Georgia State University
ScholarWorks @ Georgia State University

Philosophy Theses | Department of Philosophy

4-19-2010

Can Bayesianism and Inference to the Best Explanation be Friends?

Rush Tyler Stewart
Georgia State University

Follow this and additional works at: http://scholarworks.gsu.edu/philosophy_theses

Recommended Citation
Stewart, Rush Tyler, "Can Bayesianism and Inference to the Best Explanation be Friends?" (2010). *Philosophy Theses.* Paper 59.

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

CAN BAYESIANISM AND INFERENCE TO THE BEST EXPLANATION BE FRIENDS?

by

RUSH T. STEWART

Under the Direction of Andrea Scarantino

ABSTRACT

Bas van Fraassen argues that inference to the best explanation (IBE) is a probabilistically incoherent rule (1989). Anyone following IBE is open to being Dutch booked. According to one of the most interesting and popular responses to van Fraassen's argument, van Fraassen misrepresents IBE in probabilistic terms. With the proper probabilistic representation, it is claimed, IBE is not inconsistent with Bayesian rationality constraints. Building on the work of IBE's proponents, I first propose a minimal account of what makes one explanation better than another. I then argue that, even on this minimal account, the alternative probabilistic model of IBE does not work, and hence fails to successfully respond to van Fraassen's argument.

INDEX WORDS: Bayesianism, Inference to the best explanation, Probability, Confirmation, Explanation, Explanatory virtues, Coherence, Dutch book, Information, Bas van Fraassen

CAN BAYESIANISM AND INFERENCE TO THE BEST EXPLANATION BE FRIENDS?

by

RUSH T. STEWART

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2010

Copyright by
Rush Tyler Stewart
2010

CAN BAYESIANISM AND INFERENCE TO THE BEST EXPLANATION BE FRIENDS?

by

RUSH T. STEWART

Committee Chair: Andrea Scarantino

Committee: George Graham
Stephen Jacobson
Sebastian Rand

Electronic Version Approved:

Office of Graduate Studies
College of Arts and Sciences
Georgia State University
May 2010

For Ava Jean Baker

ACKNOWLEDGEMENTS

In writing this thesis, I was aided, in various ways, by many people. I would like to thank them. First, my advisor, Andrea Scarantino, read and offered extensive and extremely helpful comments on several drafts. He was very generous with his time. I'm sure his approach to philosophy will continue to influence me in the future. Second, I have learned a great deal from all of my professors at GSU, especially all of my committee members: George Graham, Stephen Jacobson, and Sebastian Rand. Stephen Jacobson, though, deserves special thanks. Third, my fellow students in the Master's program were a tremendous boon in researching and writing my thesis. Thanks, in particular, to Ryan DeChant, Matt Duncan, Trip Glazer, Dylan Murray, and Brad Summers. Finally, I would like to thank my family for constant support.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

CHAPTER

1. INTRODUCTION 1

2. THE PROBLEM: IBE IS PROBABILISTICALLY INCOHERENT 1

3. IBE STRIKES BACK 4

4. A NEW PROBLEM: FA IS UNSOUND 10

5. OBJECTIONS 19

6. CONCLUSION 24

BIBLIOGRAPHY 25

1. INTRODUCTION

Bayesianism and inference to the best explanation (IBE) are widely influential accounts of induction. It would be a serious problem if they were inconsistent with each other. That is exactly what Bas van Fraassen argues in *Laws and Symmetry*. There, he purports to show that IBE, in offering a policy for updating beliefs in conflict with Bayesianism, violates rationality constraints. As he puts it, IBE is probabilistically incoherent. Van Fraassen concludes that allowing IBE-based explanatory considerations to factor into belief revision is irrational.

Defenders of IBE reject van Fraassen's argument, claiming that Bayesianism and IBE can "be friends" (Lipton 2001). In particular, they object to how van Fraassen understands IBE, and offer alternative formulations that aim to vindicate IBE's rationality. Here, I argue that the most promising counterproposal fails. If I am right, the rational status of IBE as a model of induction remains in question.

2. THE PROBLEM: IBE IS PROBABILISTICALLY INCOHERENT

Van Fraassen writes: "we should not listen to anyone who preaches a probabilistic version of Inference to the Best Explanation, whatever the details. Any such rule, once adopted as a rule, makes us incoherent" (169). His argument is that a probabilistic version of IBE is at odds with Bayesian conditionalizing, and is consequently irrational. Let us see how this works.

The basic schema for IBE is, for a given set of data, to infer the probable truth of the hypothesis that explains the data better than competing hypotheses do (Okasha, p. 691). The core idea behind IBE, then, is that the quality of an explanation *qua* explanation is a guide to the probability (of truth) of that explanation (Harman 1965; Lipton 2001; Psillos 2003; Lipton 2004). According to Bayesianism, beliefs come in degrees that can be represented with numerical

values. An agent is rational only if these values satisfy the following properties: 1. They conform (synchronically)[1] to the axioms of the probability calculus, and 2. They are updated, on new evidence, according to Bayes' rule. [2]

In making his case against IBE, van Fraassen imagines a situation in which an "alien" die is rolled and an agent, Peter, is considering different hypotheses about the bias of the die coming up *ace*. Let P(.) be a function from sentences to [0, 1], that represents an agent's belief system, where 1 represent certainty and 0 represents utter disbelief (Douven, 1999).[3] For instance, if Peter believes H_2 to be more probable than H_1, then $P(H_2)>P(H_1)$. As the die is rolled and evidence comes in, Peter updates P(.) to reflect his change of credence in the various hypotheses of bias. For hypothesis H and evidence e, Bayesianism prescribes the following updating rule:

(**Bayes' Rule**) $P'(H)=P(H|e)$

The rule states that the new probability of a hypothesis, P'(H), called the *posterior*, should be equal to the old probability of that hypothesis conditional on the evidence, P(H|e). Once the evidence e comes in, P(H) should change to P'(H), or P(H|e). The formula for calculating P(H|e) is given by Bayes' Theorem, a basic theorem of the probability calculus:

(**Bayes' Theorem**) $P(H|e)=P(H) \times [P(e|H)/P(e)]$

P(H) is referred to as the *prior*, P(e|H) as the *likelihood* (of the evidence given the hypothesis), and P(e) as the *expectedness*. In the alien die example, Peter assigns prior probabilities to the various hypotheses of bias. He then updates those probabilities as evidence comes in according to Bayes' rule.

Peter, as van Fraassen's story goes, is also a proponent of IBE. As such, he adds bonus

[1] Synchronic coherence is coherence at a given time. Diachronic coherence is coherence across time.
[2] Bayesians maintain that these probabilistic constraints are in addition to constraints of logical consistency. For example, one should not believe both that Tesla hated Edison and that Tesla did not hate Edison.
[3] I am here presupposing a subjectivist interpretation of Bayesianism. As Elliott Sober puts it, "[m]ost contemporary Bayesians have given up on objective Bayesianism, and have gone the subjective route" (2002, p. 2). In section 5, I will consider whether anything changes with respect to my main argument under an objectivist interpretation.

points to the posterior probability P'(H) of a hypothesis, *after* conditionalizing on the evidence, according to how well it explains the evidence *e*. That is, Peter adopts a rule "which fixes a posterior probability for each hypothesis of bias, depending not only on the initial probabilities and the series of outcomes, but also on a feature which he calls explanatory success" (p. 166). While van Fraassen does not give the details of the rule, he says "it really gives *bonus* probabilities to certain hypotheses." For example, suppose that after the die comes up *ace* four times in a row, simple Bayesian updating leads Peter's credence in a fifth toss coming up *ace* to be 0.87 (P(fifth toss coming up ace|first four tosses came up ace)=0.87), given his priors. But not only do hypotheses of high bias predict four *aces* in a row with high probability, they also, retrospectively, offer good explanations of why only *aces* came up. Van Fraassen says that the hypothesis of perfect bias towards *ace* is the most explanatory "so it definitely gets a *bonus*." As a result, Peter, as advocate of IBE, increases his credence in more explanatory hypotheses "more than [he] would anyway" (p. 166). He redistributes his probabilities, raising his probability in a fifth *ace* from 0.87 to, say, 0.9. In summary, Peter goes through a two-stage process of updating his beliefs for any evidence: first, he conditionalizes using Bayes' Rule, then, he bumps up the posterior of the better explanations according to IBE.

Van Fraassen shows that anyone adopting such a two-stage process of belief revision is open to being Dutch-Booked. A Dutch Book is a series of bets that an agent voluntarily agrees to that ensures a net loss for that agent. Dutch Book arguments presuppose that degrees of belief match betting prices. For a unit wager ($1) on event X, the agent assigns probability p to X iff p$1 is what she values a bet that pays $1 if X and nothing otherwise. The agent is willing to sell the bet for p$1 or greater, and buy the bet for p$1 or under. If her degrees of belief do not conform to the probability calculus, a Dutch Book can be made against her, whereas it cannot be

made if they conform to it (Hajek, 2008).

Van Fraassen's proof that Peter's IBE updating rule leads to a Dutch Book is an instance of the well-known dynamic, or Lewis-Teller, Dutch Book proof that anyone who adopts an explicit plan to update with a rule other than straightforward Bayesian conditionalization can be Dutch-Booked. Liability to sure loss is taken to be irrational by Dutch Book arguments, so any IBE-based updating rule is irrational. Van Fraassen does not deny either that there are explanations or that some explanations are better than others. What he denies is that *inference* to the best explanation is a good updating rule. In other words, Bayesianism and IBE conflict when it comes to belief revision. He concludes that because Bayesianism is correct, IBE is not.

3. IBE STRIKES BACK

Two reactions to van Fraassen's argument are possible. First, one might buy the argument and accept that Bayesianism and IBE are inconsistent. In that case, one is forced to reject either Bayesianism, or IBE, or both. Second, one could reject van Fraassen's argument and deny that Bayesianism and IBE are inconsistent. Lipton (2001) calls this the *irenic* approach. Taking the irenic approach requires repudiating either van Fraassen's way of understanding Bayesianism, or his way of understanding IBE, or both. Versions of the irenic approach are dominant in the literature (Douven 1999; Okasha 2000; Lipton 2001; Niiniluoto 2003; Lipton 2004; Ganson 2007). One of the most popular responses to van Fraassen's argument is to deny that IBE violates Bayesian rationality constraints by rejecting his way of understanding and modeling IBE within the Bayesian framework. Proponents of IBE, according to this line of response, need not advocate tinkering with the posterior by adding bonus points after Bayesian conditionalizing. Instead, explanatory considerations enter into Bayes' theorem directly by means of the prior and likelihood. Beliefs are updated according to straightforward Bayesian

conditionalization, and a Dutch Book can consequently be avoided. On this view, van Fraassen misrepresented IBE.

To evaluate this charge, we need to clarify the intended content of IBE. We need an account of what it means to infer to the best explanation. Explanations are, fairly uncontroversially, answers to *why* or *how-possibly* questions and a means by which we acquire understanding.[4] Going much further than this, however, is far from easy. Proponents of IBE generally leave open what specifically counts as an explanation and what makes one explanation better than another. That is, they concede that IBE awaits a fully adequate articulation. For good reason: there is not a received view of explanation.[5] All models currently on the table yield some disappointing results. The models have varying defects, such as capturing only some explanations while excluding others, or including some things that should not count as explanations at all. In this section, I will sketch a minimal account of explanatory superiority (an account that strives to be as uncontroversial as possible) based on the work of some of IBE's proponents and then lay out the proposed irenic rejoinder to van Fraassen.

Instead of attempting to insert a theoretical account of explanation into IBE, Peter Lipton (2004) makes two crucial distinctions that do not depend on the exact nature of explanation. First is the distinction between *actual* and *potential* explanations. The conditions that are required for a hypothesis to be a *potential* explanation are the same as those for *actual* explanations with one exception. There is no requirement that *potential* explanations be true.[6] So all actual explanations

---

[4] Some theoretical accounts of scientific explanation go on to place substantive constraints on the content of answers to *why* and *how-possibly* questions.

[5] See Salmon (1990) for an excellent overview of the major models of scientific explanation.

[6] One way to get a handle on this distinction is to couch it in terms of the familiar Deductive-Nomological (D-N) model of explanation. According to the D-N model, (1) an explanation must be a valid deductive argument, (2) the explanans must contain essentially at least one general law, and (3) the explanans must have empirical content (Salmon 1990, p. 12). These would be the conditions for *potential* explanations. There is an additional condition for *actual* explanations: (4) the sentences constituting the explanans must be true. In terms of the D-N model, *potential*

are potential explanations, but not *vice versa*. For explanandum e, the set of potential explanations is E={x: x potentially explains e}.

Lipton argues that since the goal of inference is truth, IBE cannot be inference to the best *actual* explanation. If it were, we would have to know that explanations are true before inferring them. So IBE has to be inference to the best *potential* explanation, the xεE such that for any yεE, where y≠x, x is better than y. However, this is uninformative about what makes one explanation *better* than another.

Lipton introduces a second important distinction between two ways of interpreting *best*: as *likeliest* or as *loveliest*. On the one hand, "inference to the best explanation" might mean inference to the *likeliest explanation*, the xεE such that P(x)>P(y), for any other hypothesis yεE, where y≠x. On the other hand, IBE might mean inference to the *loveliest explanation*, the explanation that, if true, would provide the most understanding. If we interpret *best explanation* as *likeliest explanation*, the reconciliation of IBE and Bayesianism is at hand: we infer the probable truth of the most probable explanation, where the probability of the explanation is determined through straightforward Bayesian updating. The problem with interpreting *best* explanation as *likeliest* explanation is that it makes IBE trivial. As Lipton puts it, "[s]cientists do infer what they judge to be the likeliest hypothesis, but the main point of a model of inference is precisely to say how these judgements are reached, to give what scientists take to be the *symptoms of likeliness*. If Inference to the Best Explanation is along the right lines, explanations that are lovely will also be likely, but it should be in terms of loveliness that the inference is made" (2001, p. 13; emphasis added). It is unclear, interpreting *best* as *likeliest*, what work

---

explanations are (1) valid arguments satisfying (2) and (3). *Actual* explanations satisfy (4) as well, so they are sound arguments.

explanation is doing for IBE as a model of inference. What's more is that IBE would seem to be an utterly superfluous rule – we already have Bayesianism to calculate probabilities.

This leaves us with the second way of interpreting *best* explanation, namely, as *loveliest* explanation, or as the explanation that provides the most understanding. Here, we confront anew the lack of consensus on what makes one explanation better than another in terms of providing understanding. Perhaps the least controversial way to proceed is to follow the lead of those writers on explanation (including Lipton, one of IBE's chief proponents) who focus on what virtues an explanation has to exhibit. Then, we can then think of explanation x as being better than explanation y if x displays more explanatory virtues, or displays them to a higher degree, than y. Even though there are some differences in proposed lists of explanatory virtues, there is also a lot of agreement. For the purposes of this paper, I will rely on Lipton's (2001) list, which includes the popular virtues of *scope, precision, mechanism, unification,* and *simplicity*. Different models of explanation accord these virtues different weight in trying to establish what makes one explanation better than another. For example, the Unification model of explanation (Kitcher, 1989), according to which explanation is a matter of reducing the number of independent assumptions and making hypotheses more general, takes *unification* to be *the* defining feature of an explanation. The Causal-Mechanical model (Salmon, 1990), which sees providing explanations as a matter of identifying underlying causal mechanisms, takes *mechanism* to be of central importance. What the contending models of explanation share is the idea that the quality of an explanation is determined by the extent to which it satisfies some subset of those virtues.

Explanation is a means for understanding. Better, that is, *lovelier*, explanations produce more (potential) understanding. Because the quality of an explanation is determined by the explanatory virtues, explanatory virtues are properties of explanations that tend to produce

understanding. That seems fair. Consider precision. If one explanation is more precise than others, it *ceteris paribus* offers more potential understanding. Similarly, if an explanation lays bare a causal mechanism (mechanism), or applies to a diverse range of phenomena (scope), it offers more potential understanding than explanations that do not. These virtues are intuitive matches for the sorts of things that produce greater understanding.

Here, then, is my proposal for a minimal account of explanatory superiority. I follow Lipton in understanding the *best* explanation as the *loveliest*, the explanation that offers the most understanding. But loveliness is not a brute fact about explanations. It obtains because of the explanatory virtues of *scope, precision, mechanism, unification,* and *simplicity.* For clarity, I will make one further assumption. One could imagine that the quality of an explanation is primarily a matter of satisfying a proper subset of these virtues. In that case, it would be possible for an explanation x to be better than y, because, even though y satisfies a greater number of the virtues, x better satisfies the particular virtues that are assigned a greater weight. For the purposes of this paper, I will assume that all of the explanatory virtues are weighted equally, though I do not think my arguments in the next section depend on this assumption. The crucial contention of IBE is that explanatory superiority indicates probabilistic superiority. In other words, *loveliness* is a guide to *likeliness*.

In making his case for the friendship of Bayesianism and IBE, Samir Okasha (2000) considers the following example.

> (**DOC**) A mother takes her five-year-old child to the doctor. The child is obviously in some distress. On the basis of the mother's information, the doctor forms two competing hypotheses: that the child has pulled a muscle, and that he has torn a ligament; call these H_1 and H_2 respectively. A keen advocate of IBE, the doctor examines the child carefully, and decides that H_2 offers the better explanation of the observed symptoms. (p. 702)

The doctor justifies her reasoning, Okasha goes on to say, by first pointing out that pulled muscles are very rare in pre-adolescent children, whereas torn ligaments are fairly common. She then adds that the symptoms are more likely if the child has torn a ligament than if the child has pulled a muscle. The first thing to notice is that these are probabilistic considerations. The doctor is just saying that both the prior and likelihood of H_2 are greater than the prior and the likelihood of H_1. Okasha does not specify what explanatory considerations lead the doctor to infer H_2 or to assign the probabilities that she does. In fact, his claim is difficult to evaluate because he offers no positive account of what an explanation is or what makes one explanation better than another. However, if the above account of loveliness is on the right track, we know what sort of considerations would have to be in play. The best explanation has to be the explanation that best satisfies the explanatory virtues of scope, precision, mechanism, unification, and simplicity. Unfortunately, it is not obvious that these virtues will distinguish between H_1 and H_2 as suggested by the doctor. Below I will consider variations of the DOC scenario to make this more obvious. But let us complete the IBE rejoinder first.

In responding to van Fraassen's Dutch book argument against IBE, Okasha straightforwardly puts the proponent's proposal for modeling IBE in Bayesian terms. Call this argument advocating friendship between IBE and Bayesianism the *Friendship Argument* (FA).

> (**FA**) The correct way of representing IBE […] views the goodness of explanation of a hypothesis *vis-à-vis* a piece of data as reflected in the prior probability of the hypothesis P(H), and the probability of the data given the hypothesis P(e|H). The better the explanation, the higher is one or both of these probabilities. Relative to this account, favouring a hypothesis on the grounds that it provides a better explanation of one's data than other hypotheses, and indeed making it a rule to do so, is perfectly consistent with Bayesian principles. For fixed P(e), Bayes's theorem tells us that P(T|e) is an increasing function of P(T) and P(e|T) and nothing else; so ending up with the highest degree of belief in the theory which explains the data best is exactly what the good Bayesian conditionalizer should do. No better reconciliation between Bayesianism and IBE could be hoped for. (Okasha, 703-704)

Lipton (2004, p. 108) explicitly endorses Okasha's approach.[7] According to them, van Fraassen overstates his conclusion in saying that *any* probabilistic version of IBE makes us incoherent. FA proposes a way for IBE to operate that does not leave one open to a Dutch Book. The doctor, in following IBE, just conditionalizes in the standard Bayesian way, adding no bonus points after conditionalizing, though the prior and the likelihood are fixed in accord with explanatory considerations. Because explanatory considerations determine the prior, the likelihood, or both to be higher for better explanations, and, for fixed P(e), the posterior is a function of only those quantities ($P(H|e)=P(H) \times [P(e|H)/P(e)]$), better explanations will be more probable, and IBE and Bayesianism can be reconciled. Okasha concludes, "It appears that the conflict between IBE and Bayesianism alleged by van Fraassen depends entirely on an idiosyncratic way of representing IBE in probabilistic terms" (p. 703).

4. A NEW PROBLEM: FA IS UNSOUND

Lipton is right that IBE should be understood as *inference to the loveliest potential explanation* if we are to avoid triviality. Yet, by avoiding triviality, this formulation of IBE makes problematic assumptions about the relationship between explanation and confirmation. Okasha says that it is hard to see what it could mean to say that H_2 is a better explanation than H_1 if neither $P(H_2)>P(H_1)$ nor $P(e|H_2)>P(e|H_1)$. I disagree. Where loveliness is a matter of potential understanding, likeliness is a matter of the probability of truth. The two notions are distinct, and, I'll argue, they sometimes pick out different explanations. Non-actual explanations can offer a lot of potential understanding; they can be very lovely. Moreover, there are likely but unlovely explanations.

[7] However, Lipton himself seems to argue just that the *loveliest* explanation is *reliably* the *likeliest*. He relaxes the requirement that IBE and Bayesianism always favor the same explanations. According to him, IBE is the correct (descriptive) account of how we actually reason, and justified insofar as it approximates Bayesian reasoning.

My diagnosis of this difficulty for FA is that it is wrong about the probabilistic nature of the explanatory virtues. In an earlier essay, van Fraassen observes that while he and Salmon propose very different accounts of explanation, "on one central point [they] are in complete agreement: [they] agree that to give an explanation is to give relevant information" (1983, p. 214). According to them, the explanatory virtues, the properties responsible for making certain explanations better than others, are actually informational virtues. As I pointed out earlier, IBE's central claim is that the *loveliest* explanation is also the *likeliest* explanation. I will argue, on the contrary, that there is often a tradeoff between a hypothesis' providing relevant information, i.e. being lovely, and its having high probability, i.e. being likely. The bottom line is that IBE avoids van Fraassen's Dutch Book at the expense of making false assumptions.

Consider a formalized version of the Friendship Argument (FA), where the minimal account of loveliness described above is presupposed:

(1) Explanatory virtues ("the goodness of an explanation") are reflected in the prior, P(H), and the likelihood, P(e|H).
(2) The better the explanation H, the higher the prior P(H), the likelihood P(e|H), or both.
(3) For fixed P(e), the posterior P(H|e)=P(H) x [P(e|H)/P(e)] (i.e. the likeliness) is an increasing function of the likelihood P(e|H) and the prior P(H) and nothing else.

Therefore,

(4) The loveliest explanation is also the likeliest, or loveliness is a guide to likeliness.

FA is, strictly speaking, invalid.[8] Instead of (2), FA needs (2'):

(2') For any two potential explanations H and H', if H is a better explanation than H', then the *product* of the prior and the likelihood for H is greater than the product of the prior and the likelihood for H' ([P(H) x P(e|H)]>[P(H') x P(e|H')]).

[8] The problem is with premise (2). For explanatory superiority to be a guide to probabilistic superiority, it is not enough that one term in a two-term multiplication (Bayes' theorem) be higher due to explanatory considerations. Return to DOC. Assume H_2 is a better explanation than H_1. By (1), this explanatory superiority is reflected in the priors and likelihoods. By (2), either $P(H_2)>P(H_1)$, or $P(e|H_2)>P(e|H_1)$, or both. (2), along with (1) and (3), can be true and yet (4) false. (2) is true if any disjunct is true. That is consistent with $P'(H_2)<P'(H_1)$. For example, (2) is true in any case in which $P(H_2)>P(H_1)$, but $P(e|H_2)<P(e|H_1)$ such that $P'(H_2)<P'(H_1)$. (1) and (3) are true in that case too, but (4) is not.

(2') allows the proponent of IBE greater room than requiring that both the prior and likelihood of the better explanation be greater than those of competing explanations – a surefire way to guarantee a greater posterior.[9] (2') seems to be what Okasha has in mind. If so, FA is valid. Premise 3 is uncontroversial. The premise that concerns me here is (2'). I will argue that (2') is false, and that FA is consequently unsound. In what follows, I will discuss two variations on DOC in order to demonstrate this.

DOC SCENARIO 1. Consider the following variation on Okasha's original DOC scenario. Suppose that the doctor in DOC forms, in addition to H_2, a third, less informative hypothesis, H_3, that the child's arm hurts because it has been damaged. This explanation is very likely, though not certain: the child could be suffering sensations of pain without nociceptor excitation or damage to the arm, for example. H_3, however, is *not* lovely. It provides very little understanding, certainly less than H_2. H_2 provides the *mechanism* underlying the symptoms and specifies the sort of damage (*precision*). H_2 is lovelier than H_3, that is, it satisfies more of the explanatory virtues, though it is less likely than H_3. Okasha, in a footnote, hints that he might not count a hypothesis such as H_3 as an explanation.[10] He does not, however, say why. H_3 provides *some* potential understanding of the symptoms. If Okasha wants to deny that H_3 is a potential explanation, he needs a theoretical account that distinguishes explanations, both good and bad, from non-explanations. Even on candidate accounts of explanation, such as the Causal-Mechanical model, which make more substantive (and controversial) assumptions than the

[9] To see why (2′) is preferable, suppose that the child's symptoms are more likely if the child has pulled a muscle than if the child has torn a ligament. If the doctor diagnoses the child with a pulled muscle based just on this fact, though, she is in error. For it is also the case that pulled muscles are extremely rare in young children. So rare, in fact, that the torn ligament hypothesis is more probable despite the lower likelihood. To ignore the prior leads one astray of correct probabilistic reasoning. Ignoring the prior is committing the so-called "base-rate" fallacy (Psillos, p. 4). The point is that H_2 can be more probable and, if FA is to be valid, a better explanation even if it is not the case that both the prior *and* likelihood are greater than those of competing hypotheses. (2′) allows for cases in which $P(H_2)$, for example, is greater than $P(H_1)$ to the extent that $P'(H_2)>P'(H_1)$ despite $P(e|H_2)<P(e|H_1)$.

[10] The actual example of a likely but unlovely explanation Okasha mentions is "Opium puts people to sleep because of its dormative powers" (2000, p. 705).

minimal account we are working with, it seems that H_3 *does* count as an explanation – the damaged arm causes, *via* excitation of nociceptors, the sensation of pain – it is just less informative than other explanations. Thus, despite H_2's satisfying more of the explanatory virtues, the product of the prior and the likelihood of H_2 is not greater than the product of the prior and the likelihood of H_3. So the *better* explanation is not more *likely*.

One might complain that H_2 and H_3 are not competitors, and so inferring both does not violate the spirit of IBE. In fact, H_2 entails H_3, and by inferring H_2, one must rationally infer H_3. That's right so far as it goes. Nonetheless, this example evinces something about the behavior of explanatory virtues with respect to the Bayesian framework, namely, that satisfying more explanatory virtues does not guarantee a higher probability for a hypothesis. Furthermore, this probabilistic behavior of the explanatory virtues is not restricted to cases of non-competitive hypotheses. We could switch the example to comparing H_4, the hypothesis that the child damaged a *muscle* (leaving imprecise the sort of damage), with H_2, the child tore a ligament. H_2 is a better explanation than H_4, with which it competes, for the same reasons why it is better than H_3, but it probably fails (and will surely fail in some such cases) to be probabilistically favored over H_4. This is because there are many more ways for H_4 to be true, since H_4 is a disjunction of *all* the ailments that involve damaged muscle. Consider now a second example.

DOC SCENARIO 2. Suppose that there are six children in the doctor's waiting room complaining of pain in their limbs. Suppose further that this is an unusually high number given that the doctor usually sees only a couple of these kinds of cases *per week*. She is about to examine the first child. She considers H_2. Eager to find a lovely explanation of all the symptoms of all of her patients that day, the doctor cleverly devises another hypothesis. Let H_5 be the hypothesis that the child contracted dengue fever from a dengue fever outbreak. Symptoms of

dengue fever include muscle, joint, and bone pain. H_5 not only offers an explanation of the current patient's symptoms, but also of the symptoms of the other children in the waiting room, and of the symptoms in any similar cases in the immediate future. The hypothesis of dengue fever posits just *one* causal mechanism for a wide range of cases. Compared to H_1 or H_2, hypothesis H_5 satisfies more of the virtues of scope, precision, mechanism, unification, and simplicity. It has a broader scope because it has implications for the cases in the waiting room; it is just as precise; it provides a causal mechanism; less independent assumptions are required to explain all of the cases; the hypothesis is less complex than the conjunction of independent reasons and diagnoses for all five cases. In short, H_5 provides more potential understanding than H_2. The problem with H_5 is that dengue fever is extremely rare in the U.S. For the doctor, who is well aware of the prevalence of dengue in the U.S., the prior $P(H_5)$ is very low. Despite its loveliness, H_5 simply is not likely.

I remarked earlier that, in the original DOC scenario, the explanatory virtues of scope, precision, mechanism, unification, and simplicity do not distinguish between H_1 and H_2. With respect to explanatory virtues, a pulled muscle and a torn ligament are alike as explanations in DOC (at least as far as the information Okasha provides goes). If one explanation is more likely than the other, but neither is *lovelier*, then it is not the case that explanatory considerations *all by themselves* fix the probabilities in Bayes' Theorem. Some likelihoods are calculable without taking into account the explanatory virtues (Salmon 2001b). In such cases, we need make no appeal to explanatory considerations in fixing the likelihood. In DOC, the likelihood of the observational data given that the child has a torn ligament, for instance, is calculable from the medical cases in which someone had both a torn ligament and those observed symptoms. It is unclear exactly what explanatory considerations are guiding the probability assignments in DOC.

Consider another case in which explanatory virtues play no role in probability assignments. David Lewis famously proposes the "Principal Principle," a normative constraint on subjective probability in cases concerning "chance," or "relative frequency in the long run" (1986, p. 134). "Let *t* be any time. Let *x* be any real number in the unit interval. Let *X* be the proposition that the chance, at time *t*, of *A*'s holding equals x […] Then, [P(A|X)=x]" (p. 137). So if a six-sided die is fair, the (subjective) probability of it coming up 1 ought to be 1/6. No explanatory virtues are taken into consideration.

Okasha does not believe that explanatory considerations are always required to fix probabilities anyhow. In considering the possibility that P(H) and P(e|H) can both be high and yet H fail to explain e, let alone provide the best explanation of e, Okasha says, "not all cases of updating by Bayesian conditionalization involve explanatory considerations […] not all cases of conditionalization are cases of IBE" (2000, p. 705). Granting that we can fix the prior and the likelihood without explanatory considerations in some cases, it is not plausible that in cases of IBE somehow *only* explanatory considerations are at work. Okasha never gives a reason to think so. So it looks like considerations of explanatory virtue do not exclusively fix the probabilities, and Okasha probably does not believe otherwise. But if the explanatory virtues do not fix the probabilities all by themselves, the door is open for the *loveliest* explanation to be different from the *likeliest* explanation.[11] This is just the case in *DOC SCENARIO 2*.

Let's take stock. In order for FA to be sound, first, it must be the case that the virtues relevant in determining the *loveliest* explanation are, in fact, "reflected in" the prior or the likelihood as premise (1) of FA states. For example, if x is lovelier than y because it satisfies

[11] A strange claim that Okasha makes is that IBE illuminates where Bayesianism is silent. The cases he has in mind are ones in which new hypotheses are invented in response to experience. But if he maintains that explanatory considerations fix probabilities, we are left wondering why those explanatory considerations do not fix the probabilities in the purported cases where Bayesianism is silent.

unification and *precision*, whereas y does not satisfy any virtues that x does not, then *unification* and *precision* must affect the value of the product of the prior and the likelihood of x or y. Second, it must be the case that the impact that the explanatory virtues make on the prior and the likelihood is positive, as premise (2') of FA states. For example, the product of the prior and the likelihood of hypothesis x, by satisfying *unification and precision*, is greater than that of y, where the two hypotheses are alike with respect to all explanatory virtues except *unification and precision*. The problem is that premise (1) and premise (2') of FA are in conflict. Let me explain.

In an earlier paper on explanation, van Fraassen notes a widely accepted principle of confirmation:

> If theory T provides information that T' does not provide, and not conversely, then T is no more likely to be true than T' (1983, 166).

A theory's providing more information, according to van Fraassen, is a matter of (potentially) telling us more about the world. Think of the explanatory virtue of precision in the context of DOC. The more precise an explanation is, the better that explanation is, the more potential understanding it offers. The hypothesis that the child tore a ligament is less precise than the hypothesis that the child tore the radial collateral ligament. The latter hypothesis, which is the better explanation, entails the former. However, the probability of a hypothesis can only be less than or equal to the probability of any hypothesis it entails.

The same considerations apply, *mutatis mutandis*, for many other explanatory virtues. Making a hypothesis apply to a larger class of phenomena (*scope*) means, in general, making it less probable. Unification is a matter of decreasing independent assumptions and increasing generality. More unified explanations thereby entail more than less unified explanations: more unified explanations entail at least whatever those less unified explanations entail, and, by being more general, they entail things about more cases than those less unified explanations do. These

explanatory virtues increase the informational content of a hypothesis. In so doing, they do not increase the probability of a hypothesis. Depending on how it is construed, simplicity, on the other hand, may not be informational. For example, simplicity might be a matter of making *fewer* claims about the world.

Thus, FA faces a dilemma. Either the explanatory virtues that are informational will be reflected in the prior and the likelihood, or they will not be. Let's take the first horn first. The explanatory virtues are reflected in the prior and the likelihood. (2') is false if the relevant explanatory considerations do not fix the priors and the likelihoods such that the product of those quantities is *greater* for more explanatory hypotheses. But if the relevant explanatory considerations are informational, they do not serve to raise the prior or the likelihood for better explanations (this is the case in *DOC SCENARIO 1*). So (2') is false.

On the second horn, the informational explanatory virtues such as *precision*, *scope*, *mechanism*, and *unification* are not reflected in the prior and the likelihood (*DOC SCENARIO 2* looks like a case in which explanatory virtues are not responsible for the priors). This means that premise (1) of FA, the explanatory virtues are reflected in the prior and the likelihood, is false. This seems to be the case in Okasha's original DOC example. Taking the second horn also poses a problem for (2'). If any combination of *precision*, *scope*, *mechanism*, and *unification* is relevant in determining the loveliest explanation, they are unable to influence the relevant probabilities such that the *loveliest* explanation comes out the *likeliest*. Since this list is extensive, including most of the explanatory virtues (or many important explanatory virtues, even if Lipton's list is a bit off), (2') is probably false on the second horn. Part of the allure of IBE is the ability to distinguish between empirically equivalent theories on the basis of the explanatory virtues. It turns out, however, that many of those virtues do not increase probability.

According to Bayesianism, P(.) represents an agent's belief system. If explanatory considerations do not raise the probability of a hypothesis, then they do not provide more reason to believe a hypothesis. So the fact that one hypothesis is a better explanation of *e* than some other hypothesis is not thereby a reason to believe the hypothesis. This goes against the very spirit of IBE, FA, and (2').

The foregoing diagnosis predicts that we could construct any number of cases of the following sort. H_1 and H_2 are hypotheses of equal probability and equal explanatory quality. Neither satisfies the virtue of mechanism, for example, but each satisfies the four other explanatory virtues equally. Now, adjust H_2 so that it satisfies mechanism too, while H_1 still does not. Does it follow that $P'(H_2)>P'(H_1)$? No. For we have adjusted H_2 so that it makes an additional claim about the world, so P'(H_2, which satisfies *mechanism*)$\leq$P'(H_2, which does *not* satisfy mechanism). But, by assumption, P'(H_2, which does *not* satisfy mechanism)=P'(H_1, which does *not* satisfy mechanism). So P'(H_2, which satisfies *mechanism*)$\leq$P'(H_1, which does *not* satisfy mechanism). In such a case, the *lovelier* explanation is not more *likely*. The same goes for the virtues of scope, unification, and precision.

Van Fraassen writes, "As long as the pattern of Inference to the Best Explanation […] is left vague, it seems to fit much rational activity. But when we scrutinize its credentials, we find it seriously wanting" (1987, p. 131). I think this is just what we have found. FA is unsound. Even on the least controversial account of explanatory quality I could muster, making precise what IBE amounts to hurts Okasha's case. In light of the considerations in this section, maintaining the central tenet of IBE, that the loveliest explanation is the likeliest, is very difficult. It looks as if a healthy friendship between Bayesianism and IBE is not in the offing. To be fair, I have to moderate my conclusion to the extent that proponents of IBE admit they have neither a fully

worked-out account of explanation nor of IBE, but it is up to the proponents of IBE to fill in the details. I hope this section shows the extensive difficulties IBE runs into as a model of induction.

5. OBJECTIONS

FA explores just one possible response, albeit the most promising in my book, to van Fraassen's original Dutch Book argument against IBE. Consequently, my arguments against the viability of FA do not suffice to establish either the irrationality of IBE or the incompatibility of IBE with Bayesianism. Below I discuss some of the alternative responses.

Model IBE a Different Way in the Bayesian Framework. There are alternative proposals for how explanatory considerations enter into Bayesian belief revision. Instead of fixing the prior and the likelihood, explanatory considerations might enter in Bayesian conditionalizing as evidence. For example, the fact that H_2 is the best explanation of the data – call this fact *b* – is evidence for H_2. Suppose that H_1 and H_2 are equally probable. Once we consider *b*, $P'(H_2|b)$ will be greater than $P'(H_1|b)$. We would not violate conditionalization as we would were we to follow the policy van Fraassen proposes. The only rule we follow in revising our beliefs is Bayes' rule. (For a similar proposal, see Ganson 2007).

This account is also problematic. P(b) will be the same for H_1 and H_2. That means the only way for *b* to affect the posteriors is *via* the likelihoods. So it must be the case that $P(b|H_2)>P(b|H_1)$. This reply thereby assumes the explanation-confirmation link. Is *b* (the claim that H_2 is the best explanation) more likely given H_2 than it is given H_1? We at least need an argument. This link has been called into question by van Fraassen's point about explanatory virtues being informational. To simply assert such a model is to beg the question in context of responding to van Fraassen's critique. And, to reiterate, the devil is in the details for IBE.

Furthermore, this model does not guarantee that the loveliest explanation will be the likeliest – the cornerstone of IBE. In the setup, I assumed that H_1 and H_2 were equally probable. In that case, by conditionalizing on *b*, H_2 is the likeliest explanation. Nothing prevents it from being the case that before conditionalizing on *b*, H_1 enjoys a much higher probability than H_2, so much greater that conditionalizing on *b* does not yield $P'(H_2)>P'(H_1)$. If that's so, then only in some proper subset of cases will the loveliest explanation be the likeliest. In that case, the core of IBE is lost. If the better explanation does not have an equal or higher prior probability than its competitors, then this proposal does not guarantee it has a higher posterior.

Reject Subjective Bayesianism. According to subjective Bayesianism, an agent is rational only if his or her beliefs conform (synchronically) to the axioms of the probability calculus, and they are updated, on new evidence, according to Bayes' rule. Nothing about these requirements makes it necessary that the best explanation comes out the most probable. Only coherence is required to satisfy Bayesian rationality constraints. It is in violation of neither the axioms of the probability calculus nor subjective conditionalization that the best explanation, in terms of explanatory virtues, has a lower conditional probability on the evidence than some rival explanation. This leads Weisberg (2009) to recommend rejecting subjective Bayesianism for an objective Bayesianism that takes explanatory considerations as *a priori* constraints on probability assignments. P(.) would then represent the objectively correct *a priori* probability distribution.

There are at least two things to say in response to this objection. First, the points above apply to either subjective or objective Bayesianism. If we think of "fixing the prior and the likelihood" in subjectivist terms, the claim is that explanatory considerations are the means (or among the means) by which we arrive at our personal probabilities. If we think of it in objectivist terms, explanatory considerations are supposed to be *a priori* constraints on the probabilities.

Either way, the explanatory virtues do not behave probabilistically as Weisberg needs them to. For example, it is not an *a priori* truth that, *ceteris paribus*, a more precise explanation of e is more probable. In many cases, precision makes a hypothesis less probable.

Considering that not all of the explanatory considerations augment the prior or likelihood of the best explanation, indeed, some pull in the opposite direction, we cannot say *a priori* that the loveliest explanation is the likeliest. If we admit that explanatory considerations are reflected in the prior and the likelihood, how they, in conjunction with other factors such as the principal principle, affect those quantities will be *a posteriori* and context dependent. Perhaps in some cases a confluence of factors lead to the loveliest explanation having the highest probability. But to commit to IBE, to inferring hypotheses that provide the best explanations of the data, is to commit to a rule that we should not, by Bayesian lights, follow in many cases. For, as I've tried to show, the explanatory considerations do not behave in a way that ensures the loveliest explanation will be the likeliest.

Second, a *modus tollens* on the conditional "if subjective Bayesianism is right, then IBE is wrong" is unwarranted. What exactly IBE is supposed to be is not settled yet. As we have seen, when left vague, IBE seems sensible. That sensibility is under pressure when scrutinized. The proposal on the table is to understand explanatory considerations as *a priori* constraints on P(.), such that $P(H_2|e \wedge K)>P(H_1|e \wedge K)$ where H_2 is a better explanation of e than H_1 and K is background knowledge. Weisberg: "Regrettably, I have no detailed proposals to offer in this regard." (137). Proponents of IBE lack a full account of what determines the set of potential explanations and what determines that one explanation is better than another. On this proposal, they also lack an account of how explanatory considerations function as *a priori* constraints on

probability assignments.[12] Earlier, a minimal account of explanatory quality was proposed to aid in evaluating the credentials of IBE. The same difficulty with evaluation arises here because of the proposal's lack of content. I contend that, until a plausible account is provided, rejecting subjective Bayesianism is not in order.

Make a Normative/Descriptive Distinction. I have only shown that in some subset of cases, explanatory considerations do not behave as they must for (2') to be true and FA sound. I have not shown that in all, or even most, cases (2') is not true. If IBE delivers the wrong answer in some cases, however, it is not normatively correct. To deliver the *wrong* answer, according to Bayesianism, means incoherence, for Bayesian rationality constraints are coherence constraints. (2') has to be true in all cases to insure that IBE does not deliver the wrong answer. Nevertheless, there are responses open to proponents of IBE. First, one might maintain that IBE does have some normative force to the extent that it tracks Bayesianism, the normatively correct account. On this picture, IBE approximates the strictures of Bayesian conditionalizing. IBE may deliver the wrong result occasionally, but so long as it is reliably correlated with Bayesianism, IBE is a useful heuristic.[13] (Lipton seems to oscillate between the earlier approach and this one).

If the results of following IBE were reliably correlated with the results of Bayesian conditionalizing, then IBE might prove to be very useful. Unfortunately, as Wesley Salmon writes about explanation, "the normative connection to inference or confirmation has not been

---

[12] Weisberg argues that reconciling IBE with subjective Bayesianism infects IBE with the bad things about subjective Bayesianism. This is not clear evidence for going the objectivist route, however, for the problems with objective Bayesianism are formidable. For example, the principle of indifference (assigning equal probability to all hypotheses in a set of mutually exclusive hypotheses when we have no evidence to otherwise distinguish among them), a lynchpin of objective Bayesianism, is all but hopeless thanks to the reference class problem. Weisberg offers a suggestion that explanatory considerations can help us to cope with the reference class problem. His suggestion, he admits, is awaiting further development (p. 141).

[13] A Kahneman and Tversky study (1983) purports to show that people use heuristics that lead them astray of correct probabilistic reasoning. 85% of participants committed the "conjunction fallacy," assigning a higher probability to A&B than to A or to B, but in a way that is consistent with the spirit of IBE.

made; it has not been shown that the best potential explanation is most likely true – or most probably true – explanation" (2001b, 131). Reliable correlation stands in need of demonstration. We need a demonstration that hypotheses that satisfy the explanatory virtues *generally* are more probable. I submit that the preceding section gives reason to think that a demonstration is not forthcoming, and such an assumption is not warranted. Moreover, on this approach, the proponents of IBE must consider whether they are content with IBE being nothing more than Bayesianism's lackey (this may not be a healthy friendship for IBE).

A second response for proponents of IBE is to insist that IBE is a description of scientific inference, and is not a normative account (whether it approximates Bayesian conditionalizing or not). I am simply interested in the normative question. If proponents of IBE advocate IBE solely as a description of scientific inference, it is not clear that van Fraassen's argument should worry them. There are, nonetheless, problems for IBE as a description of scientific practice. In DOC, if the best explanation of the observational data were an ailment the treatment for which required a surgery with a very high risk of death, the doctor would seek to *confirm* that hypothesis. She would perform tests to make sure the hypothesis' predictions were born out. She would not act on the diagnosis merely because it *explained* the data the best.

At the end of his essay "Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation," Salmon discusses a very popular episode from the history of science, and makes a similar point. In order to explain the perturbations in the orbit of Uranus, scientists postulated another planet's existence. This explanation was the best available, but the scientific community withheld inference until sufficient confirmation took place. It took telescopic observation for the scientific community to infer Neptune's existence. In "Reflections of a Bashful Bayesian," Salmon continues his critique of IBE and asks, "Would it not be indulging in

wishful thinking to suppose it likely to be true just because it would be so nice if it were true?" (Salmon 2001b, p. 130). This is precisely what the scientific community does *not* do in the above episode. There is no supposition of truth of a hypothesis due to its explanatory power. The Neptune hypothesis was valuable and worth pursuing because of its explanatory power and informational content, but it took confirmation for the scientific community to infer it.

6. CONCLUSION

IBE is intuitive, no doubt, but if we look closely at the assumptions it makes about the probabilistic behavior of explanatory virtues, IBE looks much less plausible. Okasha, Lipton, and cohort are right that van Fraassen's Dutch book argument does not apply to *any* probabilistic version of IBE. They are wrong that we can preserve a substantive version of IBE because explanatory considerations enter into Bayes' theorem and fix the prior and the likelihood such that the loveliest explanation is *ipso facto* the likeliest. FA is unsound because it is wrong about the role of explanatory considerations in probability updating. Alternative attempts to reconcile Bayesianism and IBE do not look very promising.

I do not deny that we value explanation. I think explanatory power or informativeness is a desideratum for our scientific theories, but it is one we must balance with the desideratum that our theories be probable or confirmed, because informativeness and high probability can pull in different directions. We should not accept a rule that commits us to ignoring this fact. The prospects for a healthy friendship between Bayesianism and IBE are dim.

BIBLIOGRAPHY

Douven, Igor. (1999) "Inference to the Best Explanation Made Coherent," *Philosophy of Science*, 66 (Proceedings), pp. S424 – S435.

Douven, Igor. (2008) “Assertion, Moore, and Bayes,” *Philosophical Studies*, forthcoming.

Ganson, Dorit. (2007) “Van Fraassen’s Dutch Book Argument Against Explanationism” in *Causation and Explanation*, pp. 171-190, eds. J. Campbell, M. O’Rourke, and H. Silverstein.

Glass, David. (2007) “Coherence Measures and Inference to the Best Explanation,” *Synthese* 157:3, pp. 275-296.

Hajek, Al. (2008) “Dutch Book Arguments” in *The Oxford Handbook of Rational and Social Choice*, ed. Paul Anand, Prasanta Pattanaik, and Clemens Puppe. Oxford University Press.

Harman, Gilbert. (1965) "The Inference to the Best Explanation," *The Philosophical Review*, 74:1, pp. 88-95.

Kahneman, Daniel and Tversky, Amos. (1983) “Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment,” *Psychological Review*. Vol 90:4 293-315.

Kitcher, Philip. (1989) “Explanatory Unification and the Causal Structure of the World,” in *Scientific Explanation*, eds. P. Kitcher and W. Salmon, *Minnesota Studies in Philosophy of Science*, Vol. 13, pp. 410-505, Minneapolis, University of Minnesota Press.

Lewis, David. (1986) “The Pragmatics of Explanation,” *Philosophical Papers*, Vol. 2 (Oxford: OU press) in *Philosophy of Science, Vol. 1: Explanation, Law, and Cause*, ed. Sklar (2000).

Lipton, Peter. (2001) "Is Explanation a Guide to Inference?," *Explanation*: *Theoretical Approaches and Applications*, pp. 93-120, eds. G. Hon and S.S. Rakover.

Lipton, Peter. (2004) *Inference to the Best Explanation, 2nd Edition*, Routledge.

Okasha, Samir. (2000) "Van Fraassen's Critique of Inference to the Best Explanation," *Studies in History and Philosophy of Science*, 31:4 691-710.

Psillos, Stathis. (2003) “Inference to the Best Explanation and Bayesianism,” *Institue of Vienna Circle Year Book*, Vol. 10, ed. Fritz Stadler.

Railton, Peter. (1978) “A Deductive-Nomological Model of Probabilistic Explanation,” *Philosophy of Science* 45, pp. 206-226, in *Philosophy of Science: The Central Issues*, eds. M. Curd and J.A. Cover (1998).

Salmon, Wesley. (1990) *4 Decades of Scientific Explanation*, University of Minnesota Press, Minneapolis

Salmon, Wesley. (2001a.) “Explanation and Confirmation: A Bayesian Critique of Inference to the Best Explanation,” in *Explanation: Theoretical Approaches and Applications*, pp. 61-92, eds. G. Hon and S.S. Rakover.

Salmon, Wesley. (2001b.) “Reflections of a Bashful Bayesian: A Reply to Peter Lipton” *Explanation: Theoretical Approaches and Applications*, pp. 121-136, eds. G. Hon and S.S. Rakover.

Sober, Elliott. (1990) “Let’s Razor Ockham’s Razor” in *Explanation and Its Limits*, pp. 73-93, ed. Dudley Knowles. Cambridge University Press.

Sober, Elliott. (2002) “Bayesianism—Its Scope and Limits”, in *Bayes's Theorem*, pp. 21-38, ed. Richard Swinburne, Oxford University Press.

Van Fraassen, Bas. (1977) “The Pragmatics of Explanation,” in *American Philosophical Quarterly*, 14:2 in *Philosophy of Science, Vol. 1: Explanation, Law, and Cause*, ed. Sklar (2000).

Van Fraassen, Bas. (1983) “Glymour on Evidence and Explanation,” *Minnesota Studies in the Philosophy of Science*, X, pp. 165-176 in *Theory, Evidence and Explanation*, (1995) ed. P. Lipton, Dartmouth, Brookfield.

Van Fraassen, Bas. (1989) *Laws and Symmetry*, Clarendon Press, Oxford.

Van Fraassen, Bas, Ladyman, James, Douven, Igor, and Horsten, Leon. (1997) "A Defence of van Fraassen's Critique of Abductive Reasoning: Reply to Psillos", *Philosophical Quarterly* 47, pp. 305-321.

Van Fraassen, Bas. (2005), Review of Lipton, “Wouldn’t It Be Lovely? Explanation and Scientific Realism,” Review Symposium for Lipton, *Inference to the Best Explanation*, 2nd ed. *Metascience* 14, pp. 344-352.

Van Fraassen, Bas. (2007) “REPLY: From a View of Science to a New Empiricism” in *Images of Empiricism: Essays on Science and Stances, with a Reply from Bas van Fraassen*, ed. Bradley Monton. Oxford University Press.

Weisberg, Jonathan. (2009) "Locating IBE in the Bayesian Framework," *Synthese*, 167: 125-143