The defender thus exposes himself to the charge that it is those other features rather than the explanatory virtues that do the real inferential work. Meeting the matching challenge will thus exacerbate the guiding challenge. The situation is not hopeless, however, since there are at least two ways to argue that loveliness is a guide to judgments of likeliness. At least many other accounts of inference fail to get the extension right: they are inapplicable to many scientific inferences and incorrect about others. If it is shown that Inference to the Best Explanation does better in this respect, then this is a powerful reason for supposing that loveliness is indeed a guide to likeliness. Secondly, if there is a good match between loveliness and likeliness, as the guiding challenge grants, this is presumably not a coincidence and so itself calls for an explanation. Why should it be that the hypotheses that scientists judge likeliest to be correct are also those that would provide the most understanding if they were correct? Inference to the Best Explanation gives a very natural answer to this question, similar in structure to the Darwinian explanation for the fact that organisms tend to be well-suited to their environments. If scientists select hypotheses on the basis of their explanatory virtues, the match between loveliness and judgments of likeliness follows as a matter of course. Unless the opponents of the model can give a better account of the match, the challenge has been met.

The Bayesian and the Explanationist should be Friends

I want now to say something about the relationship between Inference to the Best Explanation and the kind of Bayesian perspective that Salmon favours. I am not well placed to do this, because I am no expert on Bayesianism, but I want to sketch the irenic view advertised at the beginning of these remarks, according to which the two approaches are compatible. If this suggestion is in fact incoherent, I count on Salmon to say why, and I will be a sadder but wiser epistemologist.

If one wishes to map Inference to the Best Explanation onto the Bayesian scheme, a natural thought is that the distinction between the loveliness and the likeliness of an explanation corresponds to the Bayesian distinction between prior and posterior probability, the probability hypothesis H has before and after the evidence E comes in, respectively. Unfortunately, things do not seem that neat, since while likeliness may well correspond to posterior probability, I do not think that loveliness can be equated with the prior. Perhaps the easiest way of seeing this is to note the relational character of loveliness. A hypothesis is only a good or bad explanation relative to a specific explanandum. Contrastive explanations of the sort I mentioned above make the point vividly, since a good explanation of P rather than Q may not be a good explanation of P rather than R, but the point applies also to non-contrastive cases as well, since clearly a good explanation of P will not in general be a good explanation of S. Prior probability is also a relative notion -- it is relative to previous evidence -- but it is not relative to the new evidence E on which the Bayesian would have the enquirer conditionalise in order to move from prior to posterior. Loveliness, by contrast, is relative to that new evidence.

Another tempting connection would be to link loveliness not to the prior but to the Bayesian notion of likelihood -- to the probability of E given H. (I must here apologise for a growing clump of terminology that is now ripe for confusion: likelihood is not to be confused with what I have called `likeliness'! The list of `Some Key Concepts' at the end of this essay may somewhat reduce the probability of confusion.) The identification of loveliness with likelihood seems a step in the right direction, since both loveliness and likelihood are relative to E, the new evidence. But I am not sure that this is quite correct either, since H may give E high probability without explaining E. Indeed H may entail E yet not explain it as some of the counterexamples to the Deductive-Nomological model of explanation show. (Nevertheless, as Samir Okasha has pointed out to me, it might be that whenever H1 is a lovelier explanation of E than H2, the likelihood of H1 is greater than the likelihood of H2.)

It appears that loveliness does not map neatly onto any one component of the Bayesian scheme. Some aspects of loveliness, some explanatory virtues -- including scope, unification and simplicity -- are related to prior probability; others seem rather to do with the transition from prior to posterior. But what does this mean? My thought is this. In many real-life situations, the calculation that the Bayesian formula would have us make does not, in its bare form, meet the general requirement of epistemic effectiveness, a requirement I introduced above in my discussion of features of accounts of explanation that would make them ill-suited for use in the context of Inference to the Best Explanation. Here the point concerns the terms in the Bayesian formula. Simply put, we do not always know how to work out the probabilities that are required in order to move from prior to posterior probability simply on the basis of a (presumably tacit) grasp of the abstract principles of the probability calculus. My suggestion is that explanatory considerations of the sort to which Inference to the Best Explanation appeals are often more accessible than those principles to the enquirer on the street or in the laboratory, and provide an effective surrogate for certain components of the Bayesian calculation. On this proposal, the resulting transition of probabilities in the face of new evidence might well be just as the Bayesian says, but the mechanism that actually brings about the change is explanationist.

To make this case out in detail is a big job, certainly too big for here and probably too big for me. I suggest it primarily so as to elicit Salmon's assessment of its own prior probability. (Though I cannot in fairness expect him to rate it highly simply on the grounds it provides a lovely explanation of various inferential practices!) But I want briefly to suggest how explanatory considerations might help to lubricate the Bayesian mechanism, in four ways:

1. The first role for explanatory considerations is with the determination of likelihood, which is needed for the transition from prior to posterior probability.

2. The second is with the determination of the priors, the input to conditionalising.

3. The third concerns the determination of relevant evidence.

4. The fourth explanatory considerations may be connected to a scientific preference for fertile or fruitful theories.

Determination of Likelihood

One way in which explanatory considerations might be part of the actual mechanism by which enquirers move from prior to posterior probabilities is by helping enquirers to assess likelihoods, an assessment essential to Bayesian conditionalising. For although likelihood is not to be equated with loveliness, it might yet be that one way we judge how likely E is on H is by considering how well H would explain E. This would hardly be necessary in cases where H entails E, but in real life inference this is rarely the case and, where H does not entail E, it is not so clear how in fact we do work out how likely H makes E (and how likely not-H makes E). Here explanatory considerations might help, if in fact loveliness is reasonably well correlated with likelihood. What would be required, I think, is that lovelier explanations tend to make what they explain likelier (even if high likelihood is no guarantee of good explanation), and that we sometimes exploit this connection by using judgements of loveliness as a barometer of likelihood.

For example, when we consider the loveliness of a potential causal explanation, we may consider how the mechanism linking cause and effect might run, and in so doing we are helped in forming a judgement of how likely the cause would make the effect and how unlikely the effect would be without the cause. This mechanism may also be at work in the context of contrastive explanation. When Semmelweis was investigating the causes of childbed fever, he repeatedly considered to what extent various hypotheses explained the contrasts in his data, such as contrasts between rates of the fever in different wards and within a single ward under different conditions. The suggestion is that Semmelweis was aided in coming to a view of likelihoods by considering how well those data would be explained by the competing hypotheses. The case would have to be made out, but on this occasion I only wish to make the suggestion clear. Inference to the Best Explanation proposes that loveliness is a guide to likeliness (a.k.a. posterior probability); the present proposal is that the mechanism by which this works may be understood in part by seeing the process as operating in two stages. Explanatory loveliness is used as a symptom of likelihood, and likelihoods help to determine likeliness or posterior probability. This is one way Inference to the Best Explanation and Bayesianism may be brought together.

Determining the Priors

Another obvious place to look for a way explanatory considerations might in practice play an important role in a Bayesian calculation is in the determination of prior probabilities. I begin with a general observation about the role of priors in Inference to the Best Explanation. Salmon claims, surely correctly, that choices between competing potential explanations of some phenomenon are often driven by judgements of which of the explanations has the higher prior. This is one important source of his suspicion about Inference to the Best Explanation: the choices here seem actually to be based on judgements of which is the likeliest explanation, judgements which in many cases depend on which potential explanation is judged to have the highest prior, not on which is the loveliest explanation (cf. 30). My reply to this is to emphasise my agreement about the crucial role that priors play in this way, but to deny that this is in tension with Inference to the Best Explanation. Consider what the Bayesian himself says about the priors. He of course does not take their crucial role to undermine the importance of the Bayesian formula, roughly because today's priors are yesterday's posteriors. That is, the Bayesian claims that today's priors are generally themselves the result of prior conditionalising. Similarly, the defender of Inference to the Best Explanation should not deny that inference is mightily influenced by the priors assigned to competing explanations, but she will claim that those priors were themselves generated in part with the help of explanatory considerations.

This means that, insofar as my first suggestion -- that explanatory considerations play a role in conditionalising -- has merit, explanatory considerations also have a role to play in the determination of priors, since priors are partially determined by earlier conditionalisation. But we may also see how explanatory considerations enter into the determination of priors in other ways. This is where considerations of unification, simplicity, and their ilk would naturally come into play. The Bayesian is I think happy to acknowledge the role that these sort of factors may play in fixing prior probabilities, and the prospects seem promising for showing that some of these may in practice be determined by considering explanatory quality.

Determining the Relevant Evidence

The third possible point of contact between the Bayesian and the explanationist concerns the determination of what is the relevant evidence. Bayes theorem describes the transition from prior to posterior, in the face of specified evidence. It does not, however, say *which* evidence one ought to conditionalise on. In principle perhaps, non-demonstrative inference should be based on `total evidence', indeed on everything that is believed. In practice, however, investigators must think about which bits of what they know really bear on their question, and the need also to decide which further observations would be particularly relevant. So it seems that a Bayesian view of inference needs some account of how the evidential input into the conditionalising process is selected, and here there seems yet another area where the explanationist may contribute. To give just one example of how this might work, consider how we sometimes discover supporting evidence for a hypothesis by seeing what it would explain. My suggestion is that we sometimes come to see that a datum is epistemically relevant to a hypothesis precisely by seeing that the hypothesis would explain it. (Arthur Conan Doyle often exploited this phenomenon to dramatic effect: in `Silver Blaze', the fact that the dog did not bark would have seemed quite irrelevant, had not Sherlock Holmes observed that the hypothesis that a particular individual was on the scene would explain this, since that person was familiar to the dog.) Thus explanatory considerations may help to determine which bits of evidence should enter the process of Bayesian conditionalisation.

Explanation and Fertility

Finally, Salmon has helpfully suggested to me that explanationist considerations might be particularly well suited to account for scientists' preference for fertile hypotheses. This important point has a rather different conceptual location than the first three roles for explanationism that I have flagged. Those were all ways explanatory considerations might enter into the Bayesian process of fixing priors and conditionalising on the evidence. The output of this process is a posterior probability, and Salmon and I have both regularly assumed that among competing hypotheses, scientists will prefer the one with the highest posterior probability, the likeliest explanation. In fact, however, I think that neither Salmon nor I quite believe this, because we both believe that probability is not the only aim of inference. Unlike instrumentalists and constructive empiricists, but alongside Popper, we think that scientists also have a preference for theories with great content, even though that is in tension with high probability, since the more one says the more likely it is that what one says is false. This interest in scope and fertility is captured neither by basic Bayesianism nor by Inference to the Best Explanation as I have expounded it, but I agree with Salmon that it is a promising area in which explanationist considerations may operate, since scientists may judge theoretical fertility or promise by assessing the explanatory potential of the hypotheses they are evaluating.

The Quasar and the Bone

Before concluding these remarks, I would like briefly to respond to two examples that Salmon discusses, the inference from double quasar images to gravitational lensing (16-17), and the inference from the surprisingly old worked bone found in the Yukon to the hypothesis that the bone was frozen for many years before being worked (18-21, 27-28). The main point for Salmon of both examples is I think to provide cases which appear at first to fit the Inference to the Best Explanation schema well, yet which on closer examination cannot be analysed in explanationist terms.

The case of gravitational lensing is an example of inference to a common cause. As Salmon says, scientists observed two quasar images with identical spectral characteristics, where the supposition of the coincidental existence of two corresponding quasars was extremely improbable. Two alternatives were considered. One was that there really are two similar quasars but that this is not coincidental, since they evolved from a common ancestor; this explanation is rejected because there is no plausible mechanism by which this evolution might have occurred. The other is that the double image is caused by a single quasar, whose light is relativistically bent by huge galactic masses. This is the hypothesis that was inferred; it was supported by relativity theory and by the subsequent discovery of a cluster of galaxies in the right place to generate the calculated bending.

As Salmon suggests, this is a case that appears to fit Inference to the Best Explanation extremely well, in part because the bending hypothesis seems to have been accepted because it would explain the observed double image, and preferred to the common evolution hypothesis because, in the complete absence of an evolutionary mechanism, that explanation was far uglier than the hypothesis of gravitational bending. Why then does Salmon claim that the applicability of Inference to the Best Explanation to common-cause inferences is illusory? I think he has two main reasons. One is that such inferences require that one begin by assessing the probability of brute coincidence -- since this must be less than any inferred explanation -- and this assessment is not made on the basis of explanatory considerations. Saying that something is a coincidence does not explain why the two events occurred together -- it is more like saying there is no explanation -- but it may have a lower or a higher probability. If the two events are independent, then the probability of the coincidence will be the product of the probability of each event, a figure not determined by explanatory considerations. Salmon's second main reason for rejecting an explanationist analysis of the lensing case is that he holds that, even once the coincidence hypothesis is dismissed, the choice between the remaining hypotheses, though now a choice between explanations, is still not made on explanationist grounds. As we have seen, Salmon claims that scientists prefer the likeliest explanation, not the loveliest one.

I agree with Salmon that the decision whether or not to infer any common cause will depend on our assessment of the probability of a genuine fluke or coincidence. And since this probability seems the probability of there being no explanation, it is difficult to see how an explanationist approach can account for this assessment. But how does any approach account for the probability of coincidence? We must judge the probability of each event, and this judgement must on any account flow from our background knowledge. According to the explanationist, this background knowledge was itself generated with the help of explanatory considerations, and in light of this she may I think happily admit that the judgement of the probability of coincidence has only this indirect connection to inferences to the best explanation. This would be no more threatening to Inference to the Best Explanation than admitting that a hypothesis, inferred on explanationist grounds, may then be used to derive a prediction, though that prediction is not inferred because of what it explains.

There is also however an additional and more ambitious response the explanationist might make to Salmon's first point. In the absence of a common cause, the coincidence of two quasars with exactly the same spectral characteristics is presumably highly improbable, but on reflection it does not appear that the likeliness of a common cause is determined simply by multiplying the probability of each quasar's existence. For consider the probability of another quasar with a different but equally probable spectrum. If S1 is the spectrum of the two quasars under the hypothesis of coincidence, and S2 is the spectrum of my new quasar, then the probability of S1 and S1 (two quasars with the same spectrum) is the same as the probability of S1 and S2 (two quasars with different spectra), yet I take it that only the first combination is taken to be an improbable coincidence. (This is similar to the familiar point that the probability of being dealt thirteen cards from a shuffled deck that are all spades is no less than the probability of being dealt any other specific hand of thirteen cards.) What this suggests is that there is more to coincidence than the product of the probabilities, and the tempting thought is that this additional content may be analysed in part in terms of explanation. For it is only the fact of two quasars with exactly the same spectra that cries out for an explanation. (As Jason Grossman has pointed out to me, one reason for this may be that we tend to compare the probability of two quasars with the same spectrum with the (large) probability of being a member of the large set of two quasars with different spectra, while we tend to compare the probability of two quasars with different but specified spectra with the (small) probability of two quasars with any other two specified spectra.)

Salmon's second claim is that even once we dismiss the possibility of brute coincidence, the choice between explanations is not based on considerations of loveliness, but only likeliness. The example, however, does not suggest this, for as Salmon points out, the lensing explanation is far lovelier than the evolution explanation, in part because we can imagine no appropriate evolutionary mechanism. It is true that our confidence in the lensing explanation depends in large measure on its fit with relativity theory, but by this stage it should be clear why I do not regard this as a difficulty for Inference to the Best Explanation, insofar as relativity theory itself has explanationist attractions. Salmon's main reason for claiming that the choice between evolution and lensing is not based on explanatory reasons is, I think, just that he is convinced that it is based on Bayesian considerations. What I have been arguing, however, is that the one does not exclude the other. Given the apparent role of explanatory judgements in this example (one of the reasons Salmon presented it), the example thus does not appear to tell against the explanationist programme.

Similar observations apply to the worked bone example, where scientists must choose between various explanations of the discovery of a worked bone in the Yukon that very substantially precedes any other evidence of human habitation in that region. Although the scientific details are fascinating, Salmon's main purpose in developing the example is I think to make plausible the claim that the choice between competing explanations of the early date of the bone is made on Bayesian grounds: we estimate prior probabilities of each hypothesis, conditionalise, and then prefer the hypothesis with the highest posterior probability (29). But once we allow the possibility that Bayesian and explanationist mechanisms are compatible, the example is not so problematic for the explanationist. As Salmon remarks, of the three live contenders in this case, two of them are much uglier explanations than the preferred hypothesis that the bone was worked only after a long period in permafrost, because they leave what Salmon usefully calls `explanatory gaps' (23-24). The hypotheses of much older human habitation leaves unexplained the complete absence of found traces for the intervening period; the hypothesis that the older habitation was temporary leaves unexplained the absence of any independent evidence for the extinction of this community or its return to the old world.

I do not want to appear complacent. To show that Bayesian reasoning is compatible with explanationist reasoning is not the same as showing directly that explanationist reasoning is going on. (Though I do think the cases suggest this.) And in the bone case, there is I think a source of residual doubt. Unlike the quasar evolution hypothesis, which is explanatorily unsupported by mechanism, the hypothesis that people moved into the Yukon a long time ago and then moved out is as causally articulate as the hypothesis that the bone was preserved in permafrost for a long time before being worked. So one may feel that there really are no explanationist grounds for preferring one to the other. What it comes down to, it seems, is rather the greater prior probability of permafrost preservation, given the absence of independent evidence of temporary habitation, where this has nothing to do with the explanatory loveliness. But of course I would resist this feeling. Firstly, there is my mantra that priors are based on background beliefs themselves partially generated on explanatory grounds. We must also take seriously Salmon's explanatory gaps. For Inference to the Best Explanation and Bayesianism alike, inference is to some extent a holistic process. A hypothesis may provide a perfectly good explanation of the focal phenomenon, yet still be rejected, and on explanationist grounds, because of its failure to explain other things. When we reject a hypothesis because it creates more explanatory problems than it solves, as seems to be the case for the hypothesis of temporary inhabitation, we are rejecting it on explanationist grounds.

Conclusion

Early on in these remarks, I presented three general challenges that face Inference to the Best Explanation, the challenges of identification, of matching and of guiding. The challenge of identification is the challenge of providing an account of explanation and of what makes one explanation better than another that is articulated and correct. Salmon rightly observes that we are a long way from meeting this challenge, but we both believe that progress has been made in this area and that the project is well worth pursuing, whatever one thinks of Inference to the Best Explanation. The challenges of matching and of guiding are the challenges of showing that there is a strong enough correlation between loveliness and likeliness to support Inference to the Best Explanation and that enquirers exploit this correlation by using explanatory virtues as symptoms of likeliness. Here Salmon has been even more sceptical, suggesting not only that these things have not been shown, but they cannot be shown, since one cannot show what is not the case. I agree of course that these challenges have not yet been properly met, but I hope that my reply may lead Salmon to suspend judgement at least on whether they can be met. I have attempted to defuse Salmon's challenging objections to Inference to the Best Explanation in a variety of ways, but especially by suggesting why arguments for Bayesianism are not in themselves arguments against explanationism, since assessments of fertility, the determination of relevant evidence and of prior probabilities, and the Bayesian transition from prior to posterior probability may perhaps all be aided and abetted by the sorts of explanatory considerations that the account of inference to the best explanation promotes.

Acknowledgements -- I am very grateful for comments from Jason Grossman, Giora Hon, Samir Okasha, Sam Rakover, and especially from Wes Salmon.

References

Cartwright, Nancy, 1983. *How the Laws of Physics lie*. Oxford: Oxford University Press.

Fraassen, Bas van, 1980. *The Scientific Image*. Oxford: Oxford University Press.

Goodman, Nelson, 1983. *Fact, Fiction, and Forecast*, 4th ed.. Indianapolis: Bobbs-Merrill.

Harman, Gilbert 1965. `The Inference to the Best Explanation', *The Philosophical Review*, **74**, 88-95.

Hempel, Carl, 1965. *Aspects of Scientific Explanation*. New York: Free Press.

Hume, David, 1777. *An Enquiry Concerning Human Understanding*, L.A. Selby-Bigg & P.H. Nidditch (eds.), 1975. Oxford: Oxford University Press.

Lipton, Peter, 1991. *Inference to the Best Explanation*. London: Routledge.

Lipton, Peter 1993. `Contrastive Explanation', in D. Hillel-Ruben (ed.), *Explanation*, Oxford: Oxford University Press.

Peirce, Charles Sanders, 1931. *Collected Papers*. C. Hartshorn & P. Weiss (eds.), Cambridge, MA: Harvard University Press.

Sober, Elliott, 1988. *Reconstructing the Past: Parsimony, Evolution and Inference*. Cambridge, MA: MIT Press.

Thagard, Paul 1978. `The Best Explanation: Criteria for Theory Choice', *The Journal of Philosophy*, **75**, 76-92.

Some Key Concepts

Of Explanation:

Actual Explanation -- Hypothesis that meets all the requirements on explanation, including that of truth.

Potential Explanation -- Hypothesis that meets all the requirements on explanation, except possibly that of truth. Hence all actual explanations are also potential explanations, but not conversely.

Loveliness of an Explanation -- A measure of how good a potential explanation is, of how much understanding it would provide if it is or if it were an actual explanation. According to a strong version of Inference to the Best Explanation, the loveliness of a potential explanation is a guide to its likeliness or posterior probability.

Of Inference:

Prior Probability -- Probability of the hypothesis before the evidence under consideration is known.

Posterior Probability -- Probability of the hypothesis after the evidence is known. On a Bayesian approach, the transition from prior to posterior probability is governed by Bayes' Theorem.

Likeliness -- Posterior probability.

Likelihood -- The probability of the evidence, given the truth of the hypothesis. Needed for the application of Bayes' Theorem.