Metascience (2005) 14:331–361
DOI 10.1007/s11016-005-3431-7

REVIEW SYMPOSIUM

WOULDN'T IT BE LOVELY: EXPLANATION AND SCIENTIFIC REALISM

Peter Lipton *Inference to the Best Explanation* (second edition). London: Routledge, 2004. Pp. xii + 220. £18.99 PB.

By James Ladyman

The second edition of Peter Lipton's classic text contains new and important material on the causal model of explanation, the relation of inference to the best explanation to the Bayesian account of scientific reasoning, how exactly explanation guides inference, and why we ought to think that explanatory virtues are truth-tropic. Lipton is a wonderfully clear writer and a thorough and subtle philosopher, and his book is both a student-friendly introduction to the issues addressed, and essential reading for expert epistemologists and philosophers of science. Appeal to the notion of inference to the best explanation is ubiquitous in defences of scientific realism, but also elsewhere in philosophy where the explanatory virtues of theories are often the only purported grounds for accepting or rejecting them. Despite this, most authors are far from explicit about the details of inference to the best explanation, and Lipton's book is the most sustained investigation of the relationship between explanation and inference currently available. Furthermore, Lipton is exemplary in his engagement with the problems his arguments face, and judiciously modest in his claims, though not so modest as to court triviality. Hence, the book is replete with interesting and careful arguments. Everyone interested in epistemology or philosophy of science ought to read this book. That said, in my discussion below I will concentrate on what I regard as problems with some of Lipton's arguments.

The model of explanation which he develops is contrastive and causal. Lipton is clear that he does not think all explanations are causal, but he does think that many are, especially in science, and, in his investigation of the connection between inference and explanation, he confines himself to causal explanations. He argues that contrasts in explanations are an irreducible feature of them. This is key to his account because it allows him to assimilate Mill's methods – most significantly the method of difference – to it. So, he argues, to explain P is to explain why P happened, not Q, and to do so by citing a causal factor whose antecedent presence led to P rather than Q. Lipton is hopeful that whatever account of causation turns out to be the right one will be compatible with the use he makes of causal explanation. It is interesting to consider whether this neutrality is plausible. It is not uncommon for philosophers of science to adopt both realism about unobservable entities and Humeanism about natural necessity, whether in relation to laws or causation. However, I think that van Fraassen is right to associate scientific realism with a commitment to the metaphysics of modality. Note that among those philosophers who have sought to give an account of explanation that is not prone to the standard objections to the deductive-nomological account, the prevalent strategy is to invoke a thick notion of causation and to declare that the explanans must invoke the real cause of the explanandum to avoid counter-examples due to irrelevance or symmetry. The Humean view of causation, which reduces singular causation to generic causation, and generic causation to laws, which are construed as mere regularities, may collapse Lipton's causal theory of explanation into the deductive-nomological model against which he argues. Furthermore, from a Humean point of view it is difficult to see why we should take explanation seriously in epistemology. To take one of Lipton's examples, why would the presence of tracks in the snow be grounds for inferring that someone had walked there recently, if we thought that tracks and the walking were merely accidentally conjoined? Lipton thinks that realists can appeal to the greater warrant they have for beliefs about the observable world in virtue of their beliefs about the unobservable world, but arguably they can only claim an advantage with respect to warrant for beliefs about unobserved observables if the causal explanations they rest upon are taken to be more than the subsuming of events under accidental regularities.

Lipton claims that Hume shows that "the impossibility of justifying induction does not depend on a particular account of our practices, but only on the fact that they are inductive" (p. 145). This is contentious because it threatens to reduce Hume's argument to triviality, since, if the only thing that all inductive arguments have in common is that they are deductively invalid, then it looks as if Hume is simply pointing out that induction is not deduction. Although some philosophers have sought to show that inductive arguments can be reformulated as deductive ones, many others have argued that there is no reason to suppose that the only form of justification is deductive. Furthermore, there are many deductively invalid inferences which we are not inclined to call inductive. For example, the inference "this relation is transitive and reflexive therefore it is symmetric", is invalid, but not inductive. We ought to restrict inductive inferences to a proper subset of deductively invalid ones. In particular, it is arguable that Hume had in mind inferences from the observed to the unobserved. Unfortunately, this does not amount to a formal characterisation since many invalid inference forms could be inductive ones in some suitably constructed context. For example, "if p then q, q therefore p" (affirming the consequent) can be a reliable inductive inference as in "if a number drawn at random is less than 1000 then it is less than 1001, this number drawn at random is less than 1001, therefore it is less than 1000".

However, thinking of inductive inferences as those that extrapolate from the observed to the unobserved helps us make sense of why Hume thought they all rely upon the principle of the uniformity of nature, which gives no support to the fallacious inference about relations above. (The lesson of Goodman's new problem of induction is that the world might be uniform in ways we don't expect, so the uniformity of nature cannot underpin a formal theory of induction.) Mill equated the uniformity of nature with determinism but one of the lessons of quantum mechanics is that, in principle at least, nature could be uniform in the weaker sense of there being stable but ineliminably probabilistic propensities for events. So perhaps we can do better than Lipton suggests when he says "there is a sense in which the success of induction is miraculous or inexplicable on any account of how it is done" (p. 145). The success of induction would not be miraculous if the world had a determinate (but perhaps indeterministic) causal/nomological structure, and that it has such a structure is the best explanation of inductive success, and so we can justify the use of inference to the best explanation at the object level, using an inference to the best explanation at the meta-level of its past success. Of course, this is circular in the sense that someone who abstains from inference to the best explanation at the object level will not be persuaded by its use at the meta-level, but nonetheless someone who endorses inference to the best explanation at the object-level need not find its success miraculous. (Lipton agrees that circularity need not always be worthless.)

It is interesting that Lipton thinks that "there is no way to show *a priori* that [inductive inferences] will be successful, because to say that they are non-deductive is just to say that there are possible worlds where they fail" (p. 145). If he is talking about some particular inductive inferences then their fallibility surely entails that he is right. However, necessitarians about laws might argue that the rationality of inference to the best explanation is knowable *a priori* because it is necessary (and possibly also analytic) that it is rational. In any case, there are other reasons for rejecting the idea of a world in which induction fails completely. First we need to distinguish two ways of taking this claim. The first interprets it as meaning that there is a possible world such that when reasoning about it from, as it were, the outside, induction fails. So we can imagine a world in logical space where all sorts of events randomly happen with no rhyme nor reason. Such a world would be very distant from ours because we clearly live in a world with a fair amount of regularity (even if it is all going to break down in the near future). The existence of such a world is of little interest because it would lack sufficient regularity for there to be any epistemic agents in it. On the other hand, a more interesting possibility would be a world like ours, in so far as there are epistemic agents who are something like us, but where induction fails utterly from the inside. However, we have no reason to believe there could be such a world. Such agents would need to have stable enough physiologies to do experiments and think up theories, and their environment would need to have stable enough objects for them to eat lunch and record their ideas. Indeed what does it mean to imagine objects at all like those in our world, but in a world where induction fails completely? It seems to be unintelligible because for something to count as an object is for it to exhibit sufficient stability and invariance in its behaviour in particular respects and to particular degrees. In sum, I think that Lipton's account would be more plausible if it was tied to realism about natural necessity and inference to the best explanation was employed at the meta-level to justify the latter.

Lipton structures his project around the distinction between the descriptive and normative aspects. The former is the task of adequately describing inductive inference, and the latter is the task of explaining how it is justified. Being true to his own theory, Lipton explains the descriptive virtues of his model of explanation by contrasting it with the deductive-nomological model, and in so doing he makes considerable progress with the descriptive problem. In so far as we expect philosophy to reach some kind of reflective equilibrium with common sense and pre-philosophical judgements, the descriptive and normative problems are not entirely distinct, because any theory of rationality that implies that all our inductive inferences are irrational could be discounted, and hence if someone can give an account of the form our actual inductive reasoning takes, then reasoning in accordance with that account cannot be systematically irrational on pain of violating reflective equilibrium. However, granting that Lipton's descriptive claims are broadly right, this is not yet enough to establish as much by way of normative claims as some philosophers would like. One question that has been brought to prominence by van Fraassen's recent epistemological writings is whether, when we consider the rationality or justification of induction, we are asking if it is ever permissible to make ampliative inferences, or whether doing so is ever rationally required. Lipton discusses van Fraassen's views in his later chapters and thinks he mis-describes scientific practice since many scientists do appear to be guided by explanatory considerations to believe in unobservables.

However, van Fraassen could agree with the descriptive claim that often our inductive inferences are guided by explanatory considerations, and he accepts that to be so guided is not prohibited by the canons of rationality. However, he argues that nobody is ever rationally compelled to believe something because it is the best explanation of the phenomena. Furthermore, he claims that inference to the best explanation, while not irrational, is nonetheless only pragmatically motivated in general. As it turns out, being guided by explanatory considerations has led us to arrive at empirically adequate theories, and that gives us some reason to search for explanations in the future, but we should not admit explanatory considerations as reasons for belief if we are good empiricists. Lipton thinks that van Fraassen proposes an arbitrary restriction on the use of inference to the best explanation by allowing it forinferences about unobserved observables but not for unobservables. On the contrary, van Fraassen always allows inference to the best explanation whether in the domain of the observable or the unobservable, but never regards it as rationally compelling. It may be objected that it is capricious to use inference to the best explanation widely, but to always abstain from inferring the truth of the conclusion in the case of unobservable entities. However, there is a salient difference between inferring the existence of an unobserved observable and inferring the existence of an unobservable, namely that in the former case it is usually the inferring of the existence of an unobserved token of an observed type that is at issue (we have seen people leave tracks in the snow before). The history of science gives us further reasons to be wary of committing ourselves to the existence of the unobservables postulated to explain observable phenomena.

An important element of Lipton's account is the distinction between the likeliest and the loveliest explanations. If inference to the best explanation is simply inference to the most probable explanation, then talk of inference to the "best" explanation would be misleading and eliminable. So one of Lipton's main contentions is that the loveliness of an explanation is a guide to its likelihood, so that the explanatory virtues of theories are indicative of their truth. On his view then, qualitative judgements about the understanding provided by an explanation – that it is deep, elegant, simple, unifying and so on – are guides to its likelihood. This is the essence of the rapprochement that Lipton suggests between his view and Bayesianism according to which the cognitive realisation of Bayesian reasoning involves thinking about probabilities indirectly via thinking about explanations.

With his contrastive account of explanation Lipton embraces the interest-relativity of explanations which van Fraassen takes to be one of the main reasons for regarding explanatory power as a pragmatic and not an epistemic virtue of theories. He discusses what he calls "Hungerford's objection" to the epistemic credentials of inference to the best explanation, namely, that since what is the loveliest explanation is relative to agents' judgements about what makes for a good or bad explanation, we have no reason to infer it. Lipton counters by pointing out that inference is similarly relative to the agent's beliefs and evidence, and hence argues that the interest-relativity of explanatory loveliness does not detract from the latter's role in explaining what makes some inferences warranted. Another objection is inspired by Voltaire, namely, that there is no reason to suppose that the explanations which we find lovely have anything to do with what is likely; in other words, that only if we live in the loveliest of all possible worlds will the loveliest explanations also be the true ones. This is really the heart of the problem for Lipton since there are numerous cases from the history of science of explanations which fitted the background metaphysical beliefs of the time, but which turned out to be false. He argues that we have reason to believe our explanatory values are increasingly truth-tropic as we learn from our mistakes.

One of Lipton's most interesting arguments is against what he calls "the argument from under-consideration". The idea is that the antirealist who employs this argument presupposes that scientists may be able reliably to rank theories they have thought of comparatively with respect to likelihood of truth, but also denies that this amounts to a reliable absolute ranking. If we assume the "no-privilege" premise which states that there is no reason to suppose that the process which generates theories for ranking will be likely to generate the true theory, then comparative ranking does not amount to reliable inference because the true theory may be one they have not even entertained. Lipton's response is to argue that the no-privilege premise is false and that comparative ranking amounts to absolute ranking. With respect to the former he appeals to the fact that background theories contribute to reliable comparative ranking of other theories and so we have reason to believe that the are true. However, the antirealist could reply by claiming that it is only the empirical adequacy of background theories which matters. With respect to the latter, he claims that to concede the reliable *comparative* ranking of contrary theories is to concede the reliable *absolute* ranking of them because contraries entail contradictories. In other words, if q is a contrary of p, then q entails not-p and so to rank q above p is to rank not-p above p; that is, to rank the likelihood of p being false above the likelihood of p being true. However, this will not work when we are talking about ranking theories with respect to the likelihood of their being approximately true, because, in general, q being approximately true will not entail p not being approximately true. If, for example, we consider the wave and particle theories of light (classic candidates for a real case of underdetermination from the history of science), then clearly, although the wave theory was ranked above the particle theory by 1850, both ought now to be regarded as approximately true. In practice, any realistic scientist ought only to rank a theory as higher than its competitors with respect to approximate truth.

There is much more in Lipton's superb book than I have been able to address here including a solution to the paradox of the ravens, an account of how prediction and accommodation of evidence differ with respect to theory confirmation, and a discussion of the base-rate fallacy and the no miracles argument for scientific realism. Lipton continues to illuminate many aspects of epistemology and philosophy of science.

Department of Philosophy
University of Bristol
Bristol, UK

By Igor Douven

When it appeared in 1991, Peter Lipton's book on Inference to the Best Explanation (henceforth IBE) was the first in-depth discussion of that rule of inference. The book was nearly complete in that it gave careful attention to almost any issue of relevance to IBE. One thing *not* discussed in that book is the relationship between IBE and the Bayesian approach to confirmation. This was already in 1991 somewhat of a lacuna, given that what then seemed to be the most serious critique of IBE assumed some key tenets of Bayesian confirmation theory (cf. van Fraassen, 1989, Chapters 6 and 7). Over the years, the lacuna has become even more conspicuous, for in the past decade Bayesian confirmation theory has firmly established itself as the dominant view on confirmation; currently one cannot very well discuss a confirmation-theoretic issue without making clear whether, and if so why, one's position on that issue deviates from standard Bayesian thinking. It is thus not an exaggeration to say that, while the second edition of Lipton's book contains much valuable new material, the most important addition is his discussion of Bayesianism and how it relates, or should relate, to IBE. It is on this addition that I will concentrate below.

Although Lipton only briefly mentions van Fraassen's Bayesian argument against IBE, his discussion of the relation between Bayesianism and IBE can usefully be read as an attempt to reply to that argument. Let me therefore recall at least the conclusion of the argument, which is, roughly, that it is incoherent to change one's degrees of belief via any rule other than Bayes's rule (that is, the rule according to which one's new degree of belief in A after learning B should equal one's degree of belief A conditional on B before one came to know the latter (for any A and B)), and that hence IBE is either incoherent or, if it simply boils down to Bayesian updating, redundant. Lipton seeks to slip between the horns of this dilemma by arguing that, while IBE is not a competitor to Bayesianism, it would be wrong to say it boils down to Bayesianism; IBE has something valuable to offer to the Bayesian that she can accept without being forced to abandon the Bayesian framework. In Lipton's terms, not only *can* the explanationist – his word for someone thinking explanation is of confirmation-theoretic relevance – and the Bayesian be friends, they *should* be friends. I take this to mean, at least, that Bayesians should (also) be explanationists (though strictly speaking people can of course be friends and yet disagree about almost everything).

Before considering whether there is something to be gained by the Bayesian by being an explanationist, let us first ask what it could *mean* for a Bayesian to be an explanationist. In order to apply Bayes's rule and determine her probability for A after learning B, the Bayesian agent will have to determine the probability of A conditional on B. For that she needs to assign unconditional probabilities to A and B as well as a probability to B given A; the former two are mostly called prior probabilities of A and B (or just priors), the latter the likelihood of A on B. (I'm here following the "official" Bayesian story. Not all of those who sympathize with Bayesianism adhere to that story. For instance, according to some it is more reasonable to think that conditional probabilities are "basic" and that we derive unconditional probabilities from them; cf. Hájek, 2003 and references given therein. Others think that we often start calculating probabilities with a mix of basic unconditional and basic conditional probabilities; see Uffink and Douven, 2003). How is the Bayesian to determine these values? As is well known, probability theory gives us more probabilities once we have some; it does not give us probabilities from scratch. Of course when A implies B or the negation of B, or when A is a statistical hypothesis that bestows a certain chance on B, then the likelihood follows "analytically" (the claim assumes Lewis's 1980 Principal Principle and whether this principle is analytic is controversial; hence the scare quotes) but this is not always the case, and even if it were, there would still be the question how to determine the priors. This is where IBE might have a role to play. And it is indeed Lipton's main suggestion that a Bayesian should determine her prior probabilities and, if applicable, likelihoods on the basis of explanatory considerations. A Bayesian who does determine them in this way might be called an explanationist Bayesian (or, if you like, a Bayesian explanationist).

In order to assess properly the value of Lipton's suggestion, we must be clearer about how exactly explanatory considerations ought to, or even could at all, guide one's choice of priors and likelihoods. While Lipton's informal style makes his book accessible to a wide audience, I feel that the chapter about Bayesianism in particular would have benefited from a slightly more formal approach. At a minimum, I would have liked to see a more precise account of how one is to base one's probabilities on explanatory considerations; Lipton repeatedly says *that* such considerations should guide the determination of probabilities, but is rather unspecific about *how* they are to guide that process.

Let me start by noting that the answer to this question is not as obvious as one might at first think. Suppose I am considering what priors to assign to a collection of rival hypotheses and I want to follow Lipton's suggestion. How am I to do this? An obvious – though still somewhat vague – answer may seem to go like this: Whatever exact priors you are going to assign, you should assign a higher one to the hypothesis that explains the available data best than to any of its rivals (provided there is a best explanation). Note, though, that my neighbour, who is a Bayesian but thinks explanation has nothing to do with confirmation, may well assign a prior to the best explanation that is even higher than the one I assign to that hypothesis. He may even consistently, and not just in the present case, do this, not because in his view explanation is somehow related to confirmation – it isn't, as I said – but, well, just because. In this case, "just because" is a perfectly legitimate reason, because *any* reason for fixing one's priors counts as legitimate in a Bayesian context. According to standard Bayesian epistemology, priors (and sometimes likelihoods) are up for grabs, meaning that one assignment of priors is as good as another, provided both are coherent (i.e., obey the axioms of probability theory). (The situation would be different if the once popular Principle of Indifference were a requirement of rationality. Roughly stated, that principle counsels that, absent a reason to the contrary, we give equal priors to competing hypotheses. Given this principle, "just because", among others, does not count as a valid reason for assigning a higher prior to the best explanation (but explanatory considerations might). As is well known, however, the Principle of Indifference may lead to inconsistent assignments of probabilities and so can hardly be advertised as a principle of rationality; see, e.g., Gillies, 2000, Chapter 3.) More generally this means that for any Bayesian agent A who takes into account explanatory considerations and who consistently assigns, on the basis of those considerations, a higher prior to what she regards as being the best explanation in a given case, there will be or at least may be a Bayesian agent A' who thinks explanation is orthogonal to confirmation and who consistently assigns a prior probability to the best explanation that is at least as high as the one A assigns to it. Now Lipton's recommendation to the Bayesian to be an explanationist is apparently meant to be entirely general. But what should my neighbour do differently if he wants to follow the recommendation? Give the same prior to any best explanation that I, his explanationist neighbour, give to it, that is, *lower* his priors for best explanations? Or rather give even higher priors to best explanations than those he already gives? It is not clear from what Lipton says what the answer should be.

In response, Lipton might say that the recommendation is not really meant for those who already assign higher priors to best explanations, even if they do so on grounds that have nothing to do with explanation; as long as one does assign higher priors to those hypotheses, everything is fine, or at least finer than if one does not do so, regardless of one's reasons for assigning those priors. The answer to the question how explanatory considerations are to guide one's choice of priors would then presumably be that one ought to assign a higher prior to the best explanation than to its rivals, if this is not what one already does (if it is, one should just keep doing what one is doing).

That still leaves the question *why* Bayesians ought to follow the recommendation. One can think of various ways in which a Bayesian might do better by following it than by not doing so. For instance, by following the recommendation she might stand a greater chance of eventually assigning the highest posterior to the truth, or it might lead her to assign a high probability to the truth more rapidly. But these are just possibilities. That one is really helped by the recommendation in either of the aforementioned ways, or perhaps in some third way, seems an empirical claim, and it is not one for which Lipton offers any support. (It is not even entirely clear how to investigate this and other (see below) empirical claims Lipton makes in the book. In doing so, are we to employ Bayesian confirmation theory? If so, should we let explanatory considerations play a role (in whatever way)? This is not to suggest that an empirical investigation of the claim that Bayesians are better off by also being explanationists is impossible, but just that such an investigation is likely to be an intricate affair (see Douven, 2005 for more on this).)

Let us turn, then, to another and, in my view, more interesting suggestion that Lipton makes to flesh out the slogan that explanation is a guide to determining probabilities; namely, the suggestion that IBE can serve as a heuristic to determine, even if only roughly, priors and likelihoods in cases in which we would otherwise be clueless and could do no better than guessing. This suggestion is sensitive to the well-recognised fact that, standard Bayesian thinking to the contrary notwithstanding, we are not always able for any hypothesis to assign a prior to it, or to say how probable a given piece of evidence is conditional on a given hypothesis. Consideration of that hypothesis' explanatory power might then help us to figure out, if perhaps only within certain bounds, what prior to assign to it, or what likelihood to assign to it on the given evidence.

Bayesians, especially the more modest ones, might want to respond that the Bayesian procedure is to be followed if, and only if, either (a) priors and likelihoods can be determined with some precision and objectivity or (b) likelihoods can be determined with some precision and priors can be expected to "wash out" as more and more evidence accumulates or (c) priors and likelihoods can both be expected to wash out. (It is sometimes said that priors will *always* wash out, but that is wrong; under certain specific conditions they do wash out, however. For a detailed discussion of these matters see Earman, 1992, Chapter 6, where it is also explained that likelihoods can wash out, too.) In the remaining cases – they might say – we should simply refrain from applying Bayesian reasoning. *A fortiori*, then, there is no need for an IBE-enhanced Bayesianism in these cases. And some incontrovertible mathematical results indicate that, in the cases that fall under (a), (b), or (c), our probabilities will converge to the truth anyhow. Hence in those cases there is no need for IBE either.

Here the rejoinder could be that explanatory considerations *are* a route, though one so far typically neglected by Bayesians, to the (more or less) precise and objective determination of probabilities. But if they are, then that is certainly not *a priori*. And again Lipton's book cites no empirical evidence to believe this claim is true.

We can consider yet another way in which according to Lipton IBE might be of help to the Bayesian. Psychological research in the past decades has amply demonstrated that we can do quite poorly *qua* Bayesian reasoners, and are prone to commit probabilistic fallacies such as the well-known "conjunction fallacy" (see, for instance, Tversky and Kahneman, 1983). Referring to this research, Lipton makes the intriguing suggestion that IBE may well help us to realize, or at least approximate, the Bayesian reasoning in situations in which we are apparently bad at doing it "directly". Here, too, he adduces scant evidence that could support the claim, but let us nonetheless grant that it is correct. Then, while in my view the suggestion is intriguing, as I said, I am not sure how interesting it is from a purely philosophical perspective (as opposed to from a psychological perspective). After all, most philosophers seem to take an interest in Bayesianism or IBE (or both) principally in the contexts of scientific and philosophical reasoning; what rules, if any, we follow in our daily lives seems to be much less of a concern to them. But scientists and philosophers may be expected to be able and willing to do the math required for Bayesian reasoning (which often is of a rather elementary nature), at least when it matters to their research; at a minimum, they will know how to let their computers do the math for them. So, in the contexts of interest to (most of) us, there seems to be no call for a kind of shortcut that helps people realize or approximate Bayesian calculations; in those contexts, people simply do the calculations. (This is not to deny that some mathematical problems involved in Bayesian reasoning may be hard to solve even with the help of a computer. But in those cases no one will expect much from any explanationist shortcut.)

I conclude that while Lipton seems right that explanatory considerations can have a role in Bayesian reasoning, and is to be commended for pointing to various ways in which such considerations *might* be helpful, it would seem rather premature to claim that they *should* have a role there. The normative claim still awaits an empirical underpinning that Lipton has not tried even to begin in his book.

Department of Philosophy
Erasmus University
Rotterdam, The Netherlands

By Bas C. van Fraassen

Lipton's valuable monograph has already seen wide use as a text and received considerable attention in the literature. This revised new edition provides a welcome extension, taking account of some of the difficulties, as well as adding new material, including a entire chapter on the Bayesian approach to our epistemic and doxastic life. While engaged throughout with ongoing controversies in epistemology and philosophy of science, the book is a pleasure to read and easy to use as textbook for both undergraduate teaching and graduate seminars.

Let me say at once that I'm not a neutral reviewer. Lipton defends Inference to the Best Explanation (henceforth IBE) and offers critiques of my some of my arguments for rejecting it. I will not respond to these here, but I will offer some critique in return.

The first edition was reviewed in a number of philosophical journals. Some, such as Hobbs' in *Philosophy of Science*, contain serious and far-reaching criticisms which clearly haven't intimidated Lipton, who remains steadfast in his convictions. I do not want to repeat commentary already so easily available, so will focus on just three items. The first is the methodology announced in Lipton's Prefaces, the second his contention that "the Baysian and the explanationist should be friends", and the third is the question broached again in the revised Chapter 9, "is the best good enough?".

LIPTON'S METHODOLOGY

In the second edition's Preface Lipton emphasizes that what he intends to do remains "to explore our actual inferential practices by articulating and defending the idea that explanatory considerations are an important guide to inference". The phrase "inferential practices" conveys a conviction which I'll address in a moment. "Important" could just mean "large", but it soon becomes clear that it acts here in an endorsing role, for Lipton means to defend as well as explore (as he says again at the beginning of his Conclusion).

This does not sit easily with the intention expressed in the (still included) Preface to the first edition. To avoid entering in this monograph upon Hume's problem about induction (presumably the subject of his forthcoming *The Humean Predicament*), Lipton distinguishes two tasks in epistemology, justification and description: "Even if we cannot see how to justify our inductive practices, surely we can describe them". More importantly, so much philosophical discussion in this area has leaned on assumed simplistic characterisations of how we manage our opinion that one might well despair of any progress until the philosophers' fables are replaced by some serious investigation of this sort. So the intention is admirable, but I think its announcement is somewhat misleading.

Reviewers have tended to focus on the defence and seen it as such (e.g. Vogel, calls the book a "sustained articulation and defence" of the view that "explanatory success provides a good reason to believe a theory"; similarly Harman). But we should not just take the descriptive enterprise in our stride. Lipton says that it is amazingly difficult to give "a principled description of the way we weigh evidence". But does he really make a serious effort to describe with some neutrality on the issues? The language is loaded. Could there really be such a distinct task as "descriptive epistemology" in a form that fits his practice? What does the term "actual inferential practice" refer to, does it signal an empirical thesis about how we manage our opinion (by inference rather than in some other way)? What qualification is signalled by "important" in "important guide" or "principled" in Lipton's "principled description"?

To avoid equivocation, let us use "induction" to refer to what we all do, which is to form opinion that goes beyond our evidence. That includes logical errors as well as consistent extrapolations, superstition and stereotyping as well as acceptance of scientific theories, leaping to conclusions as well as careful experimental design. Let us use a capital letter, "Induction" to refer to what philosophers discuss under this heading. Induction is a certain practice of induction subject to rules, norms, or principles of right reason, which can be formulated with some degree or other of precision. What exactly it is, and whether it ever actually occurs in the wild, is usually not clear. My worry about such phrases as "actual inferential practices" is that they bring along presuppositions linking induction and Induction.

If Induction were specified precisely, then a sociologist or psychologist could describe a given community as engaged or not engaged in it. Without such a specification, an empirical scientist could still describe induction as it occurs in various communities.

In either case the findings might or might not include signs of constraint by this or that rule in some proportion or type of cases. Now that would be description prior to evaluation! A long tradition of discussions of induction in philosophy assumes that we manage our opinion by a process that is in some significant, nontrivial sense one of inference. But to assume that is already to set the terms of debate a priori.

Lipton does not engage in any sort of empirical inquiry. What would be the pertinent antonym of "descriptive" in this context? If a philosopher applies normative concepts to examples of reasoning and opinion management – and does not just describe us or a society as engaged in applying or conforming to them – we may aptly call him or her engaged in "normative" epistemology. Is Lipton?

He often expresses his conviction that we are very good at induction. This presupposes that there are standards of success for induction, hence that it is an activity with a certain aim. Therefore this conviction already involves the application of normative concepts to examples of reasoning and opinion management. In fact, if I am not mistaken, Lipton means that we succeed in attempts to reach true conclusions and reliable, well calibrated opinion. A more modest assumption would be that we have not yet done so badly as to be in danger of extinction, or that our inductive behaviour leaves us faring pretty well in our current ecological and environmental niche. Of course, an author is entitled to set out his starting point as he chooses. But this particular starting point gives one the uneasy feeling that Lipton has already decided that induction is largely Induction and that Induction is reliable. The first is an empirical claim and the second an evaluative one, both made beforehand, keeping us inside a framework that we have by now serious reason to distrust.

In fact, Lipton is very much focused on defence, in a context in which the description is largely provided by past philosophical discussion, and pretty well taken for granted as correct. The book presents us not so much with a "principled description of the way we weigh evidence" as with a description of what Lipton sees as principled ways to weigh evidence. The subject is not so much the practices we have or norms we follow but the norms we should have.

SHOULD THE EXPLANATIONIST AND BAYESIAN BE FRIENDS?

In his chapter on the Bayesian approach Lipton does describe Kahneman and Tversky's empirical studies, mentioning them as objections to that approach. Those studies are genuinely descriptive, and purport to show that we are not actually managing our opinion by the Bayesian recipe of conditionalisation (purely logical updating, the probability analogue of Modus Ponens). For Lipton these studies provide evidence that we tend to infer to the best explanation rather than conditionalize – and he takes this as support for IBE. But support for what, precisely? Not, surely, for advocating the use of IBE! Perhaps it is support for an empirical thesis about our behaviour: see, we are IBEing all the time, that is what we do. To be adequate that thesis would have to describe us as overwhelmingly given to practices that ought to be corrected. Taken by itself, such conclusions as Kahneman's that we tend to go for spurious causal explanations when we don't understand – such as regression to the norm – simply place IBE in bad company.

While not making this same point, Lipton tries to counter such criticisms implicitly in three ways. The first is by means of empirical assertions of his own, without empirical evidence, in the absence of any relevant empirical studies. These include the repeated assertions, begun already in the Preface, that we are actually very good at making reliable ampliative inferences, even if we can't explain how we do it. The second is that the sort of clearly unreliable inferences charted by Kahneman and Tversky etc. are to be found in cases of simple and artificial character, and that we do much better in more complicated, real life situations. Psychologists should take Lipton's opinion on this as an empirical challenge, and devise more sophisticated tests. Before they do, we have no good evidence one way or the other. Our actual success in daily life may be largely due to remaining in an environment to which our inductive behaviour happens to be adapted. Stereotyping and prejudice may well have serious survival value for epistemically challenged agents. On the other hand, our more striking and important successes may be due to the sciences, where great precautions are taken to ensure that demonstration is by acceptable statistical methods instead.

This second counter seems to me to actually involve Lipton in an inconsistency. At the beginning of the chapter he outlines the Bayesian arguments to the effect that one should update one's opinion only by conditionalisation. As De Finetti explained so clearly, optimising one's expectation of gain or loss, on the supposition that there will be a penalty depending on inaccuracy in one's updated opinion, demands conditionalisation. Lipton offers no critique of these arguments. But he points out that one does have the option of changing one's prior opinion if faced with new evidence. The analogy is to Modus Ponens: if I believe that if A then B, and find that A, then I have the option of adding B to my beliefs or to delete that prior belief in the conditional. So he suggests that if conditionalisation of one's prior subjective probabilities would conflict with an inference to the best explanation, one has the choice to revise one's prior so as to favour the more explanatory hypotheses.

This combination of respect for De Finetti's argument with the admission of this friendly cooperation between the two camps seems to me straightforwardly inconsistent. For suppose that in such a situation where the prior is P, and conditionalisation on the total new evidence would yield P', Lipton points to the possibility of changing P to P* before conditionalising. Then the prior expectation of gain or loss, on the supposition that there will be a penalty depending on inaccuracy in one's updated opinion, will favour P' over P*, as well as over any conditionalisation of P*. Lipton could have attacked De Finetti's and similar Bayesian arguments, but he did not. He cannot let them stand and still propose that "the Bayesian and the explanationist should be friends" in that way.

The third counter, less explicitly, is to suggest that in general prior probabilities have to be constructed, we do not have them already, and that regard to explanatory value can (does? must?) play a role in this construction. That suggestion does not make a friend of the orthodox Bayesian of course, but it is certainly a familiar line among more liberal probabilists. I am very sympathetic to it – it is undoubtedly the correct first step out of the orthodox Bayesian straightjacket. But what argument is there to endorse the use of IBE (in some not spelled out version that could do this job) in such a construction? Again it seems to be that Lipton equivocates on the tasks of description and defence. One task could be the naturalistic one, as urged by Quine, that coincides with what empirical psychologists attempt. Do people construct prior probability functions, however vague, when faced with a new problem situation? Do they use explanatory considerations, if so? A presumed conviction that we all do things that way should have no place here. But suppose we find that people do things that way. How shall we then, as second task, assess the likeliness of priors thus constructed leading to well calibrated opinion? Supposing that we, looking on, have our probabilities in place, what should be our expectation values for the constructors' success?

If Bayesians have relevant subjective probabilities already in place, and see someone constructing his or her subjective probabilities in the manner sketched above, they will in general have a low expected value for that agent's calibration. Poor explanationists, how can they be friends with someone who looks at them this way?

IS THE BEST GOOD ENOUGH?

An explanation may have many virtues that we can recognize quite independently of its truth or falsity. Newton's theory of gravitation explained the tides – that could be claimed, and can still be claimed, without implying that this theory was true. But it was lovely ... that is Lipton's apt term for a positive evaluation of this sort. The word "best" in "inference to the best explanation" must refer to that, since it is the conclusion of that inference which asserts that a given candidate is true, going beyond the premise that is the best.

Why should the loveliness of an explanation make it more likely to be true? Lipton calls this *Voltaire's question* and devotes a chapter to it. One of that chapter's sections has as title the subsidiary question "Is the best good enough?". Inference to the best explanation consists in allocating one's belief to the best among those explanations which are available. If "better" here implies "more likely to be true", then indeed, obviously, the inferred conclusion is likely to be true provided the truth is in the range of available candidates. But why should we think that the latter provision holds?

There is one obvious answer: it must be part of the rule of IBE that we are to apply it only in cases where we believe the truth to be very likely in that range. Taken in itself this is a very good answer. No one, however critical of IBE, can deny that a choice of the comparatively likeliest hypothesis in a set which very likely contains the truth, will most likely yield a true conclusion! But this answer is a probabilistic one. In this answer, the role played by the agent's prior judgment of what seems more or less likely – the prior state of opinion – pre-empts that which proponents of IBE want for their rule. We don't need a special rule when a tautology will do!

Lipton circumvents this point by taking the question in a special form, which allows one to draw on auxiliary views about scientific practice. Assume that in their comparative ranking of available alternative hypotheses scientists are very good at giving a higher ranking to the ones more likely to be true. Might they then yet be very poor at choosing to apply this procedure just to ranges of alternatives likely to contain the truth? Lipton begins with some quick replies (as he himself calls them) in support of a negative answer. He does not take them to suffice even for this very special case of presumed experts in IBE at work. But he has two replies that he does take to settle the matter.

The first appears to be a logical point: "The nub of the argument is the claim that there is an unbridgeable gap between comparative and absolute evaluation. This gap is, however, only a plausible illusion" (p. 155). What could establish this? Imagine that you want to evaluate a given theory T1 against its negation, ~T1. We assume only that you can rank contraries as to how likely they are. A contrary T2 of T1 will imply ~T1; therefore if we rank T2 as more likely than T1, we automatically rank ~T1 as more likely than T1 as well. On the other hand, a contrary T3 of ~T1 will imply T1, so if we rank T3 as more likely than ~T1, we also rank T1 as more likely than its negation.

What follows from this? Take Eddington's famous expedition to test the General Theory of Relativity (T1). Its contrary T2 to be considered – Newton's theory of light and gravitation – implies a different deflection of starlight near the sun. Is there also a contrary here of ~T1? Certainly; it is T1 itself. In this case, the predictions of T1 did fit the evidence much better than did those of T2, so we can take it that there was a comparative ranking. But how does this affect the question of whether T1 was likely to be true? Logically speaking there are many hypotheses concerning the deflection of the light that fit the data better than either of these contraries. They are not considered since they are not implied by any theories under consideration. Hence a comparative ranking is not precluded, but no logical legerdemain will yield an absolute ranking of T1. Eddington's elation seems more plausibly ascribable to his prior absolute and conditional personal probabilities, with no need for IBE. What is missing from Lipton's recourse to logic is specifically a premise about how the range of alternatives considered, T1, T2, T3, relates to the facts – i.e. the very point at issue.

Lipton then offers a second argument, sounding near to transcendental. The comparative ranking of the considered alternatives is made under the guide of a background theory. If scientists are (as assumed in this special scenario) so good at comparative ranking, what guides them must be relevantly good. But that means that their background theory must be probably true.

> Scientists rank new theories with the help of background theories. According to the ranking premise... this ranking is highly reliable. For this to be the case, however, ... these theories must be *probably true*, or at least probably approximately true. If most of the background theories were not even approximately true, they would skew the ranking, leading in some cases to placing an improbable theory ahead of a probable one ... (p. 157)

Lipton concludes that "if scientists are highly reliable rankers ... the highest ranked theories have to be absolutely probable" (p. 158). Thus the best selected will not likely be the best of a bad lot, and IBE is likely to work very well indeed.

This argument pertains to the special case Lipton has chosen to examine, in which we can draw on some knowledge as well as assumptions about how the ranking is done. The assumption is that it is a ranking which places more probable alternatives above less probable ones. The extra knowledge drawn on is that scientists are guided in their actual ranking by background theories. Well, let us look a bit further into this actual ranking, and the assumption. In what sense is the ranking highly reliable? In that it ranks as more probable (usually) those which are more probable. What concept of probability is invoked here? The ranking expresses the scientist's probabilistic judgments, which De Finetti would be as happy to discuss as anyone. But presumably the assertion of reliability does not mean that the ranking is reliable according to the scientists' own judgement. That would trivialize it: coherence requires that they expect their own probabilities to be well calibrated. So does the ranking assumption bring with it a belief in some external standard, such as objective chance, which the ranking is meant to match? What exactly was the content of this ranking assumption?

How are the scientists assumed to perform their ranking? What precise role does the background theory play in the ranking activity, as opposed to the selection of the range of alternatives? In the Eddington example I took it that the ranking was based on how well the hypotheses fit the data obtained. Visible success would then consist similarly in continuing to fit the data, or fitting it better than the competitors do. The data, however, reveal but a very small aspect of nature. If guidance by the background theories leads to reliable ranking in this sense, one wonders by what inference they are concluded to lead to hypotheses more likely to be true (as opposed to empirically adequate, say). I imagine that Lipton thinks of the background theories' truth as the best explanation of their empirical success, but he can't very well offer that here as his reason for the inference.

Lipton relates the issues in this chapter, as well as in some of the others, to recent debates about scientific realism. In fact this chapter ends with a careful and admirably even handed assessment of the extent to which his conclusions here can affect that topic. He has made it very clear that he is not sympathetic to alternatives to scientific realism. At one point he suggests for example that a constructive empiricist must be involved in such simple self-contradictions as "my computer is on the table, which is a swarm of particles, but there are no particles" (p. 146). He carefully points out however, at many places in the book, the limited reach of his arguments and he resists simplistic overall conclusions.

In the last two pages of the book Lipton says with frank and appealing modesty that his conclusions "can be used to provide some arguments for adopting a realist rather than instrumentalist stance toward scientific theories" but that he cannot apply IBE as vehicle for the "no miracle" argument for scientific realism, nor for dismissing scepticism concerning Induction in general. We can look forward to his forthcoming book on the latter subject, which will surely be as useful, thorough, original, and readable as this one.

Department of Philosophy
Princeton University
Princeton, USA

Author's Response*

By Peter Lipton

I am enormously grateful to Igor Douven, James Ladyman, and Bas van Fraassen for their reactions to the second edition of *Inference to the Best Explanation*. Having such acute readers is a privilege and they have given me plenty to worry about. I cannot respond fully to all their points here, and not just for reasons of space, but I will touch on a number of the issues they raise in what follows, under three headings. These are the distinction between describing and justifying induction, the compatibility of Inference to the Best Explanation (IBE) and Bayesianism, and the relationship between the relative and the absolute evaluation of scientific hypotheses.

THE DESCRIPTIVE AND THE NORMATIVE

On the face of it, there is a clear enough distinction between descriptive and normative accounts of induction, descriptive and normative accounts of the way we, in van Fraassen's phrase, "form opinion that goes beyond our evidence". A descriptive account says something about how we actually form these opinions, for better or for worse; a normative account says something about how we ought to form these opinions, whatever we actually do, if we want them (say) to be true. I am disposed to take IBE both ways, both as a relatively good description of some aspects of the way we actually form opinion and also as a relatively good way to form opinion. But one could endorse IBE for only either type of account alone. Thus one might accept that explanatory considerations are in fact a guide to forming opinions that go beyond our evidence but regret the fact, holding that this practice is on balance ill-suited to our cognitive goals. Conversely, one might hold that explanatory considerations play little or no role in the way we form our opinion but that we would do better, cognitively speaking, if they did.

My book sometimes considers directly the normative prospects of IBE, for example in my discussion of the no-miracle argument

*Acknowledgement: My thanks to Paul Dicken, Igor Douven, James Ladyman, and Bas van Fraassen for comments on this essay.

for scientific realism (which I suggest is alas unsuccessful as it stands) and in my discussion of the bad lot argument against supposing that the best explanation is likely to be true. The book's primary focus, however, is supposed to be descriptive: I am trying to strengthen the case for saying that explanatory considerations are a guide to inference, for better or for worse. Van Fraassen has a number of misgivings about the way I pursue the descriptive project. One is that a neutral description of induction is a matter for empirical inquiry, for psychology or sociology, not for armchair philosophy. Another is that my descriptions are so loaded with normative presuppositions that my actual subject turns out to be "not so much the practices we have or norms we follow but the norms we should have".

The descriptive question, the question of how induction actually works, is indeed an empirical question and a multidisciplinary project, and we ought to look to psychologists, sociologists, historians, and cognitive scientists for help. And although I do occasionally appeal to the empirical – for example in my discussion of Semmelweis's investigation of the etiology of childbed fever and in my use of some of the empirical results of Kahneman and Tversky's research into "systematic irrationality" – my book is rather conventionally philosophical in tone and method. So I can see how it might give the impression either of the normative in descriptivist clothing, or of an exercise in indoor ornithology. But why can what I do not count as a contribution to the descriptive project? That project depends on getting clearer about the content of the account under investigation, in this case clearer about the "Best", and "Explanation" (and perhaps "Inference", as van Fraassen suggests) in "Inference to the Best Explanation". These are squarely philosophical tasks. And these are not the only things philosophers can do in this neighbourhood: they can also help to assess the descriptive adequacy of various accounts of induction, including philosophical accounts with a strong normative flavour.

One of the reasons why a philosopher can sometimes assess the descriptive adequacy of philosophical accounts of induction without detailed empirical investigation is that those accounts turn out on inspection to be egregious descriptions. If one can show that an account fails to discriminate between the evidential force of black ravens and white shoes, or that it entails that everything consistent with a hypothesis provides a reason to believe it, or that data that are known before a hypothesis is constructed can provide no reason to believe it, then one can show that there is a serious descriptive difficulty without engaging in serious empirical inquiry. And if an alternative account can be shown to avoid some of the blatant descriptive difficulties of other accounts, that is an argument for its descriptive superiority. And that is so even if, say for Humean reasons, it proves impossible to show that one's favoured descriptive account describes a process that is likely to take one to correct opinion.

Though no replacement for other forms of inquiry, and fallible too, traditional philosophical appeals to intuition or self-observation – for example considering whether one would have been willing to make a given inference as described – can contribute to the descriptive project. To rev this up a bit, philosophers can engage in Goodmanian "reflective equilibrium", playing off their intuitions about specific cases and about general principles or mechanisms, searching for contradictions, and then revising for consistency. But wait a minute. Does this not make van Fraassen's point, that my descriptive pretensions are really just a veneer for my normative claims? Although Ladyman does not raise this as a criticism, he remarks that if reflective equilibrium is taken to be a canon of rationality, then someone who reasons in a way that accords with the outcome of that process could not count as systematically irrational. The normative is thus built in from the start.

I agree that the descriptive and the normative projects are linked in various ways. My method for coming up with a description of induction includes looking for an account that captures what strike me on reflection as good inferences and excludes what strike me as bad ones. So if I find such an account, it is guaranteed to meet some of my own norms. But this sort of linkage does not compromise the descriptive project. Indeed it would not do so even for those who maintain, in an attempt to dissolve Hume's sceptical problem, that the rationality of our actual inductive practices, whatever they are, follows immediately and analytically from the definition of "rational". Such philosophers could still look to our own judgements as an aid to saying how people judge generally. As it happens, I do not myself accept that reflective equilibrium packs as great a normative punch as this. We endorse the inductions we endorse, and our intuitions have been tempered by the requirements of consistency, but it remains all too easy for me to step back from my practice and worry compulsively about the Humean problem of justifying any consistent inductive policy over any other. A policy may pass the test of reflective equilibrium, yet still seem unwarranted from either an externalist or an internalist point of view, since it may yet be unreliable and it may not have been certified in a way that avoids an illicit circularity of the sort that Hume warned us against. But whatever the normative strength of reflective equilibrium, it may yet be an effective means of prosecuting the descriptive project. Bring on the human sciences by all means – we need all the help we can get – but intuitions about what one would think are not to be disdained, certainly not when the project centrally includes the task of saying how one does think.

In addition to the general worry that my descriptive project may be tainted by my prior normative commitments, van Fraassen is concerned that a particular presupposition I make may well be false, namely that our undeniable practice of forming opinions that go beyond our evidence is a practice of *inference*. The point, I take it, is that to count as inferential a practice would have to be strongly rule-governed, and it is not at all clear that our non-demonstrative practices are so governed. This doubt is I think intimately related to van Fraassen's epistemic voluntarism and to his negative view of rationality as something that specifies only what is not allowed, leaving very considerable freedom over the expectations one forms on the basis of experience. Thus, although he displays brilliantly the virtues of his constructive empiricism over realism, van Fraassen does not accuse the realist of irrationality, since all parties may be expected to observe the principles of deductive and probabilistic consistency, which is pretty much all that rationality requires. Whether this conception of rationality is sufficiently thick is a interesting question, but either way there is I think no special problem for the descriptive project. For one may accept the thin conception of rationality, yet also maintain that explanatory considerations help to account for the way we bridge the gap between the principles of rationality and our actual expectations. Moreover, the core idea I explore in my book – that explanatory considerations are a guide to forming expectations – seems compatible with the denial of a strongly rule-governed inferential system. Illuminating and reasonably pervasive heuristics would be glory enough. The voluntarist and the explanationist can be friends.

BAYESIAN ABDUCTION

For this edition of my book I decided that I could no longer put off some consideration of the relationship between IBE and Bayesianism. Here my primary goal is even less than descriptive. It is to show that these two approaches to induction are compatible and so that one could in good conscience be both a kind of Bayesian and a kind of explanationist. At the same time, it is also my hope that the case for compatibility makes some contribution towards both the descriptive and normative credentials of IBE, by showing how it might fit naturally with another account of belief revision that itself has considerable independent descriptive and normative clout. The main thought is that explanationist considerations might be part of the way people "realise" the Bayesian transformations. Ladyman captures this well when he writes that "the cognitive realisation of Bayesian reasoning involves thinking about probabilities indirectly via thinking about explanations". Thus I suggest among other things that we may think about the explanatory credentials of a hypothesis as an aid to fixing its prior and its likelihood.

Douven makes a number of significant observations about my discussion of the relationship between Bayesianism and IBE. For example, he usefully distinguishes between the idea that IBE is a heuristic for working out priors and saying that we should in fact give better explanations higher priors. These are indeed distinct claims, though I am favourably disposed towards both. He is also quite right to say that although better explanations might on that basis be given higher priors, there is no probabilistic obligation to do so, and indeed that I have offered no direct argument here that to do so would be more likely to get us to the truth. That is correct: the primary aim of my discussion of the relationship between Bayesianism and IBE is far less ambitious, and Douven is right to emphasise this.

My discussion of Bayesianism makes some appeal to well-known work of Daniel Kahneman and Amos Tversky that seems to show that real people do not always obey even the most basic requirements of probabilistic consistency. For example, they show how easy it is to get subjects to rank the probability of a conjunction ahead of the probability of one of its conjuncts. The use I make of these results is, as van Fraassen observes, delicate in the context of my overall strategy; but it is not I think incoherent. I make two main points. The first is that these empirical results suggest that we do not find it all that easy to perform the sort of probabilistic calculation that Bayes's theorem represents. The second is that some of the cognitive errors that Kahneman and Tversky discuss may be accounted for by supposing that we sometimes let our desire for good explanations get out of hand and that we sometimes pay insufficient attention to other relevant information, such as information concerning base rates. My motive for making these suggestions is to make more plausible the idea both that we need help in performing Bayesian calculations on that we may well use explanatory thinking to this end.

The reason my strategy here is delicate is because, as I have said, my overall goal is to show the compatibility between Bayesianism and IBE; yet the Kahneman and Tversky cases seem precisely cases where the two ways of thinking diverge. But I take it that this divergence, though genuine, is the exception rather than the rule, and that if we see explanationist thinking as in part a way of approximating values that should go into the machinery of Bayesian conditionalisation, we can take on board the possibility that these approximations are not always good ones. Moreover, I take it that the cases where we do go astray, though psychologically natural, are in no way compelled by an approach that endorses both IBE and Bayesianism. But Douven and van Fraassen are both quite right to observe that I have not shown (nor have I attempted to show) how IBE would help us to avoid the mistakes in reasoning that Kahneman and Tversky document. Indeed one might go further here, and take results to suggest that it is not only Bayesianism that might be aided and abetted by IBE, but that perhaps, now from a normative point of view, IBE should be tempered by respect for the constraints that Bayesianism imposes.

MAKING THE BEST OF A BAD LOT

One place in my book where my discussion is directly normative is in the attempt to reply to van Fraassen's bad lot argument, and Ladyman and van Fraassen have both raised important objections to my reply. The nub of the bad lot argument is that even if one granted that we know that a scheme such as IBE enables us reliably to rank the competing hypotheses we have thought up in terms of comparative probability, there does not follow any reason to believe that the winning hypothesis is outright probable – that it is more likely than not – since we have been given no reason to believe that any of the hypotheses we have generated is likely to be true. Likeliest does not entail likely, and more generally it seems that a comparative probabilistic evaluation does not entail an absolute evaluation.

One of my arguments is that actually comparative evaluation of competing hypotheses does entail some absolute evaluation. For example, if we know that T1 and T2 are contraries and that T1 is more likely than T2, then we can also make the absolute evaluation that ~T2 is more likely than T2; that is that ~T2 has a probability greater than .5. For if T1 and T2 are contraries then T1 entails ~T2, and so ~T2 must be at least as likely as T1. The logic is unexceptionable, but van Fraassen objects that it is irrelevant, because what we want to be able to say is that T1 – the best of the lot – is more likely than ~T1, and this we still cannot say. Well, van Fraassen's logic too is unexceptionable: showing that T2 is probably false is not to show that T1 is probably true. And this is important: my logical point clearly cannot in itself show that if we know that a theory is the best we have, then we know that it is likely to be true. But it does show that comparative evaluation yields absolute evaluations – in this case that T2 is probably false and that ~T2 is probably true. Indeed it shows us that every hypothesis ranked below the winner has a probability of less than 0.5. Moreover, since T2 may itself be a negative claim, for example that there are no entities of a certain type, the contrary-contradictory move shows that unrestricted powers of ranking could yield even positive knowledge of the probable existence of unobservable entities.

Ladyman has a different objection to the move from contraries to contradictories, which is that it fails for approximate truth, since members of a pair of contraries may *both* be approximately true. To assess the impact of this point, we might construct a new ranking assumption in terms of approximate truth rather than in terms of probabilities. The assumption that we are reliable rankers then becomes the claim that we tend reliably to rank theories in order of how well they approximate the truth: better approximations tend to be ranked ahead of worse approximations. The question then seems to be whether the fact that T1 is closer to the truth than T2 shows that ~T2 is closer to the truth than T2. Since T1 entails ~T2, ~T2 must be at least as probable as T1, but must ~T2 be at least as close to the truth as T1? We don't really understand approximate truth, but the answer nevertheless seems clearly to be negative. One reason for this is that specificity is good for verisimilitude but bad for probability. If my bike weighs exactly 16 kg, to say that it weighs 15 kg is closer to the truth than to say that it weighs between 13 and 15 kg. So I accept Ladyman's point. The move from contraries to contradictories does indeed show that comparative evaluation yields absolute evaluation when the evaluation is in terms of probability, as was supposed in the original form of the bad lot argument; but it does not show this when the evaluations are rather in terms of verisimilitude. Ladyman's point is significant because, as he plausibly maintains, realistic realists ought to see scientific claims generally as claims of approximate rather than of strict truth. And this failure of deductive closure for verisimilitude may be expected to throw a spanner in the works of a variety of arguments deployed by realists and instrumentalists alike. But the contrary-contradictory move may yet have some life in it, even in the context of verisimilitude. For example, insofar as we are happy to describe scientists not as ranking with respect to degree of verisimilitude but rather ranking with respect to the probability of claims of the form "T is approximately true" (where there might be some lower bound on degree of approximate truth for a theory to count as approximately true *simpliciter*), then the contrary-contradictory move applies as before, since claims of that form, about approximate truth, would themselves be supposed to be strictly true.

Another objection I run against the original bad lot argument is that to grant reliable ranking is also to grant whatever that requires, and this would include the probable truth or approximate truth of the background theories we use to rank. Since those background theories were themselves generated by ranking, it is incoherent to suppose that we are reliable rankers yet not getting at the truth. Van Fraassen observes that this argument would not work if the only role of the background were to generate the hypotheses that we then go on to rank purely on empirical grounds. A false background might make is more likely that the hypotheses we generate are themselves false, but this would not impugn the reliability of the ranking we go on to perform on this bad lot, if the background plays no role in ranking. My reply is that the background is more influential than this: scientists use it not just to generate hypotheses, but also to assess the ones they generate. Science seems filled with discussions over the ranking of hypotheses that are fuelled by background considerations. Indeed the theories scientists already accept influence ranking in many ways, for example through their roles in experimental design and the interpretation of data, as well as through more direct comparisons of the hypotheses themselves. So it seems that the reliability of ranking depends in many ways on the verisimilitude of background belief.

Department of History and Philosophy of Science
University of Cambridge
Cambridge, UK

References

Douven, I., "Evidence, Explanation, and the Empirical Status of Scientific Realism", *Erkenntnis*, (2005), in press.

Uffink, J. and Douven, I. "The Preface Paradox Revisited", *Erkenntnis* 59 (2003), pp. 389–420.

Earman, J., *Bayes or Bust?* (Cambridge, MA: MIT Press, 1992).

Gillies, D., *Philosophical Theories of Probability* (London: Routledge, 2000).

Hájek, A., "What conditional probability could not be", *Synthese* 137 (2003), pp. 273–323.

Lewis, D., "A Subjectivist's Guide to Objective Chance", in R. Jeffrey (ed.), *Studies in Inductive Logic and Probability* (Berkeley: University of California Press, 1980), pp. 263–293.

Tversky, A. and D. Kahneman, "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment", *Psychological Review* 91 (1983), pp. 293–315.

van Fraassen, B.C., *Laws and Symmetry* (Oxford: Clarendon Press, 1989).

Hobbs, J., Review of *Inference to the Best Explanation*, *Philosophy of Science* 60 (1993), pp. 679–680.

Harman, G., "Review of *Inference to the Best Explanation*", *Mind N.S.* 101 (1992), pp. 578–580.

Lipton, P., *The Humean Predicament*, (Cambridge: Cambridge University Press, forthcoming).

Vogel, J., "Review of *Inference to the Best Explanation*", *Philosophical Review* 102 (1993), pp. 419–421.