Inference to the Best Explanation, 2nd edition

Peter Lipton, *Inference to the Best Explanation, 2nd edition*, Routledge, 2004, 219pp, $33.95 (pbk), ISBN 0415242037.

Reviewed by Lefteris Farmakis, London School of Economics and Stephan Hartmann, London School of Economics

The first edition of Peter Lipton's *Inference to the Best Explanation*, which appeared in 1991, is a modern classic in the philosophy of science. Yet in the second edition of the book, Lipton proves that even a classic can be improved. Not only does Lipton elaborate and expand on the themes covered in the first edition, but he also adds a new chapter on Bayesianism. In particular, he attempts a reconciliation between the Bayesian approach and that offered by Inference to the Best Explanation (IBE).

IBE recently has been championed as a distinctive kind of inductive inference, broadly understood, which has the dual attributes of doing justice to the actual workings of science and the demands for its rational justification. In other words, it is supposed to be both an accurate description of the inferential processes of actual science and also endowed with the property of conferring epistemic warrant to the conclusions reached by means of it. As such, exploring its merits on both counts is of great importance for contemporary philosophy of science. While Lipton touches on justificatory issues, his primary focus is the *descriptive* merits of Bayesianism and IBE, a subject that the author argues has been rather neglected. As its name indicates, IBE is concerned with the notions of *inference* and *explanation*, which are introduced in the first three chapters of the book. On the inference side, Lipton reviews various accounts of inductive inference and notes their shortcomings, which IBE hopes to overcome. On the explanation front, he argues that the absence of a satisfactory theory of explanation need not preclude us from analyzing IBE philosophically. Despite his scepticism about existing philosophical frameworks for explanation, he does argue throughout chapter 3 that a version of the causal model of explanation best supplements the general inferential pattern of IBE. This favoured version is that of contrastive explanation. These explanations are unique in that they do not attempt to answer the question 'Why this?' but the question 'Why this *rather than* that*?*' (p. 33). In other words, they do not employ causal talk to explain an event *simpliciter*, but rather an event together with *the absence* of another event, similar in kind.

Lipton's presentation of these introductory themes is rich, well informed and competent. Yet given the small number of pages, and the volume of material that Lipton addresses, it is quite demanding.

In chapter 4, Lipton embarks on a more systematic treatment of IBE. In particular, he characterizes IBE as inference to the *loveliest* explanation, not inference to the *likeliest* explanation. In this way, Lipton attempts to avoid trivialising his model by assimilating it to those accounts that seek the most *probable* explanation. Instead, he takes the loveliest explanation to be the one that, if correct, provides the most understanding (p. 59). As he avoids this analysis in terms of probability, IBE, as Inference to the Loveliest Explanation, can stand as an independent account on its own right. Finally, after his extensive argument for the distinction between the loveliest and likeliest explanation, Lipton ultimately concludes that if IBE is a

reasonable model for our inferential practices, loveliness and likeliness will tend to be coextensive (p. 61). This claim, though, is not at all obvious but rather seems to require an argument. However that may be, we argue below that a more serious problem for Lipton's project is that his argument actually undermines, instead of promoting, the independent authority of IBE, actually subjugating it to accounts that focus exclusively on likeliness, such as Bayesianism.

Although Lipton believes that IBE is a distinctive kind of inductive inference, he does not express imperialistic inclinations for it over other forms of inductive inference as Harman (1965) and Psillos (2002) do. Nonetheless, he still sets out to defend its normative status by discussing issues of *justification* in chapters 6 and 9-11. Along with the newly added Bayesian chapter, these chapters contain the most interesting arguments. Chapter 6 provides an insightful application of IBE to the raven paradox of confirmation. Lipton also argues that this solution has many benefits over the various standard solutions already in the literature. Briefly put, Lipton's point is that causal-explanatory considerations suffice to rule out confirmation of hypotheses from *prima facie* wildly irrelevant instances by bringing in relevant information about the causal properties of the latter (pp. 95-97). In another interesting application of IBE to an old problem, Lipton discusses the often made distinction between *predication* and *accommodation*. He argues that this distinction can be easily substantiated under his reading of IBE. The core of his solution consists in the hypothesis that in the case of accommodation, the best explanation of the theory's fit with the data may be that the scientist has actually fudged the theory's implications by essentially using the data to be explained (p. 170).

Lipton also spends a significant amount of time defending the claim that IBE is a truth-conducive rule of inference against van Fraassen's (1989) challenge that our best scientific theories may not be good enough to track truth. Van Fraassen's charge is that not only does IBE not guarantee that our loveliest hypotheses are true but rather the converse most likely holds. Lipton's response is considerably forceful here, arguing that van Fraassen's charge is *inconsistent*. In his argument, Lipton makes use of the central role of background beliefs in the workings of IBE. In particular, he argues that if we assume, as van Fraassen does, that our ranking of the explanatory loveliness of theories is reliable, then it follows that our theories are approximately true. Consequently, van Fraassen's argument is inconsistent (pp. 157-159).

Although this result has important consequences for the issue of scientific realism, Lipton postpones the explicit treatment of realism until the last chapter of the book. In this chapter, Lipton examines the merits of the most popular argument for scientific realism, the No-Miracles Argument (NMA). Briefly put, the NMA suggests that the approximate truth of our scientific theories is the best explanation of their empirical success; if this were not the case, their success would be a miracle. It has been argued (see, e.g., Boyd 1982 and Psillos 1999) that the NMA is an instance of the general inference pattern of IBE. However, since our scientific theories have been generated by means of first-order IBEs, the NMA is only a second order IBE. This has given rise to the charge of circularity against the NMA. Lipton, unlike other recent defenders of IBE (cf. Psillos 1999), admits that the charge of circularity has some bite. For this reason, he focuses attention on the low-level IBEs and their relation to truth. The main themes of his arguments are the unificatory virtues of the realist thesis, which extend over both the observable and the unobservable domain, and the inability

of alternatives, like constructive empiricism, to draw a principled line between these two realms (pp. 198-206).

Overall Lipton's discussion of these issues is clear, precise, and illuminating. Despite the limited space, he succeeds in offering a serious defence of the realist thesis, doing justice both to the subtleties of the issues at hand and the richness of the relevant literature. These arguments represent a substantial contribution to the field.

The most ambitious, original but also controversial part of the book is Lipton's treatment of the relationship between Bayesianism and IBE in chapter 7. Here, he defends the general thesis that "Bayesianism and Inference to the Best Explanation are broadly compatible" (p. 106), in that "Bayesian Conditionalisation . . . is run in part on explanationist tracks" (p. 107). The originality of this view lies in that, except for a few scattered remarks, nearly all systematic treatments of either Bayesianism or IBE have treated the two as mutually exclusive. In favour of Bayesianism, van Fraassen (1989) claimed that Bayesianism shows why IBE is incoherent. Similarly, Howson (2000) argued that IBE amounts to committing the 'base-rate fallacy,' and Salmon (2001) argued that only Bayesianism does not confuse the separate issues of explanation and confirmation. In favour of IBE, Okasha (2000) has highlighted the fact that IBE, unlike Bayesianism, throws light on the context of discovery, i.e. the actual workings of science, while independently of IBE there have been numerous complaints regarding the descriptive inadequacy of Bayesianism (cf. Chihara 1987, Earman 1992).

Lipton attempts to take a middle way in the debate. Although he also presses in favour of IBE, citing the descriptive inadequacies of Bayesianism, he nevertheless attempts to build a bridge between the two by exploiting the alleged overlap between the loveliest (IBE) and the likeliest explanation (suggested by Bayesianism). He constructs this bridge by citing four complementary aspects of the explanations offered by IBE and those offered by Bayesianism: 1) explanatory considerations may be viewed as supplying the actual values of *both* the prior probability of the hypothesis *and* the probabilities of the evidence to be inserted in Bayes' theorem, 2) very often IBE and Bayesian conditionalization yield the same results, in which case one can easily view IBE as a useful heuristic replacing abstract and cumbersome probabilistic reasoning, 3) explanatory considerations help determine the evidence one should conditionalize on and 4) IBE sheds light on the processes involved in the context of discovery. Hence, Lipton thinks that the Bayesian and the explanationist can be friends.

The nobility of Lipton's peace-making ambition notwithstanding, we are not convinced that his solution is viable. Our claim is that on his argument, IBE is relegated to a mere auxiliary device employed in service to the epistemically prior Bayesian theory. This point is clearly demonstrated when we consider what happens when Bayesianism and IBE give different answers to the same question, a possibility which follows naturally from the distinction between loveliness and likeliness. In that case one is confronted with the following options: (a) IBE is right, (b) or Bayesianism is right or (c) neither of the two is right. According to Lipton's claims for the autonomy of IBE, one would expect him to argue in favour of either (a) or (c). On the contrary, however, Lipton repeatedly asserts that Bayesianism gives the *normatively correct* answer and IBE merely reveals why ordinary judgements are mistaken, if they are mistaken (cf. pp. 107, 115, 117, 120). If indeed we do possess a normatively

sound theoretical framework, i.e. Bayesianism, which allows us in principle to reach the correct answer, then how does IBE contribute to the analysis of our inductive practices?

If Bayesianism does in fact provide the normatively correct answer in the face of conflict, then IBE is at best an auxiliary device that can only explain why we are often mistaken in our judgements. As Lipton's own Tversky-and-Kahneman-style examples reveal (e.g. Linda the bank teller[i] and the base-rate fallacy), IBE is more of an example to avoid when addressing the normative question, rather than a model to pursue. In other words, IBE is *not* a rule of inference complementary to the analytically prior Bayesian inference the way Lipton believes it to be.

To be sure, Lipton is primarily concerned with the descriptive task, and from this perspective one might argue that IBE fares better than Bayesianism, since it appears to be a fact that people do not engage in abstract probabilistic reasoning in their everyday affairs, scientific or not. This brings us naturally to our second, more general, complaint. We think that Lipton puts too much weight on the descriptive project at the expense of the philosophically more crucial problem of justification. We take it that the primary task of a plausible philosophy of science is that it present us not simply with an adequate description of our inferential habits but rather with firm grounds for our beliefs these habits generate. Being too modest on this count, Lipton offers a rather weak justificatory case, which shows most clearly in his admission that Bayesianism sets the standards of correctness and IBE merely fills in the descriptive gaps.

This critique is illustrated by considering Lipton's first attempt to reconcile Bayesianism and IBE, i.e. that explanatory considerations help us determine the probabilities to be inserted in Bayes' formula. As long as IBE does not offer the means to *fix* the probability assignments under consideration, no claims to objectivity can be adduced in its favour, except perhaps of the less important merit of descriptive accuracy as far as the actual determination of the priors and likelihoods is concerned (cf. Psillos 2004). Given the immense difficulties in justifying such descriptive constraints, only a strong normative case for IBE and against Bayesianism can aspire to avoid the result of IBE being just 'the second best'. Such an *incompatibilist* approach, however, would of course openly violate Lipton's plea for the compatibility between the Bayesian and the explanationist rules of inference.

Despite our reservations regarding Lipton's reconciliation between the Bayesian and the explanationist, *Inference to the Best Explanation* ultimately offers the most complete presentation and defence of IBE to date and as such it is a 'must' read for anyone who wants a deeper understanding of inductive inference, broadly understood.[ii]

REFERENCES

Bovens, L. and S. Hartmann (2003) *Bayesian Epistemology*, Oxford: Oxford University Press.

Boyd, R. (1984) "The Current Status Scientific Realism" in J. Leplin (ed.) *Scientific Realism*, Berkeley: University of California Press.

Chihara, C. S. (1987) "Some Problems for Bayesian Confirmation Theory", *British Journal for the Philosophy of Science* 38: 551-560.

Earman, J. (1992), *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory*, Cambridge, MA: MIT Press.

Harman, G. (1965) "The Inference to the Best Explanation", *The Philosophical Review*

74: 88-95.

Howson, C. (2000) *Hume's Problem*, Oxford: Oxford University Press.

Okasha, S. (2000) "Van Fraassen's Critique of Inference to the Best Explanation", *Studies*

in History and Philosophy of Science 31: 691-710.

Psillos, S. (1999) Scientific Realism: How Science Tracks Truth, London: Routledge.

Psillos, S. (2002a) "Simply the Best: A Case for Abduction" in A.C. Kakas and F. Sadri (eds.) *Computational Logic*, Berlin Heidelberg: Springer-Verlag.

Psillos, S. (2004) "Inference to the Best Explanation and Bayesianism" in F. Stadler (ed.) *Institute of Vienna Circle Yearbook* Vol. 10, Dordrecht: Kluwer.

Salmon, W. (2001) "Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation" in G. Hon and S. S. Rakover (eds.) *Explanation: Theoretical Approaches and Applications*, Dordrecht: Kluwer Academic Publishers.

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

----

[i] For a Bayesian treatment of the Linda case, see Bovens and Hartmann (2003), pp 85f.

[ii] Thanks to Peter Lipton for valuable feedback on a first draft of this review.