Philosophy of Science

Inference to the Best Explanation and the Importance of Peculiarly Explanatory Virtues

--Manuscript Draft--

Manuscript Number:	11397
Full Title:	Inference to the Best Explanation and the Importance of Peculiarly Explanatory Virtues
Article Type:	PSA 2012 Contributed Paper
Keywords:	Inference to the Best Explanation; invariance; Woodward; Lipton
Corresponding Author:	David Harker, Ph.D. Jonesborough, TN UNITED STATES
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	David Harker, Ph.D.
First Author Secondary Information:	
Order of Authors:	David Harker, Ph.D.
Order of Authors Secondary Information:	
Abstract:	Inference to the best explanation has at times appeared almost indistinguishable from a rule that recommends simply that we should infer the hypothesis which is most plausible given available evidence. In this paper I argue that avoiding this collapse requires the identification of peculiarly explanatory virtues and consider Woodward's concept of invariance as an example of such a virtue. An additional benefit of augmenting IBE with Woodward's model of causal explanation is also suggested.

Inference to the Best Explanation and the Importance of Peculiarly Explanatory Virtues

Abstract

Inference to the best explanation has at times appeared almost indistinguishable from a rule that recommends simply that we should infer the hypothesis which is most plausible given available evidence. In this paper I argue that avoiding this collapse requires the identification of peculiarly explanatory virtues and consider Woodward's concept of *invariance* as an example of such a virtue. An additional benefit of augmenting IBE with Woodward's model of causal explanation is also suggested.

1. Inference to the Best Explanation and the Threat of Vacuity

To illustrate the advantage of 'inference to the best explanation' (henceforth, IBE) over enumerative induction, Harman (1965, 90-1) invites us to consider inferences from samples to populations and the question of "when a person is and when he is not warranted in making the inference from "All observed A's are B's" to "All A's are B's"". Harman continues:

> The answer is that one is warranted in making this inference whenever the hypothesis that all A's are B's is (in the light of all the evidence) a better, simpler, more plausible (and so forth) hypothesis than is the hypothesis, say, that someone is biasing the observed sample in order to make us think that all A's are B's.

Clearly we can posit various reasons for why all the observed A's were also B's. It might be that "All A's are B's"; someone could have purposefully manipulated the sample to deceive us; perhaps our method for selecting subjects ensures, or makes it likely that, we will observe only those A's that are also B's, and so on. Furthermore, and equally patently, the *actual* reason for the observed regularity will be different in different cases. We observe only male drones, because all drones are male. Water that's pumped through an effective filter will contain no contaminants above a certain size; the absence of contaminants from the *original* water supply, however, often will *not* be the reason that the filtered water is pure. Harman supposes that such reasons can function as explanations. Let's concede that for now. Faced with competing explanations for an observed regularity Harman urges us to infer to the truth (or approximate truth) of whichever explanation is best.

Harman's proposal is thoroughly sensible – we should infer that hypothesis which is "better" and "more plausible".[1] However, without some guidance concerning how we *identify* the best, from competing explanations, and Harman has named a problem but not solved it. Insofar as IBE is regarded as a substantive theory of confirmation, its advocates can't rest content with an interpretation that advises only to infer that conclusion which is most plausible. Seemingly though Harman's phrase is sufficiently seductive, and has become sufficiently well-entrenched, that it is now hard to appreciate how vacuous the advice really is. Had Harman suggested we infer 'that hypothesis which seems most plausible in light of all available evidence', the attenuated condition of the suggestion would perhaps be more immediately apparent. If inferring to the best explanation is different, for Harman, it's hard to see how. On inspection, inference to the best explanation can appear quite insipid.

Lipton (2004), cognizant of the problem, offers a general means of responding. Unfortunately his development of that response opens him to critical objections, or so I'll argue in Section 2. The problems with Lipton's response trace to a failure to identify *explanatory* virtues, as distinct from virtues of the hypotheses that feature in the explanation. This diagnosis leaves room for a successful defense of IBE that utilizes Lipton's general strategy, but insists on peculiarly explanatory virtues, burdening advocates for IBE with the task of identifying such. Turning to the work of Woodward

[1] Harman does, in addition, suggest that better explanations are simpler, less ad hoc, and explain more. However, these concepts are insufficiently well-defined to provide helpful guidance in the face of competing explanations.

(for example, Woodward (2003)), I'll argue in Section 3 that distinctive explanatory virtues are apparent within the sciences and, furthermore, that it is not implausible to suggest that these reliably guide theory choice. Part of Woodward's project involves discriminating descriptions from explanations. An implication of this distinction is that Harman's example, above, might fall outside the scope of IBE, a possibility I discuss and welcome in Section 4. The purpose of the paper is not a complete defense of explanatory reasoning, but an attempt to motivate two important pieces of the groundwork: first, to urge that IBE requires the identification of explanatory virtues, and can't rely on the theoretical virtues of those hypotheses that are centrally involved in an explanation; second, to suggest that IBE has a limited scope, for purposes of understanding ampliative reasoning, which we might move some ways towards delineating by distinguishing descriptions from explanations.

2. Loveliness, Likeliness, Matching, Guiding

Concerned that IBE avoid appearing trite, Lipton responds in part by distinguishing two senses of 'best explanation'. The likeliest explanation, for Lipton, is that which is most likely to be correct. Informed that two theories each explain some phenomenon, we establish the likeliest explanation by evaluating which theory is best supported by available evidence. To infer to the likeliest explanation we needn't attend to anything about the *explanations themselves*; it is the well confirmedness of the respective theories that matters. The loveliest explanation, in contrast, can't be determined by attending to the merits of the underlying theory. Lipton suggests that the loveliest explanation "provides the most understanding". White (2005), endorsing Lipton's distinction,

suggests that explanations are often valued for "the degree of *satisfaction*" they deliver; explanations might disappoint because they are implausible, but also and alternatively because they can be "deeply unsatisfying". Having made this conceptual distinction, Lipton and White each suggest that IBE is a potentially important tool for investigating inductive reasoning, because explanatory *loveliness* might prove a reliable *guide* to explanatory likeliness. If this connection between loveliness and likeliness is real, we could justifiably appeal to the loveliness of an explanation for purposes of defending conclusions about which theory or hypothesis is most plausible, at least in some circumstances.

One concern with the proposal, as described, is that the concepts of understanding and satisfaction threaten to introduce a worryingly subjective dimension. What helps one person understand some phenomenon might differ from what helps another; explanations satisfy some folks, but not others. Judgments about differences in explanatory quality that ride on these kinds of consideration are unreliable markers of underlying plausibility. Lipton at least is careful to distance himself from overly psychological interpretations of the relevant concepts, but we can avoid such connotations altogether since the basic distinction suffices. Explanations can be evaluated in terms of the plausibility of the theory that motivates them, or in terms of features that are peculiar to explanations and independent of associated theories. In what follows I'll use the phrase 'explanatory virtue' to denote the latter. IBE avoids the charge of triviality by distinguishing explanatory virtues from the overall merits of a theory, and defining the rule as an inference based on the former; the plausibility of the rule, at least if it's understood

normatively, hinges on whether explanatory virtues reliably guide us towards a proper evaluation of available theories.

In furtherance of his claim that explanatory virtues need not be subjective, Lipton suggests simplicity, provision of mechanisms, scope, precision, among others, as appropriate measures of explanatory loveliness. None are unproblematic concepts, as Lipton concedes. Nevertheless, attaching loveliness here helps remove any lingering specter of subjectivity. Barnes (1995) protests, however, that these are not reliable guides to underlying plausibility. Suppose we have two competing explanations, but only one provides a mechanism. Whether we prefer the mechanistic explanation depends on the *independent plausibility* of the mechanism, suggests Barnes, rather than any intrinsic value in describing mechanisms. Lipton offers no obvious means of evaluating mechanistic hypotheses, but providing them can't be a reliable means of *improving* an explanation, or choosing between competing explanations, because even contrived and outrageous suggestions about the underlying mechanism describe a mechanism. Barnes raises similar complaints against the other putative explanatory virtues that Lipton describes.

Against the first edition of Lipton's book Barnes objections seem pertinent. Lipton (1991) asserts that "mechanism and precision are explanatory virtues" (118), "unification makes for lovely explanations" (119) and suggests that elegance and simplicity are also qualities of explanatory loveliness (68). He further argues that by attending to these qualities we are typically, reliably directed to the most plausible hypothesis. Lipton is unfortunately silent, however, on the issue of how we should *balance* the pursuit of these various

virtues, which might pull in opposing directions. If each virtue is evaluated in isolation, then Barnes objections are critical: discriminating purely on the basis of the presence or absence of a mechanism, for example, will often warrant an implausible inference. If, on the other hand, Lipton intends us to weigh all explanatory virtues and reach an appropriate balance between them, then his failure to describe how this should be conducted leaves the account disconcertingly obscure. Lipton's earlier defense is either reasonably transparent, but implausible, or quite opaque. However, Lipton's defense shifts between the two editions of his book. In the more recent he argues explicitly for a *correspondence* between theoretical and explanatory virtues, then argues independently, and on empirical grounds, that we *in fact* use the latter to evaluate the former. What is discussed as "matching" and "guiding" in the later edition are not distinguished in the earlier. Lipton hereby implies that the likeliest and loveliest explanations will each provide the best balance of various virtues, although again Lipton provides no guidance on how we are to recognize the best trade-off. Given Lipton's new strategy it becomes hard to accuse him of proposing an unreliable rule of inference, since it's a rule that by definition should guide us towards that conclusion which best instantiates all those theoretical virtues that are typically assumed important. The problems with Lipton's new strategy lie elsewhere.

One prominent theme in Lipton's book is that IBE describes our inferential practices better than alternative accounts. Lipton claims such advantages over Bayesianism, hypothetico-deductivism and Mill's methods of causal reasoning. Deficiencies with each, in terms of how well they describe our inferential practices, suggest either their

replacement with IBE or, in the case of Bayesianism, augmentation with explanatory considerations. These comparative claims have been challenged. Rappaport (1996) defends Mill's methods against Lipton's concerns. Bird (2007) argues that Lipton's objections are largely ineffective against hypothetico-deductivism. Douven (2005) argues that Lipton says too little about how and why Bayesians should build explanatory considerations into their framework. Furthermore, even if we concede that IBE better describes our inferential tendencies, we don't thereby achieve any normative justification for explanatory reasoning. What Lipton does say about the normativity of the rule is uninspiring.

According to Lipton's matching claim, explanatory reasoning is justified since explanatory considerations direct us towards that hypothesis which is most precise, has greatest scope, and so on, which Lipton suggests render that hypothesis most probable. However, Lipton offers little by way of analysis for these theoretical virtues. Consequently, because they're notoriously vague, and because it's hard to justify why they matter for purposes of confirmation, and because we don't know how to balance these often competing qualities against one another, Lipton leaves many hostages to fortune. The justification for explanatory reasoning is entirely derivative, and it is derivative on something that's worryingly vague. There is no answer as to why we should value a rule that directs us towards the simplest hypothesis, other things being equal. However, we might reasonably expect that if a theory of *confirmation* is going to place a premium on considerations of simplicity, then it should justify that decision. Leaving so

many concepts unanalyzed might leave us again wondering whether there's any real substance to IBE.

The failure to more carefully define these concepts becomes problematic again when we turn to Lipton's guiding principle. It is suggested both that, as an empirical matter, we tend to be impressed by explanatory considerations and, when confronted with competing explanations, it is the simpler, more precise, and so on, that is inferred. However, there is no obvious reason to suppose that the sense of simplicity that I employ when making a judgment about competing explanations will be the same sense that might prove a justified means of adjudicating between competing hypotheses.[2] A normative justification for Lipton's account requires either that we offer distinct analyses of explanatory and theoretical simplicity, then argue that explanatory simplicity is a reliable guide to theoretical simplicity, or we stipulate that simplicity has the same sense in each context. The former strategy is far from straightforward. The latter makes it much more difficult to argue that we *in fact* prefer simpler explanations, in the relevant sense and other things

[2] For example, in curve-fitting problems it has been argued that introducing additional adjustable parameters is appropriate only if will improve the predictive accuracy of the curve. If we define simplicity in terms of the number of adjustable parameters, then we justify a role for simplicity within certain well-defined contexts (see Forster and Sober (1994)). However, the balance between fit and number of parameters emerges from a non-obvious mathematical theorem. It seems unlikely that any 'intuitive' sense of simplicity that we might employ in evaluating explanations should guide us towards hypotheses that are more simple in this respect.

being equal. Maintaining both the guiding principle and a normatively justified interpretation of IBE becomes less plausible.

Hopes of preserving the normative dimension of IBE are further degraded when Lipton appeals to data from cognitive psychology. For example, Lipton describes the results of work conducted by Kahneman and Tversky, which demonstrated our propensity for committing the conjunction fallacy. (Asked to identify which event was most probable, given some scenario, many subjects committed the error of supposing a conjunction of two events can be more probable than one of the conjuncts.) Lipton offers this as evidence both that we are not good at Bayesian updating and that explanatory considerations play an important role in how we reason. An obvious concern is that Lipton's interpretation of the result provides an immediate example of explanatory reasoning that is unreliable. Lipton responds that in circumstances more complicated than those described by Kahneman and Tversky explanatory reasoning might be more reliable, but offers no evidence to support the conjecture.

In summary, Lipton argues that explanatory loveliness is both a reliable guide to explanatory likeliness, because considerations like simplicity and scope are features of more probable hypotheses and more virtuous explanations, and an important aspect of our inferential practices. However, the connections between these theoretical virtues and the plausibility of a given hypothesis are sufficiently vague that it is hard to admit them into a theory of confirmation as brute facts. The argument also requires us to concede that our natural proclivities, when evaluating explanations, will draw on similar considerations to those that will ultimately be deemed important for evaluating hypotheses, and that we

apply them in similar ways. Finally, in light of our demonstrated cognitive failures where we are perhaps unduly influenced by explanatory considerations, we must hope for evidence that such failures are heavily restricted to certain kinds of case. Absent such evidence and, although we might have reason to suppose we in fact employ explanatory reasoning, we'd lack any reason to suppose that we should. The normative dimension of IBE, as developed by Lipton, is both vague and tenuous. Admittedly Lipton at times seems content with defending a purely descriptive interpretation of IBE, in which we declare only that explanatory considerations in fact feature prominently in our reasoning. Typically IBE is understood as a normative thesis; a purely descriptive thesis certainly falls short of my ambitions for the rule.

Where did Lipton go wrong? I suggest it's in arguing that explanatory and theoretical virtues align. By adopting that position it becomes hard for explanatory considerations to illuminate, account for, or justify judgments about which of competing hypotheses is most plausible. The promise of IBE, as initially presented by Lipton, was with the idea that we could read off qualities of an explanation and thereby learn something important about the merits of the underlying theory. Given the matching claim, any normative justification for IBE becomes fully dependent on concepts that are not only problematic and vague, but also appear independent of explanatory considerations. Consequently, Lipton is forced to adopt an essentially descriptive interpretation of the rule. A model of IBE would be more useful and more interesting if we could identify peculiarly explanatory virtues, that cannot be identified with qualities of the underlying hypotheses, and that help us understand why certain inferences are sensible. Developed in this way

and IBE could live up to its reputation as a theory of how we should reason. Utilizing Woodward's model of causal explanation I'll now sketch a way of relating explanatory considerations to underlying plausibility that seems promising.

3. Invariance, Mechanisms and Consilience

Woodward's model is centrally concerned with change relating regularities, regularities that describe how changes in the value of one variable affect the value of another. Interventions on variables pick out causal and explanatory relations, for Woodward, if they are a reliable means of manipulating other variables within the regularity. Many regularities will satisfy this standard under some conditions but not others. For example, the ideal gas law properly captures our ability to increase the temperature of a gas by increasing the pressure, in certain circumstances. The law is thus a change-relating regularity that describes a causal relation, exploitable for purposes of explaining. The law doesn't hold universally, however. When temperatures become sufficiently low, or pressures sufficiently high, the law no longer accurately describes the relation between these variables. In such conditions we might appeal to the van der waals equation, which holds in circumstances where the ideal gas law breaks down. For Woodward, the latter is more *invariant.* Regularities are invariant if they continue to hold despite interventions on the variables that feature in that regularity. We explain an outcome by appealing to a system of regularities that is invariant under at least some interventions, and which can be combined with a range of possible initial and boundary conditions to describe how events would have differed had those conditions been otherwise. Only regularities that are invariant under some interventions are explanatory. Regularities that are *more* invariant

support a broader range of explanations, since they allow us to say more about how things would have been different if initial or background conditions were different.

Although Woodward isn't concerned with the relationship between invariance and confirmation, and even expresses some skepticism about inference to the best explanation (see note 5), I suspect there are important connections. My proposal is that it is reasonable to infer more invariant explanations, over less invariant explanations, because considerations of invariance tell us something important about the regularities that ground the explanations. My suggestion is that pursuing greater invariance will tend to produce the kinds of achievements that scientists consider epistemically significant, including our admiration for verified novel predictions, predictive success more generally, and high precision testing, our suspicion of ad hoc hypotheses, desire for both `deeper' explanations and explanations of `free parameters', as well as our pursuit of theories that have greater consilience. Despite their reputations, these concepts are poorly understood. The concept of invariance, insofar as it can illuminate these more familiar concepts, advances our understanding of confirmation.

Before offering some details, a few preliminaries are in order. First, invariance is distinct from predictive success, consilience, scope, and so on. The proposal thus shares with Lipton's defense a distinction between two types of explanatory achievement. We can evaluate an explanation in terms of its invariance, where more invariant explanations are better. Explanatory hypotheses and regularities can also be better insofar as they are less ad hoc, more precise, verified by novel predictions, and so on. If invoking the concept of invariance offers more plausible analyses for the confirmatory significance of such

considerations, then it has importance for our understanding of confirmation as well as explanation. What distinguishes my proposal from Lipton's more recent defense is that invariance is a peculiarly explanatory virtue, rather than a feature of the underlying theory or hypothesis. This creates room for a normative defense of explanatory reasoning. It is also important to distinguish a more modest from a more ambitious version of the thesis I'm proposing. The more modest rests content with providing a better account for extant confirmatory considerations. The more ambitious version assumes, or argues, that those concepts are in turn indicative of more general forms of scientific achievement. If pursuing invariance helps us achieve deeper explanations, for example, and deeper explanations indicate a more truthlike theory, then we connect a distinctively explanatory virtue to perhaps the ultimate scientific achievement. Admittedly concepts like consilience and ad hoc-ness are only poorly understood, thus difficult concepts to offer in defense of realist commitments. However, insofar as IBE might help provide more convincing analyses for various intuitions surrounding questions of confirmation, once augmented with Woodward's concept of invariance, it can simultaneously help justify its own normative credentials. It's beyond the scope of this paper to start properly exploring the connections between invariance and all the concepts I've alluded to. Hopefully a couple of examples will provide adequate motivation for the thesis.

First, let's return to Lipton's desire for mechanistic explanations and Barnes' concern that merely adding a mechanism can't itself reliably improve an explanation. The concept of invariance enables us to distinguish mechanisms that improve our explanations from those that don't. Drawing on Woodward's example, the amount of pressure applied to the

gas pedal explains the speed of my car, at least under some conditions. This change-relating regularity can be exploited for purposes of manipulating the speed of the car, and therefore for purposes of explaining the speed, even for those of us who are ignorant about *how* changing the pressure applied to the pedal brings about the change in speed. Providing a mechanism that relates these variables will not always produce a better explanation: fanciful mechanisms that have no grounding in experience describe mechanisms. Mechanisms which are more *invariant* than the crude regularity we begin with increase our ability to manipulate and control the speed of the car under a wider range of conditions. We improve our understanding of the counterfactual dependencies that describe the system. Providing a mechanism that relates distinct variables will *improve* an explanation only if it is more invariant than the regularity alone.

Providing mechanisms for causal regularities is an important scientific pursuit. Thoroughly speculative mechanisms, however, are not valued, requiring us to find means of distinguishing speculative from plausible mechanisms. The concept of invariance achieves that. Furthermore, it's at least plausible to suppose that this improved ability to manipulate a system reflects a better understanding for how a given system behaves.[3]

[3] Several authors have suggested that IBE has importance for purposes of fixing prior probabilities, likelihoods, or both, within Bayes' equation (for example, Lipton (2004), Okasha (2000), Weisberg (2009)). The rule is thus given a probabilistic interpretation. Elsewhere I've argued that advocates for this approach are vulnerable to a critical

As a second illustration, again inspired by Woodward (2003, 261-2), consider the puzzle of distinguishing consilience from conjunction. Conjoining two theories produces a new, more general theory. However, explaining events by appealing to a conjunction is no improvement over an explanation that appeals to the relevant conjunct. Conjoining Hooke's law with the ideal gas law doesn't improve our explanations for the temperature of a given gas, even though the conjunction is more general. Theories are, however, lauded for their consilience. Newton's theory of universal gravitation offered explanations for falling bodies, planetary motions and tidal effects via a unified system. Consilience involves more than just conjunction, but identifying the excess has proved problematic. Again the concept of invariance is edifying. Conjunctions provide no *additional* information about the effects of intervening on variables, beyond what's provided by one of the conjuncts in isolation. Frequently cited cases of consilience, in contrast, do provide additional information. Galileo offered explanations for bodies falling near the Earth's surface. Newton also offered explanations for bodies falling near the surface of Earth (or any other massive object), but his were invariant under changes to the mass and radius of the body on which the objects are dropped. Newton's explanations are invariant in ways that Galileo's are not. The concept of invariance accounts for the differing attitudes towards conjunction and consilience.

The concept of invariance promises valuable analyses of various confirmatory concepts. A convincing defense of this claim requires both a more careful explication of the two

---

dilemma and that IBE should instead be understood as a guide to better representations of target systems (see author).

concepts already presented, and their relation to invariance, and extended discussions of the other concepts I've alluded to. A satisfactory treatment lies beyond the scope of this paper, but hopefully I've done enough to at least induce some goodwill for the idea. Rather than develop this aspect of the project further, in the following section I'll explore an independent reason to regard Woodward's theory as a helpful crutch for IBE.

4. Descriptions, Explanations and IBE's Scope

For Woodward, explaining involves communicating relations of counterfactual dependence. Regularities that don't capture such relations can't be utilized for purposes of explaining, although they might provide useful and accurate *descriptions* of target populations. For example, "All swans are white" cannot explain why a particular swan is white, since it doesn't provide the kind of dependency to which Woodward attaches significance. The explanatory impotence of certain regularities has an important consequence for Harman's puzzle, described above. Concerned to identify those circumstances when it is appropriate to infer 'All A's are B's' given that 'All observed A's are B's', Harman suggests the inference is justified if the former provides the best explanation for the latter. If the regularity is not change relating however, then it doesn't explain at all, at least according to Woodward.

IBE is understood differently by different authors. One disagreement concerns the rule's scope. Harman (1965) and Psillos (2002) suggest the rule is more general than inductive reasoning; Lipton (2004) describes IBE instead as one important type of non-deductive reasoning. I favour Lipton's more modest attitude; some of the considerations that persuade me will be presented below. Adopting Lipton's position burdens one with

providing criteria for when IBE can, and cannot, be employed, and an intriguing platform for that project is precisely the distinction between descriptions and explanations that Woodward's model of explanations articulates. Sometimes our concerns are principally with describing a process, or kind; sometimes our concerns lie with explaining why certain events occurred, or why things are configured in a particular way. Restricting explanatory inferences to those circumstances when we are actually engaged in explaining seems sensible. It also helps insulate the rule against important objections.

Consider Hitchcock's (2007) objection, in which we imagine two coins, one fair and one biased (3:1) in favour of heads. A coin is selected at random and flipped four times, where each flip lands heads. We assume a prior probability of 1/2 that we selected a particular coin, conditionalize on the new evidence, and thereby determine the posterior probabilities. We know how probable it is that we selected either coin, but Hitchcock sensibly asks what reason IBE can offer for preferring one hypothesis over the other. Relative to the evidence, neither hypothesis is simpler, more unifying nor, more generally, more lovely. Thus while the Bayesian can give clear directives concerning which hypothesis is more probable, and by how much, advocates of IBE seemingly have little to offer. Hitchcock's concern is well-directed, but might serve to motivate the delineation described above. Whether the selected coin is fair, or not, is a question about whether we have properly described the propensity of the coin. Such descriptions will align more or less probably with the outcome of subsequent sequences of flips, which are thereby entirely relevant for purposes of evaluating the plausibility of the competing *descriptions*. By restricting IBE to the evaluation of change relating regularities, however,

the example falls outside the domain of IBE. Hitchcock is thus quite correct, I'd submit: IBE has nothing to offer in terms of illuminating such cases. The lesson is not that IBE is flawed, but that it has a restricted range of application.[4]

5. Conclusions

Inference to the best explanation faces various objections and would benefit from additional work along several dimensions. Most urgent, to my mind, is that the rule distinguish itself from a recommendation simply that we infer that conclusion which is most plausible given available evidence. A second significant challenge emerges from some very sensible criticisms: explanatory considerations are not *always* relevant to inductive reasoning, so the rule must have a more limited scope than some have suggested. The challenge is to identify those circumstances when IBE helpfully and properly models good inferential habits. In Woodward's account of causal explanation I've suggested that we may have the resources both to develop a potentially instructive and plausible version of IBE, and simultaneously start to better understand its boundaries.

---

[4] Woodward (2003, 5) also expresses doubts about IBE, arguing that the distinction between explanation and description is essential to a proper understanding of scientific methodology, but that descriptions are evidently not confirmed by appeals to explanatory qualities. Clearly, however, once we rescind hopes of developing IBE into a universal model of confirmation, Woodward's concern disappears.

References

Barnes, Eric. 1995. "Inference to the Loveliest Explanation." *Synthese* 103:251-77.

Bird, Alexander. 2007. "Inference to the Only Explanation." *Philosophy and Phenomenological Research* 74:424-32.

Douven, Igor. 2005. "Wouldn't it be lovely: Explanation and Scientific Realism." *Metascience* 14:331-61.

Forster, Malcolm and Elliott Sober. 1994. "How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions." *British Journal for the Philosophy of Science* 45:1-35

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

Hitchcock, Christopher. 2007. "The Lovely and the Probable." *Philosophy and Phenomenological Research* 74:433-40.

Lipton, Peter. 1991. *Inference to the Best Explanation.* 1st edition London: Routledge.

Lipton, Peter. 2004. *Inference to the Best Explanation.* 2nd edition. London: Routledge.

Okasha, Samir. 2000. "van Fraassen's critique of inference to the best explanation." *Studies in the History and Philosophy of Science* 31:691-710.

Psillos, Stathis. 2002. "Simply the Best: A Case for Abduction." in *Computational Logic: Logic Programming and Beyond: Essays in Honour of Robert A. Kowalski.*

Rappaport, Steven. 1996. "Inference to the Best Explanation: Is It Really Different From Mill's Methods?" *Philosophy of Science* 63:65-80.

Weisberg, Jonathan. 2009. "Locating IBE in the Bayesian Framework." *Synthese* 167:125-43.

White, Roger. 2005. “Explanation as a Guide to Induction.” *Philosophers' Imprint* 5:1-29.

Woodward, James. 2003. *Making Things Happen: A Theory of Causal Explanation.* Oxford University Press.